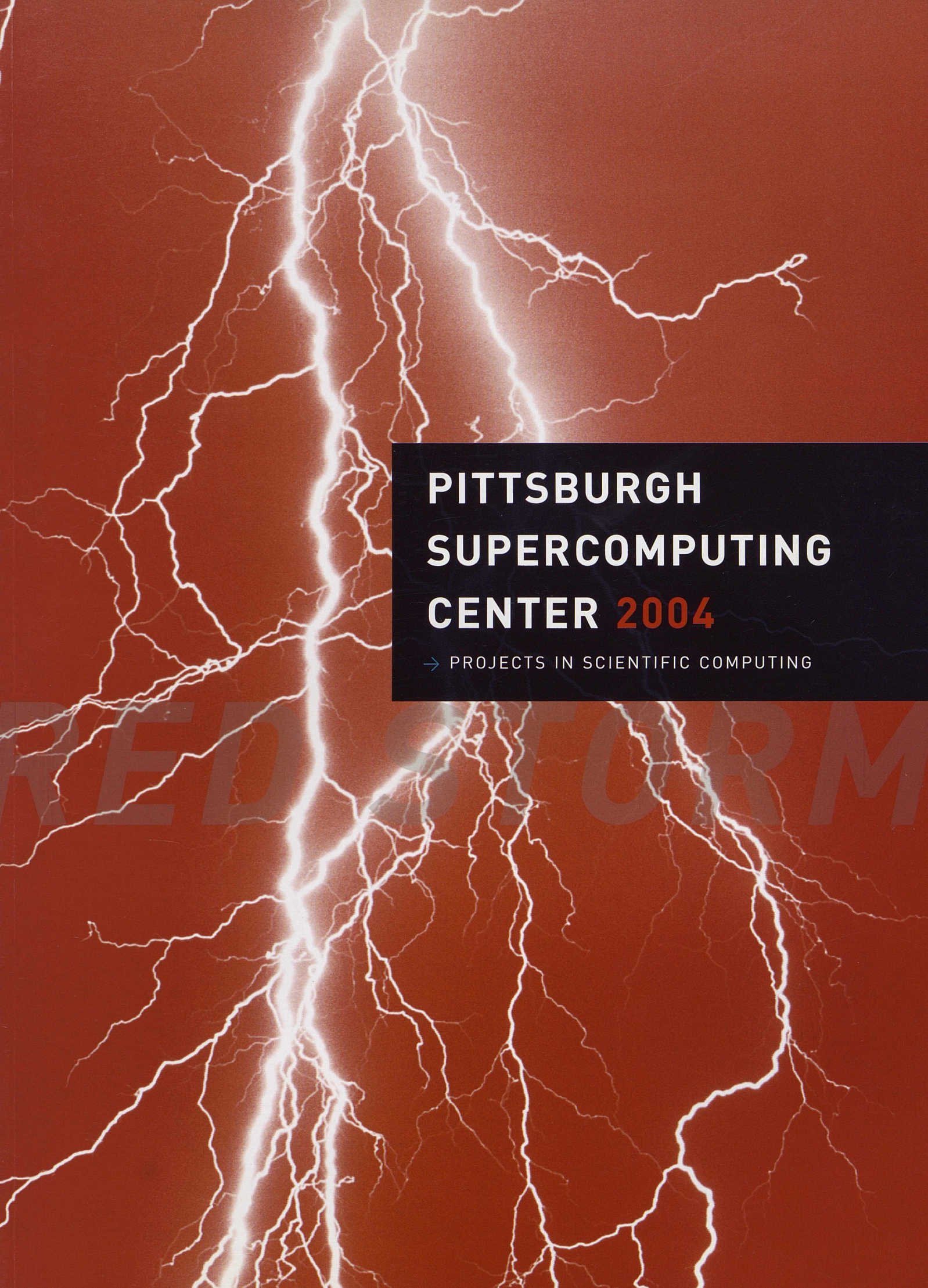
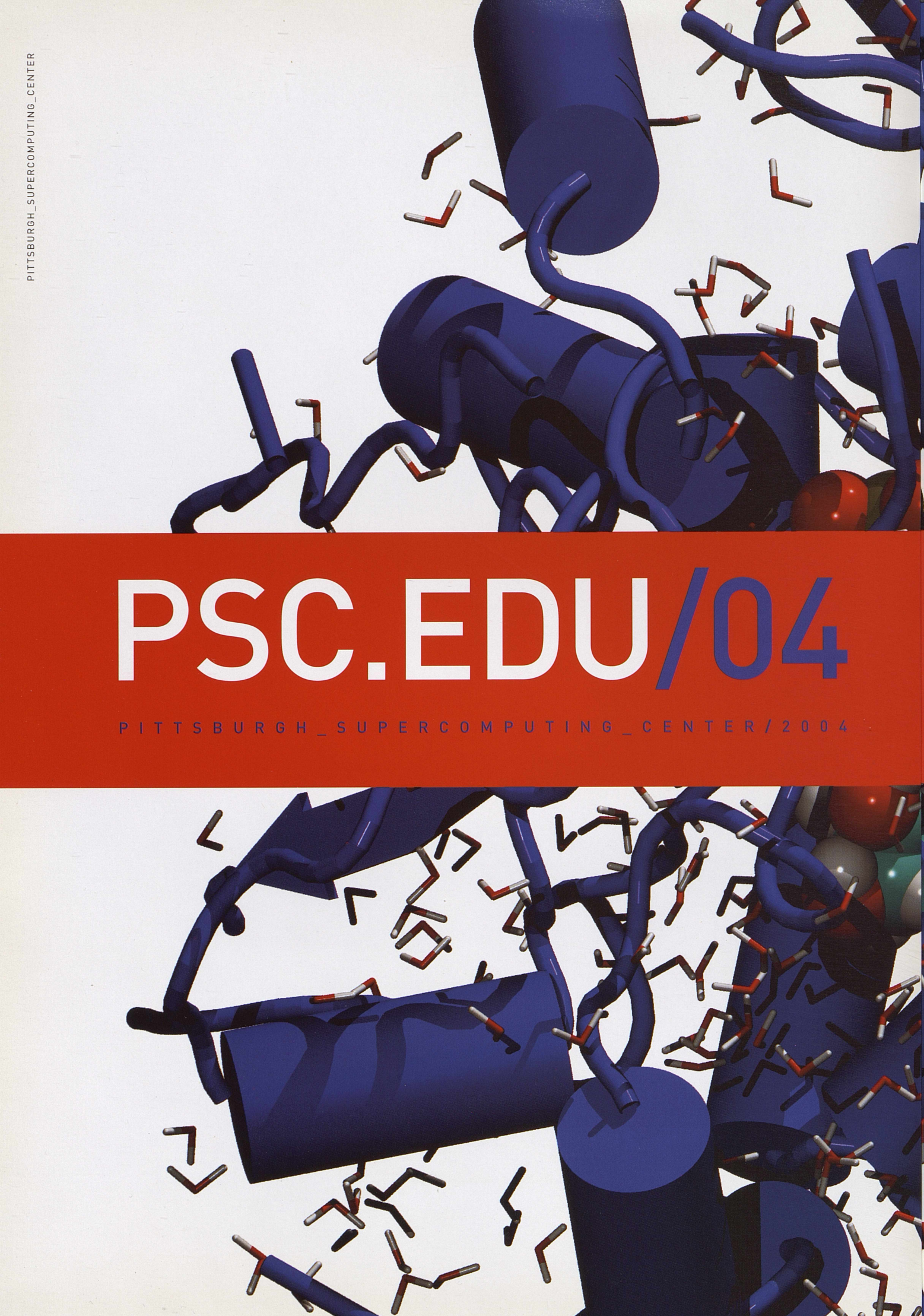

PITTSBURGH
SUPERCOMPUTING
CENTER 2004
→ PROJECTS IN SCIENTIFIC COMPUTING
RED STORM



PSC.EDU/04

PITTSBURGH_SUPERCOMPUTING_CENTER/2004

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. **WWW.PSC.EDU TEL: 412-268-4960**

FOREWORD_FROM_THE_DIRECTORS

Michael Levine (right) and **Ralph Roskies**,
scientific directors, Pittsburgh Supercomputing Center

We are very pleased to once again report on the outstanding science and engineering research enabled by the leading-edge facilities at the Pittsburgh Supercomputing Center and made accessible and available to the national community as part of the National Science Foundation's Cyberinfrastructure program. Last year, we took time to define that relatively new word. Now, cyberinfrastructure is part of the lexicon.

The TeraGrid, the NSF's multi-year effort to build a distributed national cyberinfrastructure, is now in production with a coordinated set of services for the nation's science and engineering community. The TeraGyroid project (p. 36) made early, impressive use of the TeraGrid to arrive at significant scientific results. Their work helps to unravel the complex physics of materials on the borderline between liquids and solids. PSC has played an important leadership role in TeraGrid in the development of interoperability, security and in responding to the needs of users.

LeMieux, PSC's terascale computing system, continues to be a research powerhouse, having provided 60 percent of the computing time used for NSF computational science and engineering during 2003. Added last year, PSC's HP GS1280 "Marvel" 128-processor shared memory systems have also been highly productive. Klaus Schulten and his colleagues at the University of Illinois, Urbana-Champaign, who have made notable use of LeMieux to study membrane proteins, this year used the Marvel system for a major study of a crucially important enzyme (p. 20). Their highly detailed quantum calculations revealed new understanding of ATPase, the remarkable molecular machine that synthesizes ATP, the fuel of life.

NSF foresees that an expanded leading-edge capability is required for continued scientific advance, and we are gratified that they chose PSC to house a new system, recognizing our ability to field untested systems and rapidly turn them into productive research performers. An NSF grant, announced in September (p. 4), will allow us to acquire a 10-teraflop system from Cray Inc.—based on the Sandia "Red Storm" architecture. We expect this system to become the TeraGrid resource best suited for very large-scale, demanding projects.

Among other recent scientific and technical achievements at PSC, this publication also calls attention to the work of Carlos Simmerling of SUNY Stony Brook, who used LeMieux to simulate an important protein model system (p. 24). His findings helped to fill-in details in an emerging picture of the relationships between amino acids and the folded structure of proteins.

We're especially pleased that PSC's training efforts in the expanding repertoire of supercomputing tools for biomedical research led directly to Ph.D. student Aleisha Dobbins' success at sequencing the SP6 bacteriophage (p. 28). This work contributes to a renewed interest in bacteriophages as an answer to the difficult health problem of antibiotic resistant bacteria.

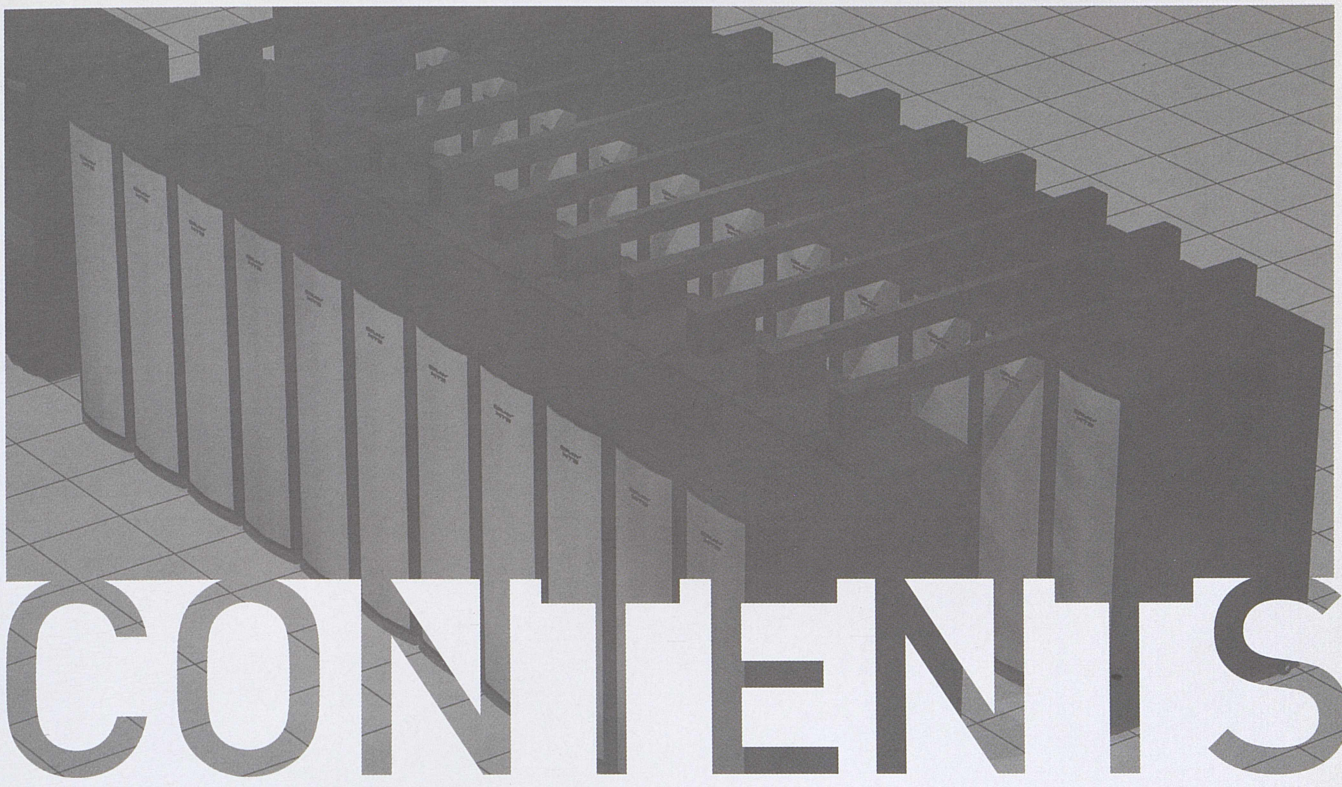
For nearly a decade, PSC has had a fruitful collaboration with The Center for Analysis and Prediction of Storms at the University of Oklahoma led by Kelvin Droegemeier. This year, Droegemeier's colleague Ming Xue used LeMieux for a watershed tornado simulation (p. 32)—the largest and most realistic ever done. This work will help to reduce the high false-alarm rate inherent in tornado warnings.

This year's publication also features the very large-scale and significant simulations of the cosmos (p. 40) carried out by Paul Bode and Jeremiah Ostriker of Princeton. Harnessing LeMieux to data from the WMAP satellite, they've produced unprecedented detail in their picture of the structure and evolution of dark matter, data that will guide upcoming astronomical observations.

All these achievements are due to the efforts of the outstanding staff of PSC and are made possible by support from the National Science Foundation, the U.S. Department of Energy, the National Center for Research Resources of the National Institutes of Health, the Commonwealth of Pennsylvania and many others.

MICHAEL J. LEVINE, SCIENTIFIC DIRECTOR

RALPH Z. ROSKIES, SCIENTIFIC DIRECTOR



CONTENTS

Red Storm Comes to Pittsburgh

PSC WILL ACQUIRE AND INSTALL THE LATEST, MOST ADVANCED CRAY SYSTEM

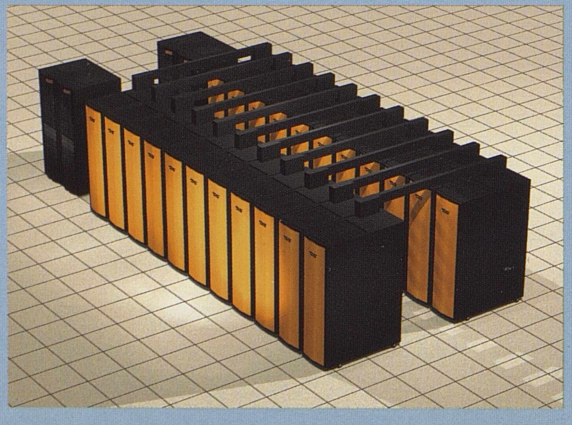
LeMieux, PSC's terascale computing system, has been an impressive performer for three years, but PSC will soon have a Red Storm with even more impressive computing credentials.

On September 29, NSF announced a $9.7 million award to PSC to acquire and install the newest high-performance system from Cray Inc. Called Red Storm by Cray, the new system has processors twice as powerful as LeMieux along with a state-of-the-art internal network that allows the processors to communicate and share data more than 10 times faster than any similar system.

PSC's Red Storm will comprise 2,000 AMD Opteron processors and have a peak performance of nearly 10 teraflops—10 trillion calculations per second. It will be the first prototype of this highly capable system to be made available to NSF scientists and engineers. PSC will evaluate this innovative architecture on representative applications, which will include blood-flow modeling, protein simulations, storm forecasting, global climate modeling, and simulations of earthquake ground vibration.

Because the network is an integral design feature of the system, Red Storm will occupy much less floor space than LeMieux, about as much as a spacious living room (12 x 28 feet) compared to the basketball-court size space used by LeMieux.

PSC's Red Storm will employ many of the same technologies as a larger, 40 teraflop Red Storm now being installed at Sandia National Laboratories in Albuquerque. Although that system has complex features required for classified research—unnecessary in PSC's open research environment—PSC and Sandia will pool their knowledge and experience with Red Storm to maximize its productivity as a scientific resource.



Jim Kasdorf, PSC director of special projects

PEOPLE ARE TALKING ABOUT RED STORM

"The Red Storm system in Pittsburgh will enable researchers to explore the limits of high-performance computing and to demonstrate the potential of this architecture for a wide range of scientific applications," says Peter Freeman, head of NSF's Computer and Information Science and Engineering directorate. "The system will complement other systems already provided by NSF to the national community and will strengthen the growing high-end computing partnership between NSF and the Department of Energy."

"We're extremely gratified to be able to introduce Red Storm for the NSF," said PSC scientific directors Michael Levine and Ralph Roskies in a joint statement. "PSC has unmatched experience in deploying new systems for the national research community. Going back to the CRAY Y-MP in 1990, we have installed over a half-dozen first and early systems of diverse architectures."

"Cray is very pleased to partner with PSC and the NSF to deliver a system, built from the ground up for high-end computing, to the broader academic research community," said Peter Ungaro, vice president of sales and marketing for Cray Inc. "Bringing the Red Storm system to PSC will provide researchers with incredibly high bandwidth and usability while leveraging the best in microprocessor technology and price/performance. We are excited to imagine how this Cray technology will be used to push the bounds of science."

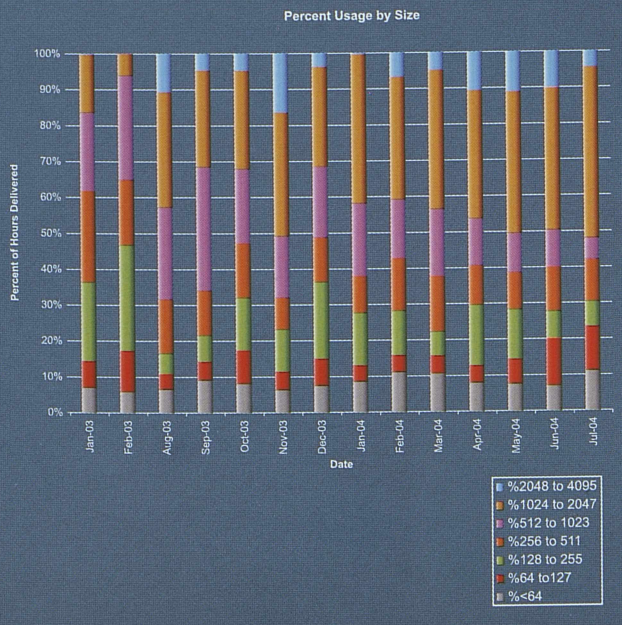
WITH THE GRID, SCIENCE HAS BECOME GLOBAL

RED STORM & LEMIEUX: STAR "SCALING" PERFORMERS

Because of its superior interprocessor communication, Red Storm will provide a powerful platform for research applications designed for efficient "scaling"—using hundreds or thousands of processors simultaneously on the same problem. It will succeed LeMieux as the prime NSF resource for the most complex, demanding projects in computational science and engineering. But Red Storm has a tough act to follow. LeMieux, NSF's first terascale system, has fulfilled this role extremely well. For 2003, it provided more than 60 percent of the computing time used for NSF science and engineering research.

Usage of LeMieux also reflects that PSC training and workshops on "scaling" have had a solid payoff. Good scaling requires careful programming, and PSC's workshops "Scaling to New Heights" and "Terascale Code Development" have trained researchers in these techniques.

LeMieux has spectacular scaling credentials—as evidenced by this graph. For the past year, more than 50 percent of LeMieux's computing hours have gone to jobs using more than 512 processors, and from May through August, to jobs using more than 1,024 processors. This shows that many PSC researchers have learned the tools of scaling. This is the most impressive use of massive parallelism among U.S. supercomputing centers.



TERAGRID GOES LIVE

Like LeMieux, PSC's Red Storm will be integrated into the TeraGrid, a multi-year NSF effort to build and deploy the world's largest, most comprehensive distributed infrastructure for open scientific research. In January the first phase of TeraGrid entered production mode and in October TeraGrid entered full production, making 40 teraflops of distributed computing power available to scientists across the country.

"We are pleased to see scientific research being conducted on the initial production TeraGrid system," said Peter Freeman, head of NSF's Computer and Information Sciences and Engineering directorate. "Leading-edge supercomputing capabilities are essential to the emerging cyberinfrastructure, and the TeraGrid represents NSF's commitment to providing high-end, innovative resources."

The TeraGrid offers storage, visualization, database, and data collection capabilities. Hardware at multiple sites is networked through a 40-gigabit per second backplane—the fastest research network on the planet. This Chicago-Los Angeles backplane links with Pittsburgh via a 30 Gbps light pipeline.

One of the most impressive feats of Grid computing to date has been the TeraGyroid project (pp. 36-39), which linked Grid resources on two continents and relied heavily on LeMieux as well as other TeraGrid sites.

"The TeraGyroid Project exemplifies what's possible with Grid technologies," said Rick Stevens of Argonne National Laboratory and the University of Chicago, TeraGrid project director. "It's a major success for the NSF vision of integrated national cyberinfrastructure, and it helps us to appreciate that—just as the economy is global—with the Grid, science too has become global."

TERAGRID SITES

Argonne National Laboratory

The Center for Advanced Computing Research at the California Institute of Technology

Indiana University

The National Center for Supercomputing Applications at the University of Illinois, Urbana-Champaign

Oak Ridge National Laboratory

Pittsburgh Supercomputing Center

Purdue University

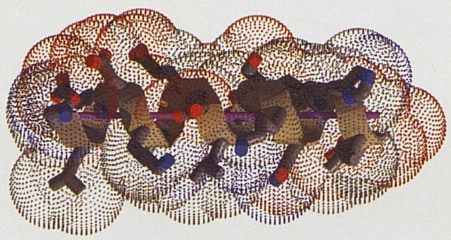
The San Diego Supercomputer Center at the University of California, San Diego

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

A COMPUTATIONAL ENVIRONMENT FOR BIOMEDICINE

ENVISIONING TECHNOLOGIES FOR BREAKTHROUGH ADVANCES IN HEALTH CARE

by David Deerfield, Michael Levine, Ralph Roskies and Joel Stiles



High-performance computing and communications (HPCC) has become increasingly central to leading-edge biomedical research. In the last decades, biomedicine has changed from a mostly data-poor, qualitative science to one that is increasingly data-rich, increasingly quantitative, and heavily influenced by high-throughput technologies. HPCC itself has also changed dramatically, from a world dominated by large purpose-built mainframes to much more powerful systems that are increasingly distributed, and are increasingly affordable at modest sizes. Adequate access to HPCC resources inspires scientists to attack problems that would otherwise be considered infeasible. High-performance computing, along with the expertise needed to utilize it, redefines the kinds of problems that scientists are willing to investigate.

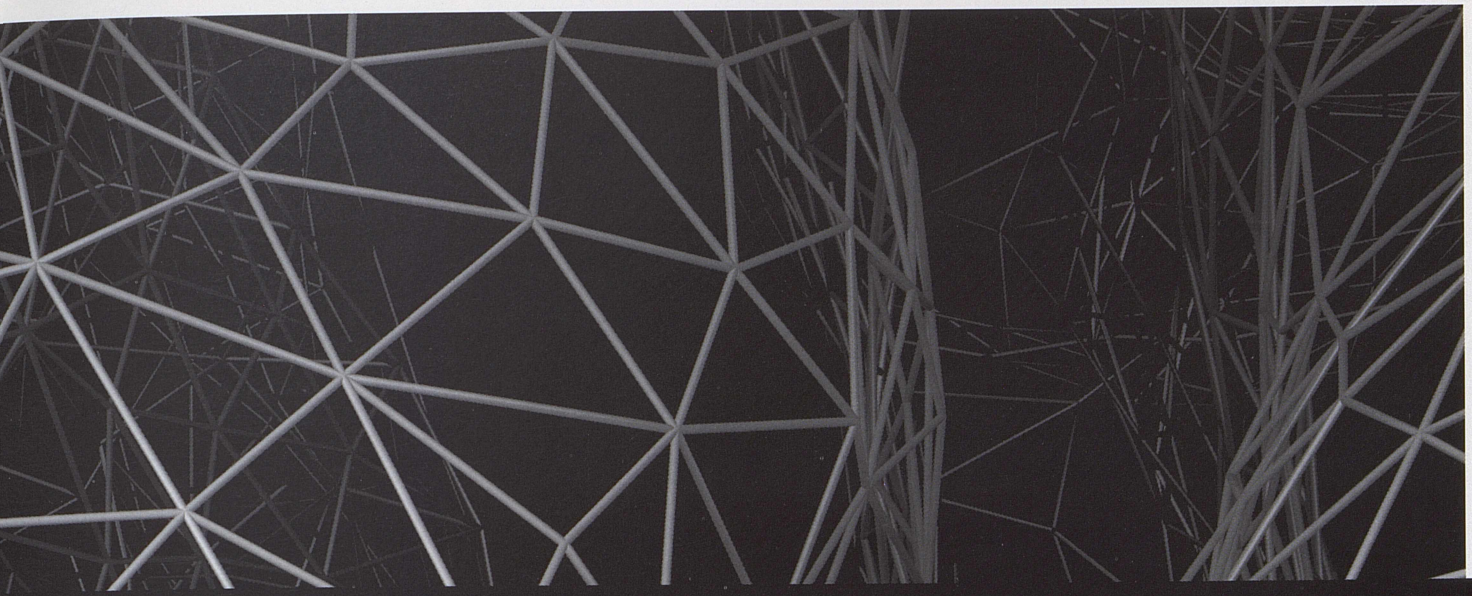
Over the past several years, leadership at the National Institutes of Health (NIH) has put a premium on HPCC and its critical role in future breakthrough advances for human health. This critical role is reflected in the NIH's recent Roadmaps, and is often motivated by the continuing explosion in range and volume of biomedical data. However, biological understanding also requires advances in modeling and simulation that will require HPCC on previously unforeseen scales.

In recent testimony before the President's Information Technology Advisory Council (PITAC), Eric Jakobsson defined a mission statement for the NIH's Roadmap on Bioinformatics and Computational Biology. Chair of NIH's Bioinformation Science and Technology Initiative Consortium and director of the National Institute of General Medical Sciences Center of Bioinformatics and Computational Biology, Jakobsson stated:

> In ten years, we want every person involved in the biomedical enterprise—basic researcher, clinical researcher, practitioner, student, teacher, policy maker—to have at their fingertips through their keyboard, instant access to all the data sources, analysis tools, modeling tools, visualization tools, and interpretative materials necessary to do their jobs, with no inefficiencies in computation or information technology being a rate-limiting step.

THE PSC BIOMEDICAL RESEARCH GROUP

The PSC biomedical group comprises 13 full-time staff, including seven PhD scientists, five master's degree programmer-analysts in computer science and engineering and a full-time administrator, along with 10 interns and four students.



While a ten-year timescale may be overly ambitious, on a longer timescale, we at PSC can envision a groundbreaking "Computational Environment for Biomedicine" (CEB) that will link all Genomics, Proteomics, and other growing databases with integrated tools for imaging, model generation, multiscale simulation and visualization, analysis, training, education, and HPCC- and grid-based computing. In addition, such a CEB will be linked to increasingly personalized information on genomic sequence, time-dependent protein expression levels, and many other quantitative measures obtained on a case-by-case basis. All of this population and individualized data coupled to a CEB will enable breakthrough clinical advances such as:

- predicting the whole body effect of a genomic mutation;
- drug design, including prediction of systemic side effects;
- halting and reversing neurodegenerative changes;
- anti-aging therapies;
- bioengineering and materials science solutions to sensory deficits, cardiovascular disease, and a host of musculoskeletal defects;
- understanding and intervening effectively in psychiatric disorders.

Of course, this will not be an easy task. In his PITAC presentation, Jakobsson also pointed out that many of today's limitations in biomedical computing stem from software that:

- *is difficult to use,*
- *is fragile,*
- *lacks interoperability of different components*

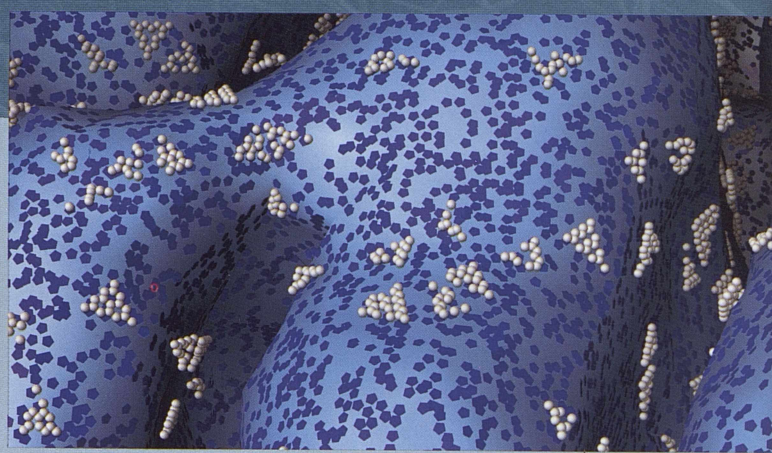
and further pointed out *the ongoing shortage of personnel trained to create and use better biological computing tools and environments.*

To a large extent, these shortcomings have been addressed by the PSC's Research Resource since its inception in 1987, when the PSC became the first extramural supercomputing center to receive funds from the NIH. Our goal for the future remains invariant—to develop and use HPCC to benefit biomedical research. The PSC Research Resource brings together experts from multiple disciplines, and provides access to cutting-edge computing by leveraging investments of other agencies. It has made software easier to use, made it more robust, presented it as part of an integrated framework, and has continually emphasized training and outreach. With the flexibility to respond quickly to unanticipated research opportunities, it has expanded the research domains that can effectively use HPCC from structural biology and sequence-based bioinformatics to include neural modeling, pathology, intra- and intercellular modeling, and visualization and analysis of extremely large imaging datasets, all components for a CEB.

We believe that the next step critical to realizing the vision of a CEB and the benefits it will bring to human health lies in realistic cell-to-organism level modeling. It is strategically positioned at an overlapping region of Structural Molecular Biology, Bioinformatics, Imaging, and Systems Biology, and is also strategically positioned between "bottom-up" and "top-down" approaches to biochemical network identification and functional analysis. Realistic physiological simulations present a grand challenge because of the wide range of underlying space and time scales, as well as the widely disparate organization and properties of different cells. Spatially realistic modeling will require new multiscale algorithms and prodigious amounts of computing, databases, storage, networking, and visualization. Such prodigious effort and costs, however, will be overwhelmingly outweighed by the resulting benefits to individuals and society at large. The PSC Research Resource will be a national leader in realizing these goals.

THE BIOMEDICAL INITIATIVE

NATIONAL LEADERSHIP IN HIGH-PERFORMANCE COMPUTING FOR BIOMEDICAL RESEARCH



In 1987, the PSC biomedical program became the first extramural biomedical supercomputing program in the country funded by NIH. Since then, with support from NIH's National Center for Research Resources (NCRR), PSC has fostered exchange between PSC expertise in computational science and experts in biology and medicine to solve important problems in the life sciences.

PSC workshops and courses on computational biology have trained more than 3,150 researchers in the use of high-performance computing for biomedical research, in such areas as sequence analysis in genome research, the structure of proteins and DNA, and biological fluid dynamics. "Our training reaches hundreds of biomedical scientists each year," says David Deerfield, director of PSC's biomedical initiative. "Techniques we've developed are helping scientists nationwide cope with the explosion of genome data."

Since its inception, PSC's biomedical program has provided computing resources for more than 1,000 biomedical research projects involving more than 2,500 researchers at 218 research institutions in 48 states. Among these are several projects featured in this booklet (pp. 20-31, 47 & 51).

MORE INFORMATION: http://www.psc.edu/biomed

SPATIALLY REALISTIC CELLULAR MODELING

PSC scientist Joel Stiles, an associate professor in the Mellon College of Science at Carnegie Mellon University, leads PSC's research in computational microphysiology. He is co-developer of MCell and DReAMM, software used in more than 100 laboratories around the world to simulate microcellular physiology.

Working with laboratory data from a patient with a neurological disorder called slow-channel congenital myasthenic syndrome, Stiles used MCell to successfully identify a previously unknown disease process. From MCell simulations, he deduced that receptors from a mutated protein in this particular patient were not only slow to close, but also slow to open, a previously unreported condition. Subsequent lab studies confirmed the finding—knowledge that can help in arriving at appropriate drug therapy as well as in research to develop better treatments.

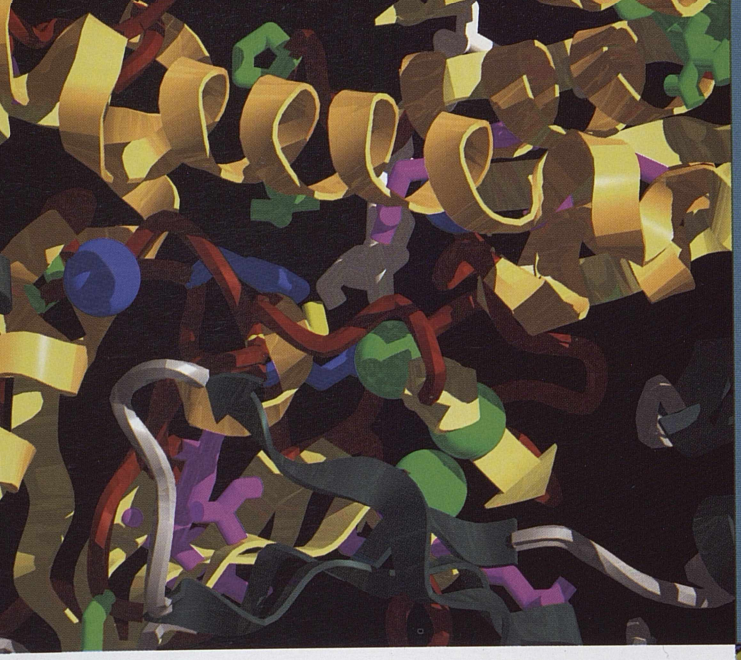
MORE INFORMATION: http://www.psc.edu/science/stiles.html

CORE RESEARCH

In addition to training and access to computational resources, PSC's biomedical group carries out its own core program of research in structural biology, protein and nucleic-acid sequence analysis, computational neuroscience and microphysiology and biomedical visualization.

PSC BIOMEDICAL COLLABORATIONS

Albert Einstein College of Medicine
Carnegie Mellon University
Duke University
Hospital for Sick Children, Toronto
Howard University
Morgan State University
North Carolina Central University
Rockefeller University
The Salk Institute
The Scripps Research Institute
University of California at Davis
University of North Carolina, Chapel Hill
University of Pittsburgh
University of Pittsburgh School of Medicine
University of Puerto Rico, Medical Sciences Campus



↓ FORCE-FIELD MODELS FOR MOLECULAR DYNAMICS

Charles Brooks of the Scripps Research Institute also does core research through the PSC biomedical initiative. His work involves the challenge of protein modeling, for which an important key is accurate "force fields"—the mathematical expression of energy relations among the atoms of the biomolecule being simulated.

Brooks is a pioneer in this area, having helped to develop the widely used force-field model, CHARMM. His recent work has been to develop atomic-level force-field models to faithfully represent the distribution of electric charge within biomolecules, critical for understanding the function of membrane proteins.

MORE INFORMATION: http://www.scripps.edu/mb/brooks

↑ COMPUTATIONAL TOOLS FOR GENOMICS AND PROTEOMICS

An expert in sequence analysis of DNA, RNA and proteins, PSC scientist Hugh Nicholas has used these techniques, in collaboration with University of Pittsburgh biologist John Hempel, to classify relationships among an important family of enzymes called aldehyde dehydrogenase (ALDH), which protect against toxicity and affect cancer treatment. The process involved, first, the alignment of 145 different ALDH sequences—one of the largest multiple-sequence alignments achieved.

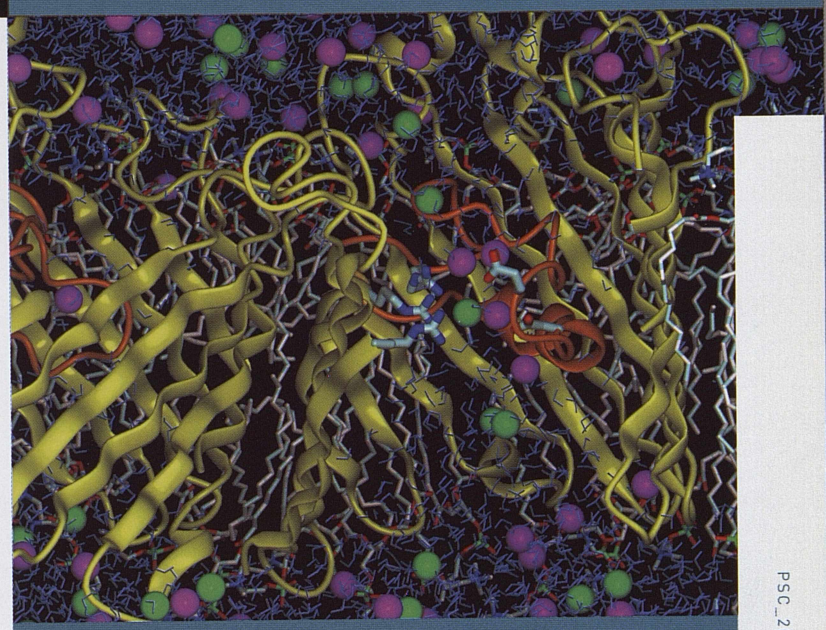
Nicholas and Hempel then applied techniques they developed to identify recurring sequence-motifs. Their work charted the family relationships among 13 distinct ALDH sub-families they identified and developed new insight into the relationship between ALDH sequence and structure.

MORE INFORMATION: http://www.psc.edu/science/hempel.html

↑ HYBRID QUANTUM AND CLASSICAL ENZYME SIMULATIONS

PSC scientist Troy Wymore collaborated with Deerfield, Hempel and Nicholas to elucidate the enzyme mechanism of a representative ALDH enzyme. He developed a methodology for combining high precision, but computationally expensive, quantum simulations with classical molecular dynamics. His simulations uncovered a novel chemical event—a proton transfer from the ALDH enzyme backbone.

MORE INFORMATION: http://www.psc.edu/science/wymore.html

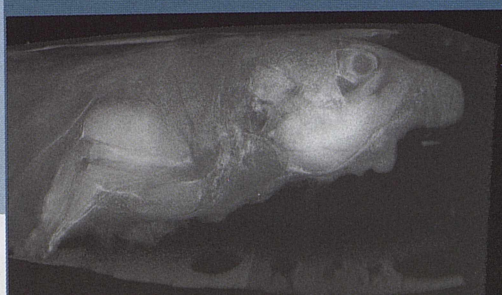


↓ VIEWING 3D DATA: THE PSC VOLUME BROWSER

PSC's biomedical initiative developed PVB, a graphical interface and rapid-retrieval system that enables users to rapidly view large 3D images over the internet. Viewing is interactive, with freedom to choose any angle or cross-sectional slice through the data to see from the inside. Because it allows viewing of uncompressed datasets from desktop computers, PVB is a breakthrough in 3D imaging.

PVB's network-delivery system allows up to 40 people to navigate independently through the data in near real-time. To gain speed, PVB also employs innovative data-compression that, for the most compressible data, achieves a ratio greater than 30:1. Used with the National Library of Medicine's Visible Human, PVB provides a versatile anatomy resource. In recent work, PSC has applied PVB to a similar large-volumetric dataset called the Visible Mouse.

MORE INFORMATION: http://www.psc.edu/biomed/research/VB/



Supercomputing in Pennsylvania

WITH COMMONWEALTH OF PENNSYLVANIA SUPPORT, PSC PROVIDES EDUCATION, CONSULTING, ADVANCED NETWORK ACCESS AND COMPUTATIONAL RESOURCES TO SCIENTISTS AND ENGINEERS ACROSS THE STATE

GLOBAL SUPERCOMPUTING COMES TO PITTSBURGH

From November 6-12 more than 6,000 people from around the world will gather at Pittsburgh's David Lawrence Convention Center for SC 2004. The annual conference of the global supercomputing community brings together equipment and software manufacturers, researchers and others to share ideas and assess new developments in high-performance computing, networking and storage.

Through the efforts of PSC executive director Beverly Clayton, who serves on the steering committee, and other PSC staff, the conference is in Pittsburgh for the second time—having been here before in 1996—with an associated economic boon estimated at $9 million.

ECONOMIC DEVELOPMENT

PSC's high-performance computing and networking help to boost the competitiveness of Pennsylvania business and industry and are among the resources the state can point to in attracting new business. Recognizing this, the Pennsylvania Department of Economic and Community Development features PSC machine time and consulting in incentive packages it offers to select companies.

As a PSC Industrial Affiliate, Pittsburgh-based PPG Industries uses LeMieux, PSC's terascale system, and the quantum-chemistry program GAUSSIAN for computational modeling in several aspects of its product lines as a global supplier of coatings, glass, fiberglass and chemicals.

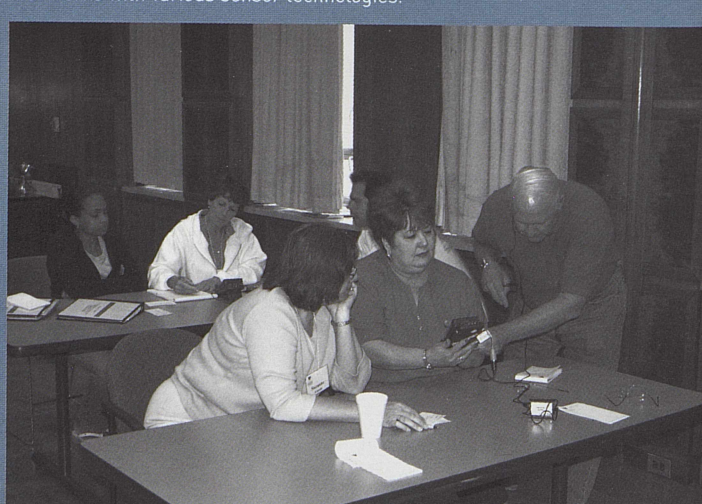
This year PSC presented a third annual customized technology-briefing day to staff from the Bechtel Bettis Atomic Power Laboratory in Pittsburgh. PSC consultants provided information on how to develop, manage and use a parallel distributed-computing environment.

Beverly Clayton, PSC executive director, coordinates PSC's program for Pennsylvania researchers.

TEACHER TECH '04

From August 2-6, PSC sponsored a weeklong workshop for Pittsburgh-area science teachers. Presented in collaboration with Rice University's Center for Equity and Excellence in Education and the National Science Foundation, the workshop introduced new technologies for helping to teach science and raised awareness about teacher roles in shaping and encouraging the next generation of scientists.

Three Pittsburgh Public Schools science teachers are using a Calculator Based Laboratory, a hand-held, data-collection system that works with various sensor technologies.



RESEARCH IN PENNSYLVANIA

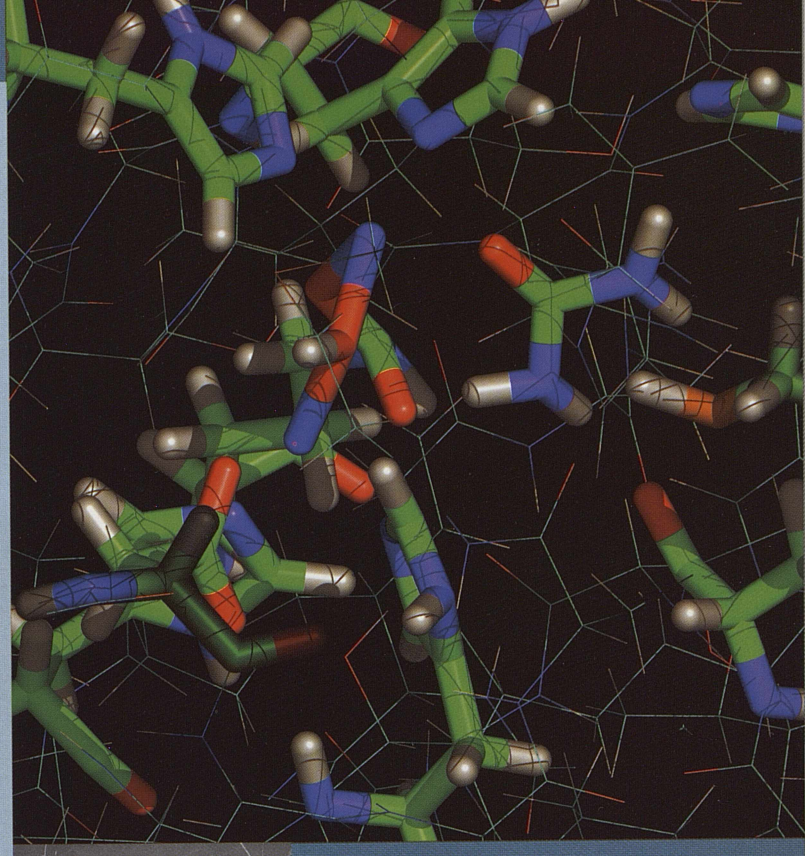
By supporting Pennsylvania university researchers, PSC resources help to attract research funds to the state. During the past year, more than 290 Pennsylvania researchers from 11 institutions used PSC computing. Pennsylvania usage of PSC's CRAY T3E totaled 1.25 million hours. In addition, Pennsylvania researchers received allocations through NSF of nearly 10 million hours on LeMieux, PSC's terascale system, along with 450,000 hours on PSC's HP Marvel systems. The projects represented here exemplify how supercomputing plays a role in scientific and engineering research in Pennsylvania.

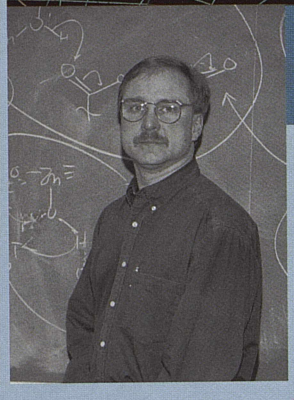
BIOMOLECULES AND THE SEARCH FOR DRUGS

Professor Ken Merz of Penn State's Department of Chemistry leads a research group that develops and uses computational methods to further understanding of biomolecules. They are especially interested in metalloenzymes, an important class of enzymes that carry out myriad biological functions. These enzymes are involved in diseases ranging from periodontal disease to arthritis and cancer, and drugs that inhibit and regulate these enzymes could be useful in treating these diseases.

Merz, who is also CEO of QuantumBio Inc, a biotechnology company he cofounded, works on developing improved computational techniques that include quantum-mechanical approaches—highly precise and at the same time computationally expensive—together with more conventional classical approaches. Some of his recent work has focused on an approach that will make it feasible to carry out fully quantum calculations on molecules as large as enzymes.

Using LeMieux, PSC's terascale system, Merz carried out a quantum-mechanical study of urease, a metalloenzyme that plays important roles in the nitrogen metabolism of plants and microbes. Urease in soil can release large quantities of toxic ammonia, affecting crop yields, and it also constitutes a virulence factor in human and animal infections of the urinary tract. This graphic (upper right) from PSC calculations shows details of the urease active-site structure.





Top: Urea in active site of urease

Left: **Ken Merz,** Pennsylvania State University

Omar Ghattas (left) and **Jacobo Bielak**, Carnegie Mellon University, finalists in the 2004 Computerworld award for Science, which recognizes innovation in information technology that benefits society.

WHEN THE EARTH SHAKES

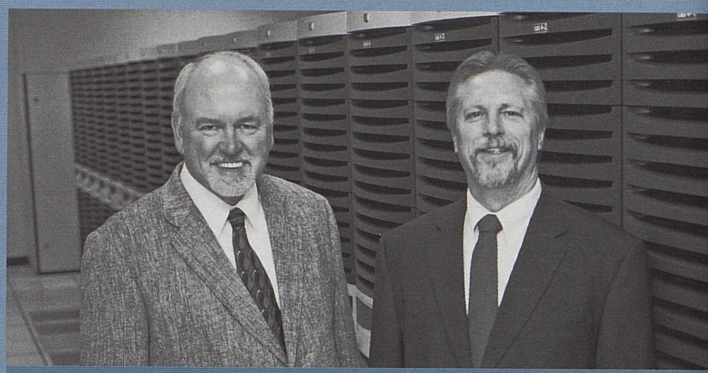
One of the main lessons of recent urban earthquakes is the need for better information about where and how much the ground will shake. A multi-disciplinary team led by Jacobo Bielak, Omar Ghattas and David O'Hallaron of Carnegie Mellon University have used LeMieux, PSC's terascale system, to simulate soil conditions for the Los Angeles earthquake basin with much greater accuracy than has before been possible. For this work, they received the 2003 Gordon Bell Prize, one of high-performance computing's most prestigious awards, and their work was recognized as a finalist for the 2004 Computerworld award in Science.

Using 3,000 LeMieux processors, they carried out the most detailed simulation yet of the 1994 Northridge earthquake, allowing soil to vibrate at two cycles per second, doubling the previous high frequency for earthquake simulation. Because of LeMieux, this group also made significant progress on the "inverse problem," a sophisticated approach that makes it possible to recover deep geological features based on seismic recordings at the surface.

MORE INFORMATION: http://www.psc.edu/science/2003/earthquake

The Super Computing Science Consortium

PENNSYLVANIA-WEST VIRGINIA PARTNERS IN DEVELOPMENT OF CLEAN POWER TECHNOLOGIES



(SC)^2 co-chairs: **Lynn Layman,** PSC, and **Bob Romanowsky,** NETL. "All (SC)^2 partners have particular areas of expertise," says Layman, "and through collaborations within (SC)^2 and with other regional organizations, we strengthen the research capabilities of all of the partners and the region as a whole."

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. 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.

PSC supplies training and computational resources along with visualization and consulting support to (SC)^2 partners. The training includes workshops in parallel programming and, during the past year, staff presentations as part of a graduate course in cluster computing at West Virginia University (WVU). The high-speed link allowed students to do homework on PSC systems.

Researchers at NETL and WVU have actively used this link to tap PSC computational resources. Since the founding of (SC)^2, more than 40 (SC)^2 researchers have used PSC systems for a range of projects, using more than a million hours of computing time, half a million within the past year. Along with the black-liquor project described here (facing page), this work includes:

- fluidized-bed combustion of silane, [http://www.psc.edu/publicinfo/netl/]
- lean-fuel mixes in next-generation power-generating turbines, [http://www.psc.edu/science/Richards/clean_power.html]
- industrial-scale technology for coal gasification, [http://www.psc.edu/publicinfo/2002/sc2/]
- a new design for a power-generating turbine, [http://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
Duquesne University
Pennsylvania State University
University of Pittsburgh
Waynesburg College
West Virginia University
Institute for Scientific Research
NASA Independent Verification & Validation Facility

MORE INFORMATION: http://www.sc-2.psc.edu

"WITH LEMIEUX WE'RE ABLE TO RUN MULTIPLE JOBS SIMULTANEOUSLY"

NEW PARTNER, MORE BANDWIDTH

In February 2004, the NASA Independent Verification & Validation Facility (NASA IV & V) in Fairmont, West Virginia became part of the $(SC)^2$ partnership, expanding the existing $(SC)^2$ network through a five-year contract with Qwest Communications. The contract increased the existing OC-3 network to OC-12 capacity and extended the network to Fairmont. The enhanced infrastructure provides for upgrading the network to OC-48 capability.

"The new architecture allows for expanded bandwidth capability to all parties at a significantly reduced cost," said Ned Keeler, director of NASA IV&V, which took the lead in the network upgrade. "Additional users may be added to the network as the north central West Virginia research community continues its growth."

"This is one more example of our continuing effort to attract leading technology providers to our region," said Dr. John Weete, WVU vice president for research. "Our goals in contributing to the enhanced network are to stimulate economic development for our entire area and increase our research capacity."

GAS FROM BLACK LIQUOR

The pulp and paper industry is among the nation's most energy-intensive manufacturing industries. Over the past 20 years, this industry dramatically increased its energy self-sufficiency by re-using black liquor, the residue from the process that extracts fiber from wood for papermaking. Existing technology burns the black liquor in boilers, primarily to recover the pulping chemicals. Gasification of the black liquor could substantially improve the efficiency of this process.

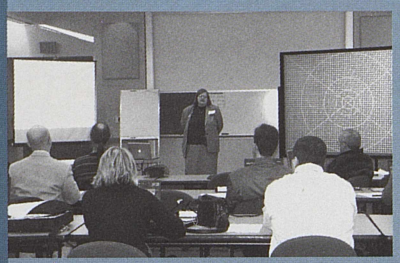
The U.S. Department of Energy has collaborated with Georgia Pacific (GP) to demonstrate full-scale black-liquor gasification. GP is replacing its existing boilers with a fluidized-bed system, which can combust and gasify black liquor—a mixture rich in carbon—both to recover chemicals and to generate steam and energy. In addition to improving energy efficiency, the system is expected to eliminate the danger of boiler explosions and dramatically reduce toxic air emissions.

To help GP with start-up of this system, NETL has used LeMieux, PSC's terascale system, to simulate the gasification process with MFIX (Multiphase Flow with Interphase eXchanges), software developed at NETL. MFIX can predict hydrodynamic behavior, track gas and solid species, and temperature distribution throughout the bed. "With LeMieux, we're able to run multiple jobs simultaneously," says NETL consulting engineer Chris Guenther, "to vary the particle size and gas flow rates. These simulations generate very large datasets, which PSC accommodates with their FAR storage system. The high-speed fiber link between PSC and NETL allows us to bring these datasets back to NETL in minutes."

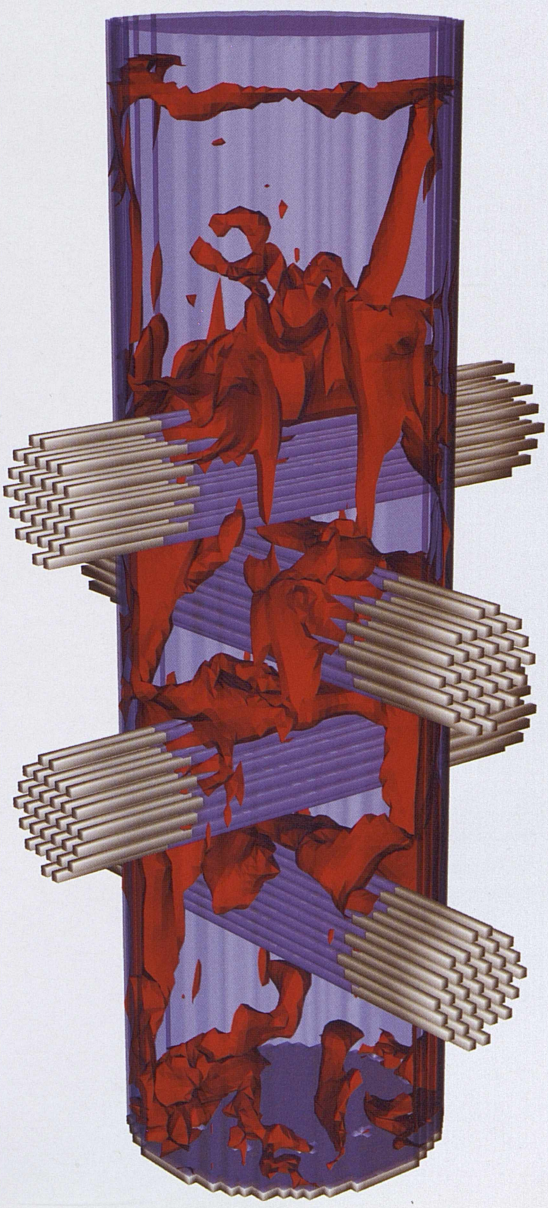
Four tube bundles—oriented perpendicular to one another — supply heat to the bed. Fluidization gas enters through the bottom and black liquor is injected into the lower region. The black liquor coats an inert bed material which combusts and gasifies as it travels upward, forming bubbles that flow through and around the tube bundles. The transient behavior of these bubbles determines the mix of gas versus solids in the flow through the bed, with the goal of keeping the bed well mixed and thermally homogeneous. To model such a complex system, it is critical to predict bubble behavior (red).

THE $(SC)^2$ VISUALIZATION WORKSHOP

On March 22, 2004, the $(SC)^2$ Visualization Working Group held a workshop for Consortium partners in Pennsylvania and West Virginia interested in scientific visualization and virtual reality. With nearly 30 participants—from Consortium members Duquesne University, Pennsylvania State University, WVU, Institute for Scientific Research, NASA IV&V, PSC and NETL—the workshop provided a forum for $(SC)^2$ partners to learn about each other's capabilities and research.



Frances VanScoy of WVU, chair of the $(SC)^2$ Visualization Working Group, welcomes participants and describes her research.



Notes & Highlights

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 network know-how. They carry out research on network performance and analysis in support of high-performance computing applications, provide engineering consulting for advanced networking nationally, and conduct seminars that disseminate knowledge to engineers around the country. In projects such as Web100 and Net100, they develop technologies that will define the networks of the future.

MORE INFORMATION: http://www.psc.edu/networking

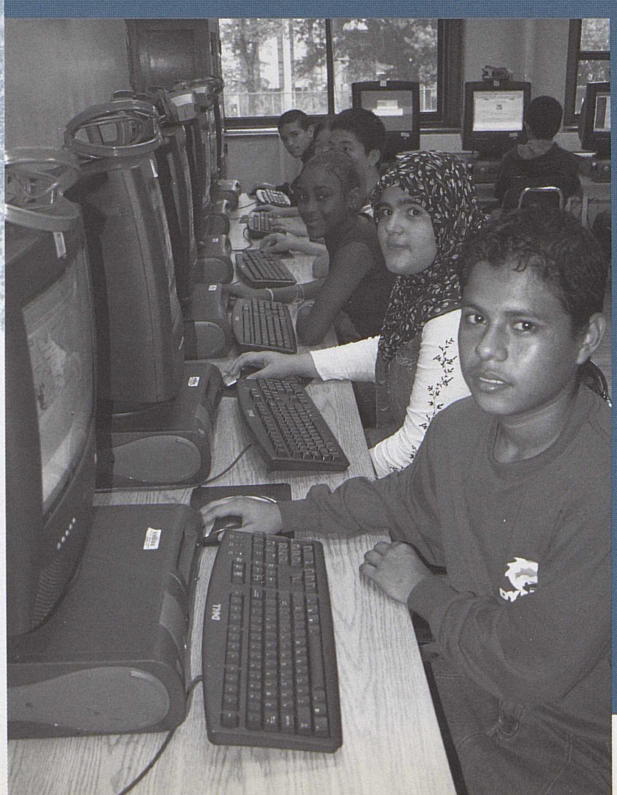
GETTING IN TUNE WITH WEB100 & NET100

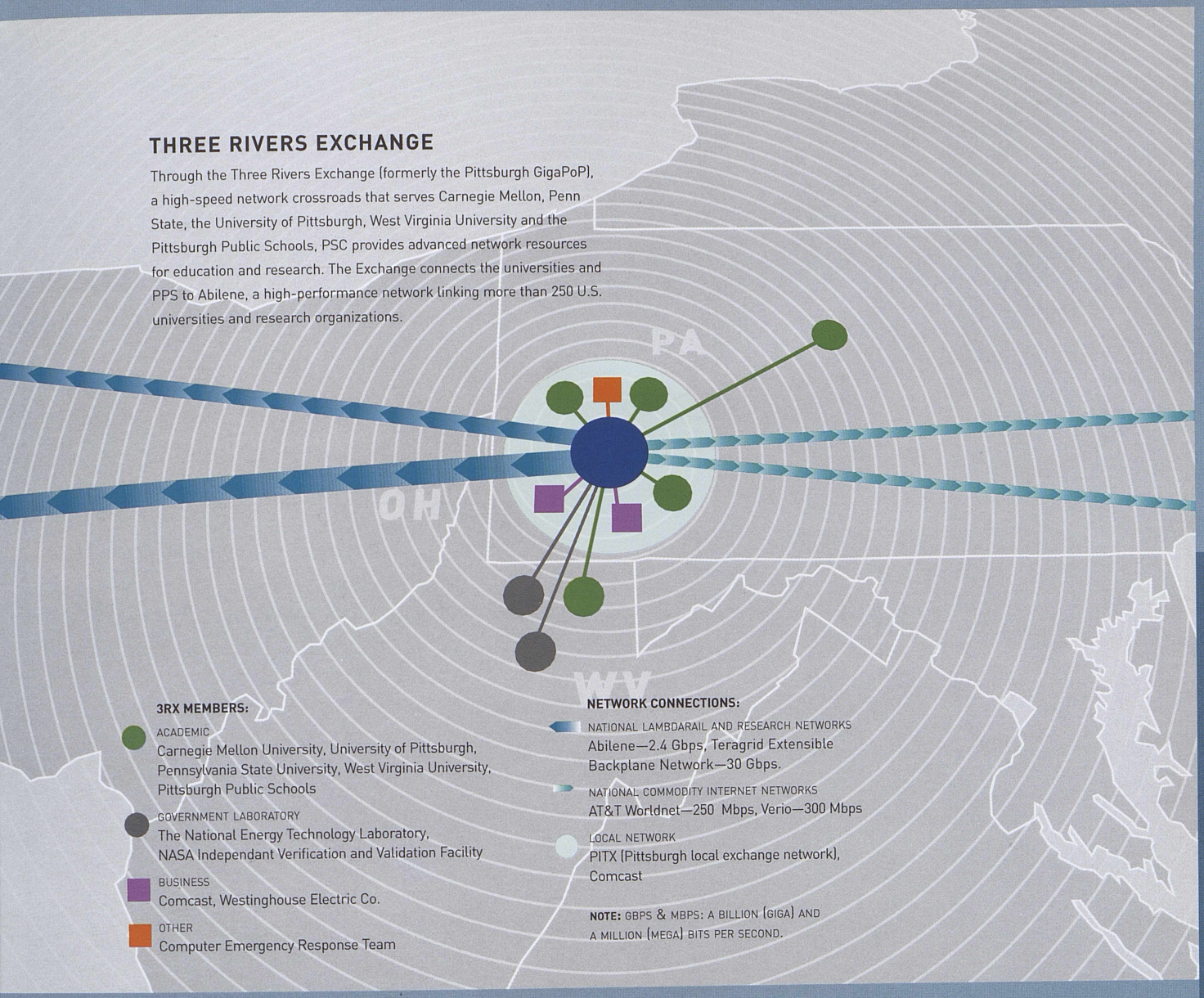
With funding from the National Science Foundation and Cisco Systems, PSC network researchers—in collaboration with the National Center for Atmospheric Research (NCAR) and the National Center for Supercomputing Applications—have developed software to "tune" the Internet protocol in computer operating systems to better exploit available network bandwidth. Called Web100, this software is now used in many research projects nationwide.

Portions of Web100 are now included in the Linux operating systems and a development release of Windows. With both operating systems, the Web100 effort is working to enable ordinary scientific users with high-performance network connections to automatically attain high-speed data transfers (100 megabits per second or higher) without help from network experts or manual tuning.

In a related project called Net100, funded by the U.S. Department of Energy, PSC's network research group collaborated with Lawrence Berkeley National Laboratory, Oak Ridge National Laboratory and NCAR to create software tools that allow operating systems to tune themselves in response to changing conditions on the network. PSC incorporated selected Net100 enhancements into Web100.

Students using the network at Frick International Studies Academy, Pittsburgh Public Schools





THREE RIVERS EXCHANGE

Through the Three Rivers Exchange (formerly the Pittsburgh GigaPoP), a high-speed network crossroads that serves Carnegie Mellon, Penn State, the University of Pittsburgh, West Virginia University and the Pittsburgh Public Schools, PSC provides advanced network resources for education and research. The Exchange connects the universities and PPS to Abilene, a high-performance network linking more than 250 U.S. universities and research organizations.

3RX MEMBERS:

ACADEMIC
Carnegie Mellon University, University of Pittsburgh, Pennsylvania State University, West Virginia University, Pittsburgh Public Schools

GOVERNMENT LABORATORY
The National Energy Technology Laboratory, NASA Independant Verification and Validation Facility

BUSINESS
Comcast, Westinghouse Electric Co.

OTHER
Computer Emergency Response Team

NETWORK CONNECTIONS:

NATIONAL LAMBDARAIL AND RESEARCH NETWORKS
Abilene—2.4 Gbps, Teragrid Extensible Backplane Network—30 Gbps.

NATIONAL COMMODITY INTERNET NETWORKS
AT&T Worldnet—250 Mbps, Verio—300 Mbps

LOCAL NETWORK
PITX (Pittsburgh local exchange network), Comcast

NOTE: GBPS & MBPS: A BILLION (GIGA) AND A MILLION (MEGA) BITS PER SECOND.

ON THE NETWORK: EXCHANGE NEWS

In partnership with the University of Pittsburgh, PSC this year became a full member of National LambdaRail, the next-generation network infrastructure. Rather than a single network, the NLR infrastructure is multiple networks that exist side-by-side in the same fiber-optic cable, but are independent, each supported by its own lightwave, or lambda.

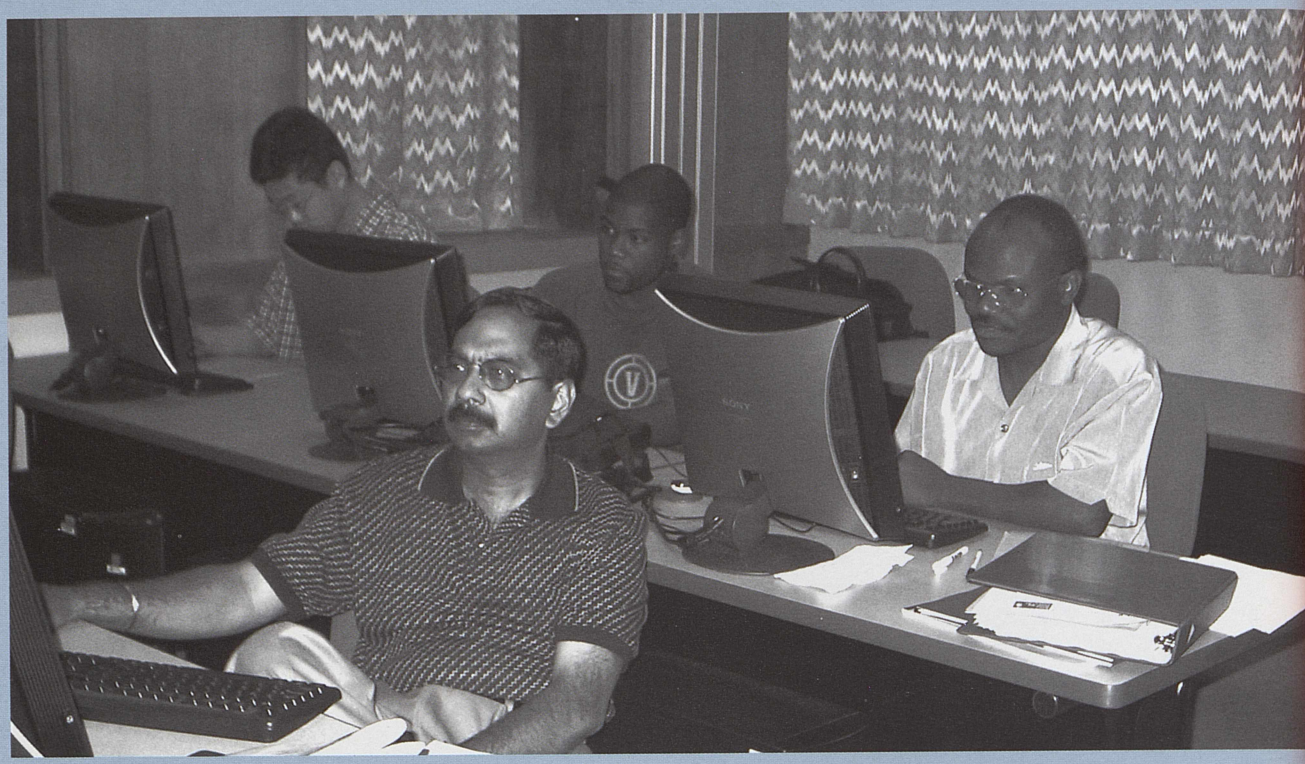
"NLR pushes beyond existing limitations of conventional Internet backbones," says PSC director of networking Wendy Huntoon. "This will allow us, for instance, to create a dedicated connection for a specific purpose, such as to support rapid data transfers between PSC and Oklahoma for a tornado forecasting experiment."

At least half the NLR resource will support networking research. PSC network engineers will use it to expand their work on Path MTU Discovery, a method to assess the best Maximum Transfer Unit—the biggest data packet a given connection can efficiently use.

In August 2004, the Exchange installed a high-speed link to high-speed Internet provider Comcast. This peering arrangement allows Comcast and the Exchange to exchange traffic directly rather than going through national Internet backbones, enhancing access to regional resources for Comcast high-speed Internet subscribers.

"This not only improves communication between regional individuals and institutions," says Huntoon, "it also lowers costs. By keeping local network traffic local, Comcast and Exchange customers reduce latency in local connections and save money from reduced Internet bandwidth use."

The Exchange also peers with PITX—a peering exchange that includes local Internet Service Providers such as Pair Networks, Telerama and Nauticom. Comcast users, therefore, should also see improved performance when connecting to these service providers.



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

PITTSBURGH SUPERCOMPUTING CENTER WORKSHOPS (2003-2004)

Introduction to Terascale Code Development

Terascale Code Development

New Methods for Developing Petascalable Codes

Developing Bioinformatics Programs

Statistical Analysis of Neuronal Data

Computational Cell Biology: Modelling and Simulation

Parallel Programming Techniques

Nucleic Acid and Protein Sequence Analysis

The PSC operational management team: (front, l to r) **David Kapcin**, manager, financial affairs; **Nick Nystrom**, manager, strategic applications; **Bob Stock**, associate director; **Rich Raymond**, manager, user support; **Sergiu Sanielevici**, assistant director, scientific applications and user support; **J. Ray Scott**, assistant director, systems and operations; (back, l to r) **John Kochmar**, facilities manager, systems and operations; **David Deerfield**, assistant director, biomedical initiative; **Jim Marsteller**, manager, Three Rivers Exchange. Not in photo: **Gwendolyn Huntoon**, **Janet Brown**, **Joel Stiles**, **Elvira Prologo**.

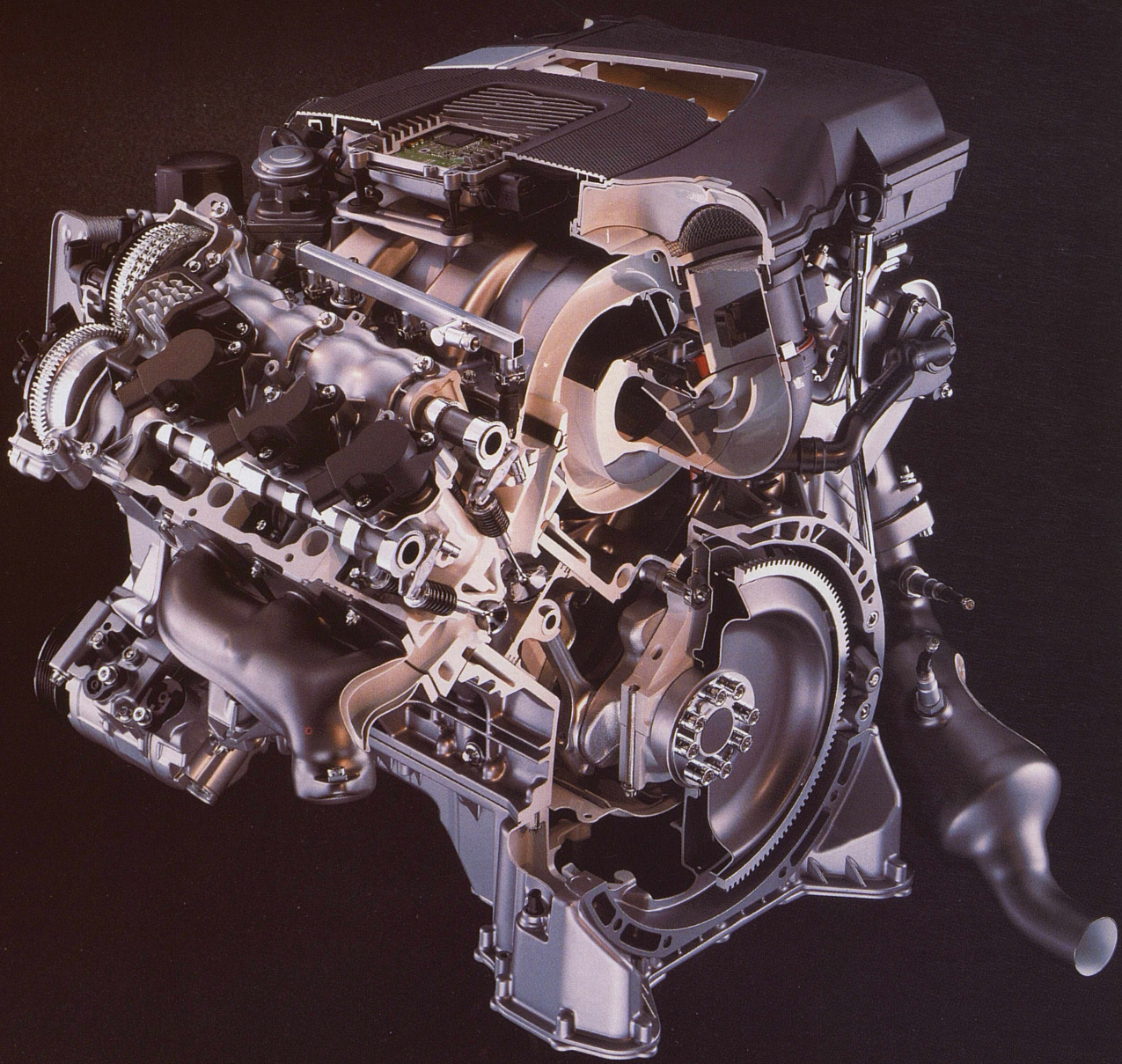
PROJECTS_IN_SCIENTIFIC_COMPUTING_2004

PROJECTS 2004
CONTENTS

STRUCTURE_OF_PROTEINS_AND_DNA

PROTEIN MOTORS,

WITH PSC'S HP MARVEL, AN EXCEPTIONAL MACHINE FOR QUANTUM COMPUTATIONS, RESEARCHERS ATTACKED A KEY PROBLEM IN PROTEIN BIOLOGY



DaimlerChrysler AG

Incorporated

Energy in the body comes from millions and millions of tiny power generators, each equipped with a crankshaft that spins round and round 24/7, producing the fuel that makes us go.

Right. And the moon is made of Gouda cheese.

Suspend your disbelief. The protein adenosine triphosphate synthase, better known as ATPase, is nature's smallest rotary motor. "You can take a spoonful of that protein," says biophysicist Klaus Schulten of the University of Illinois Urbana-Champaign, "and it generates as much torque as a Mercedes engine."

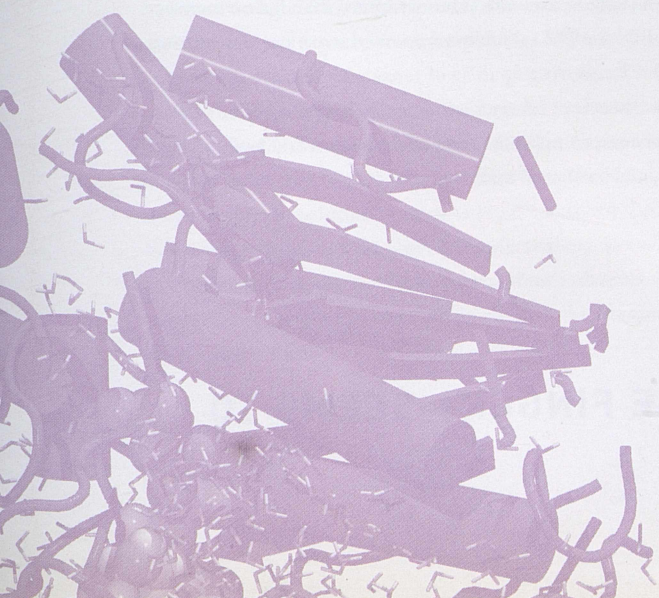
A remarkable molecular motor that in the laboratory produces torque from chemical fuel, ATPase works the other direction in humans—converting torque into ATP, the basic fuel of life, the chemical energy that fuels muscle contraction, transmission of nerve messages and many other functions. Probably the most abundant protein in all living organisms, ATPase is the power plant of metabolism. In an active day, an adult human can produce and consume its body weight or more of ATP, nearly all of it produced by ATPase.

The 1997 Nobel Prize in Chemistry recognized Paul Boyer and John Walker for their work in assembling, for the first time, a detailed picture of ATPase and how this tiny molecular machine does its job. Subsequent research has added to the picture, but many challenging questions remain.

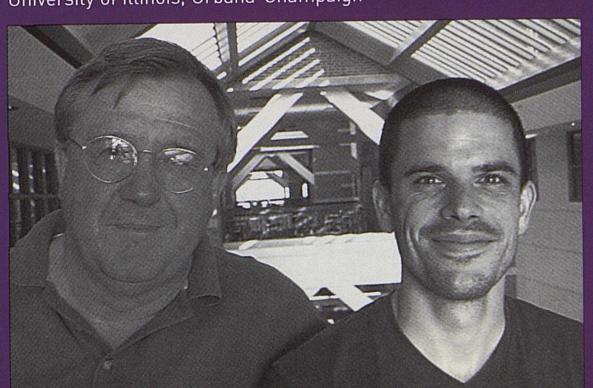
Imagine we're from Mars, says Schulten, director of the Theoretical and Computational Biophysics Group at the University of Illinois Beckman Institute, and we want to understand how a car engine works, but the engine is infinitesimally small. "It's nearly impossible to see the details in motion. The only way is to use the computer to simulate it, and then we can recognize the combustion process driving the car engine. It appears that ATPase is kind of a combustion engine too."

To see what they could see, Schulten and graduate student Markus Dittrich turned to Jonas, PSC's 128-processor HP Marvel system, which is dedicated to biomedical research. With the exceptional capability of this system, they were able to simulate "combustion" in ATPase with quantum theory — to get a picture of how electrons move from atom-to-atom during chemical reactions.

"If you want to do careful calculations," says Schulten, "you have to invest a lot of computing power. It would make no sense to approach this important problem with cheap methodology. We decided to do the most advanced calculation that people do today when simulating a biological reaction." Jonas's very fast EV7 processors and large shared memory made the quantum calculations feasible, and the resulting study turned up invaluable new information about nature's tiniest motor.



Klaus Schulten (left) and **Markus Dittrich**, University of Illinois, Urbana-Champaign



THE WHEEL SPINS ROUND AND ROUND

Like most motors, ATPase has moving and non-moving parts. There's a wheel that spins, similar to a millwheel, to turn an axle that revolves inside a hexagonal cluster, in which there are three combustion chambers (active sites), each of which, in sequence, charges up with chemical raw materials—adenosine diphosphate (ADP) and phosphate—and "fires" to produce ATP.

The wheel part of the protein is called F0. In humans and other mammals, F0 resides in the membrane of mitochondria, microscopic structures inside the cell. In bacteria, where ATPase works reversibly both as an ATP generator and an ATP-fueled motor, F0 sits in the cellular wall. In both cases, it forms a channel for protons to flow through the wall, a flow which—much like water turning a millwheel—causes F0 to rotate.

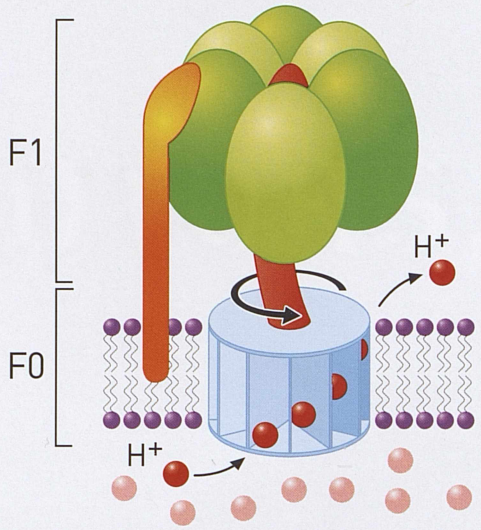
The other main part of ATPase's structure, called F1, extends into the cellular or mitochondrial interior. F1 includes a central stalk—the axle—that is coupled to and turns with F0. The other end of this axle revolves inside F1's hexagonal cluster, which contains the three active sites where "combustion" occurs.

"F0 spins, and this spins the F1 central stalk," says Dittrich, "and this leads to ATP synthesis." The axle spins in one direction (clockwise when viewed from the membrane side) during ATP production. What's not only fascinating—but also a large benefit for understanding the protein—is that the reverse reaction also works in the laboratory. Put F1, a very large protein by itself, in solution with ATP, and ATP will "hydrolyze" inside F1's active sites into ADP and phosphate and the axle will spin counterclockwise.

Because ATP hydrolysis is a chemical mirror of synthesis and more amenable to laboratory study, it offers an invaluable back-door approach to gaining knowledge about ATPase, and this reaction is what Schulten and Dittrich set out to simulate.

ZOOMING-IN ON THE SPARK PLUG

One of the intriguing questions about F1-ATPase is how its three combustion chambers cooperate with each other. During each 360-degree rotation of the F1 axle, each active site, one after another, changes structure and thereby alters its function. For the hydrolysis reaction, each site is open to bind with an ATP molecule, closes to hold it during breakdown into constituent products, and opens again during product release. With each rotation, three ATP molecules hydrolyze, one at a time. Schulten and Dittrich hoped to shed light on the atom-by-atom details of this rotary process.



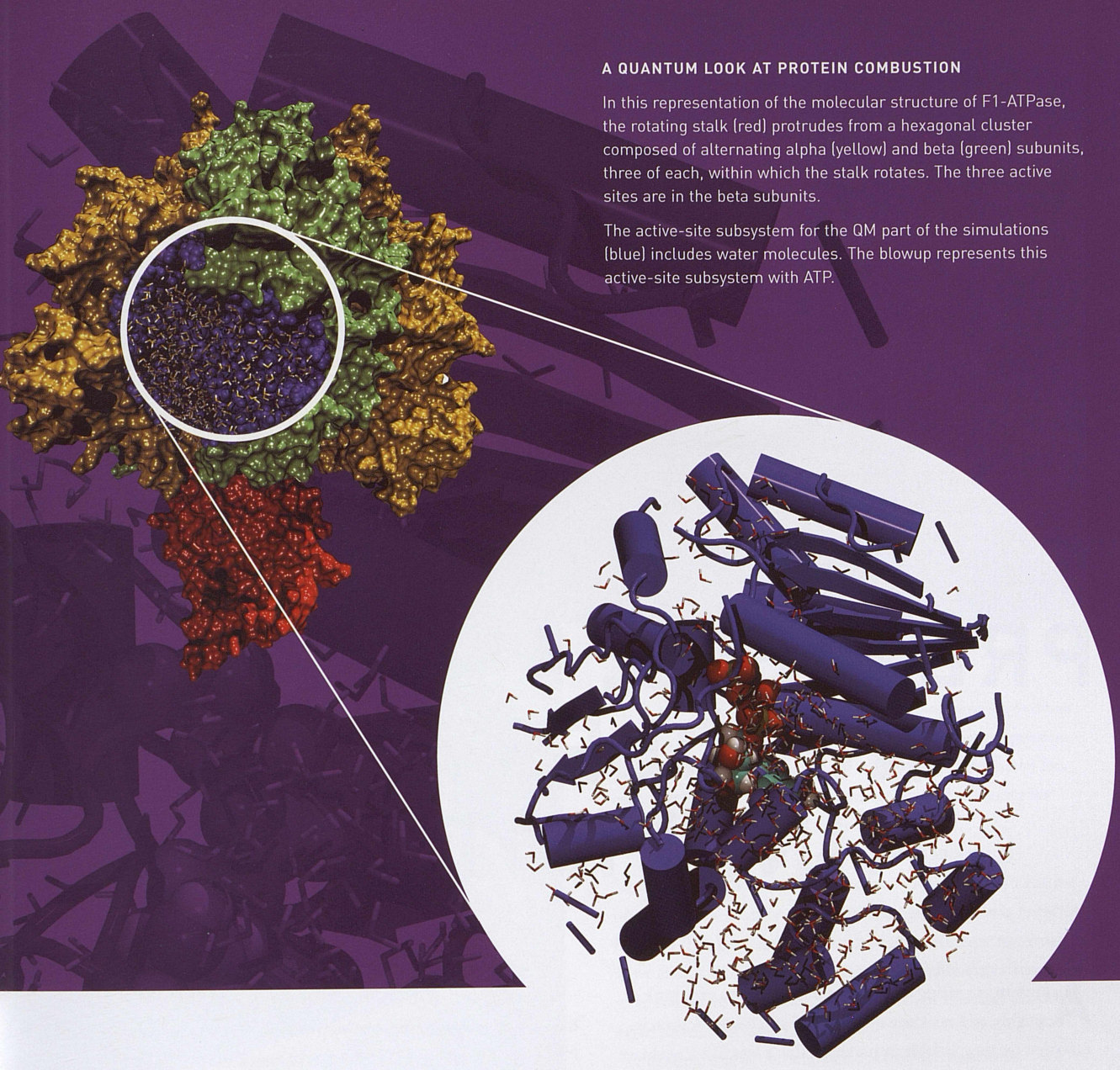
This cartoon schematic shows the F0 part of ATPase (blue) sitting in a membrane, where proton (H+) concentration outside (below) the membrane greater than on the inside causes F0 to rotate. The axle (red) of F1 also rotates and triggers rotary "combustion" in three active sites, contained within alternating subunits (green) of F1's hexagonal active-site complex.

Another challenging question concerns how the hydrolysis reaction causes the F1 axle to spin. "We don't understand how the chemical reaction is coupled to rotation of the axle," says Schulten. "That's still a big mystery."

To get at the crucial details of bonds breaking and reforming during a chemical reaction requires quantum theory. Schulten and Dittrich therefore used a method called QM/MM (quantum mechanics/molecular mechanics), which made it possible to simulate the entire subunit of F1 that houses the active site, but used quantum theory selectively like a zoom lens to zoom-in on the active site itself, where "combustion" occurs.

In total, they simulated two different configurations of the F1-ATPase active-site subunits, more than 8,000 atoms each, while the QM calculations zoomed in on the combustion chamber. Employing up to 32 of Jonas's processors at a time, with an extensive series of simulations, they used over 12,000 hours of computing time. The outcome is several key findings on a crucial biological system.

THIS AMINO ACID—THE ARGININE FINGER—SEEMS TO OPERATE LIKE A SPARK PLUG



A QUANTUM LOOK AT PROTEIN COMBUSTION

In this representation of the molecular structure of F1-ATPase, the rotating stalk (red) protrudes from a hexagonal cluster composed of alternating alpha (yellow) and beta (green) subunits, three of each, within which the stalk rotates. The three active sites are in the beta subunits.

The active-site subsystem for the QM part of the simulations (blue) includes water molecules. The blowup represents this active-site subsystem with ATP.

ATP in solution without ATPase is extremely slow to hydrolyze, taking as long as a week. With ATPase, this reaction goes about 100 billion times faster, a huge speedup that researchers have been hard pressed to explain. The Jonas simulations reinforce Schulten and Dittrich's finding (in a prior simulation study) that ATPase hydrolysis—in which water attacks ATP to break it down—is carried out by two water molecules in concert, rather than only one, as had been thought. This finding—unobservable in laboratory experiment—is a big step toward explaining ATPase's remarkable catalytic efficiency.

"In order to have this kind of concerted action of two water molecules," explains Dittrich, "they have to be arranged in a particular way. This is accomplished by ATPase and won't happen in solution because there it's unlikely the water molecules will assume this special conformation. If we didn't have this mechanism, the reaction would take place on a much slower time-scale and wouldn't lead to the observed physiological rates.

The simulations also show, unexpectedly, that there's no energy change in the active site as ATP breaks down into its reaction products—a finding that goes further than experiments in establishing that the reaction itself doesn't provide any force toward making the F1 axle rotate. "It's not the chemical event that drives the rotation," says Dittrich, "which means we have to look at other possibilities." The two remaining possibilities are when ATP binds to the active site or when the reaction products are released.

Perhaps most importantly, the simulations reveal that one of the amino acids in the active site of F1-ATPase—arginine—appears to play a key role in coordinating the timing among the three active sites. This amino acid—referred to as the arginine finger—seems to operate like a spark plug. It shifts its position depending on whether ATP or the reaction products are in the combustion chamber. "We think that the findings," says Schulten, "particularly with respect to the arginine finger, could well prove to be a crucial part of this puzzle." **MS**

MORE INFORMATION:

http://www.psc.edu/science/schulten2004.html

A SLIGHT CHANGE IN AMINO ACIDS DRAMATICALLY AFFECTS PROTEIN STRUCTURE

PRETZELS, NOODLES & MINI PROTEINS

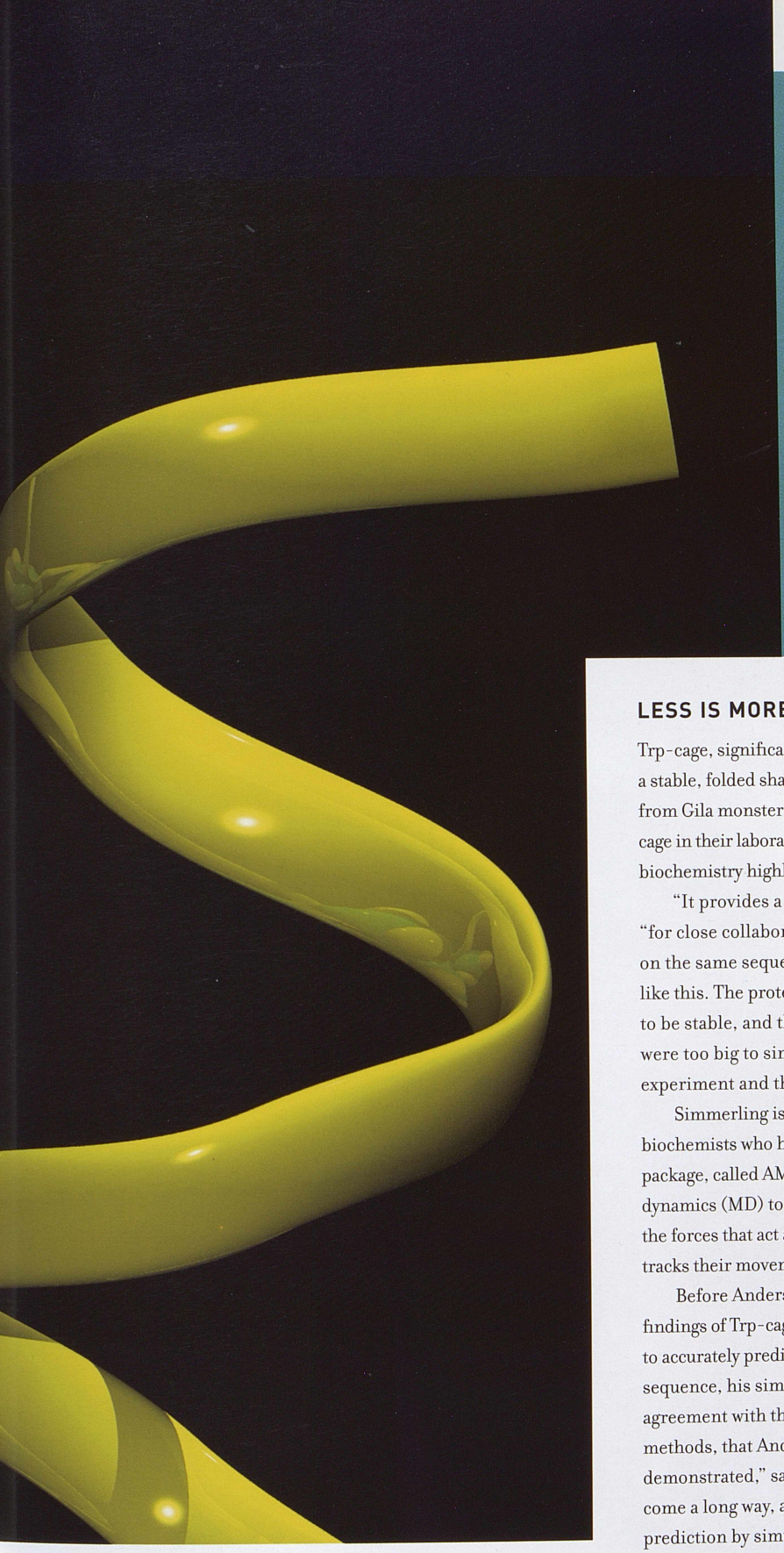
A protein is a string of amino acids, which you can think of, roughly, as a necklace of multi-colored beads. A protein, however, twists and folds into a shape that's less like a necklace than a pretzel, and each protein has its own unique pretzel shape, determined by which beads are on the necklace and in what sequence.

Scientists know that sometimes if you switch one bead for another, the pretzel doesn't change much, but other times—depending on which bead you switch or which new one you substitute—the pretzel may radically alter its shape. When a gene mutation causes this to happen in people, it may—depending on the protein—lead to a debilitating disease.

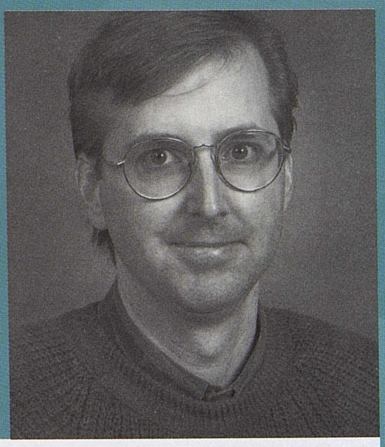
No wonder, then, that scientists want to decipher the underlying code of relationships between amino acids and protein shape, called the protein-folding problem. "It's like saying you have a sentence and it has a meaning," says computational chemist Carlos Simmerling. "How many words can you change before it doesn't have that meaning anymore? How many of the amino acids in the protein can you change before it doesn't fold properly and you have a disease?"

Simmerling, of the State University of New York at Stony Brook, and graduate student Melinda Layten used LeMieux, PSC's terascale system, to see what happens when you switch beads on the necklace. With LeMieux, he's been asking these questions with a special protein called Trp-cage, a short necklace of only 20 amino acids —compared to hundreds for most proteins. Because of its small size, Trp-cage's folding is relatively simple and presents a model case for analysis.

Working in close collaboration with a laboratory team led by Niels Andersen at the University of Washington, Simmerling made efficient use of up to 1,000 of LeMieux's processors to simulate Trp-cage, all of its atoms—first in its native state, then with several carefully chosen variations. He's found, dramatically, that one switched amino acid transforms the pretzel to a floppy noodle, but another subtle switch changes it back to a pretzel. His simulations add detail and complexity to an emerging picture of sensitive relations between amino acids and structure in this special minimalist protein.



Carlos Simmerling, The State University of New York at Stony Brook



LESS IS MORE

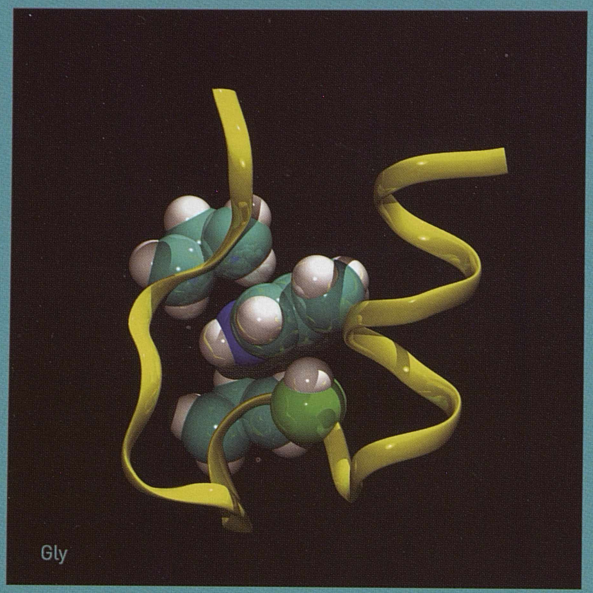
Trp-cage, significantly, is the smallest protein known that has a stable, folded shape. In 2002, starting with a longer protein from Gila monster saliva, Andersen and his team created Trp-cage in their laboratory. Their work was lauded as among the biochemistry highlights of the year.

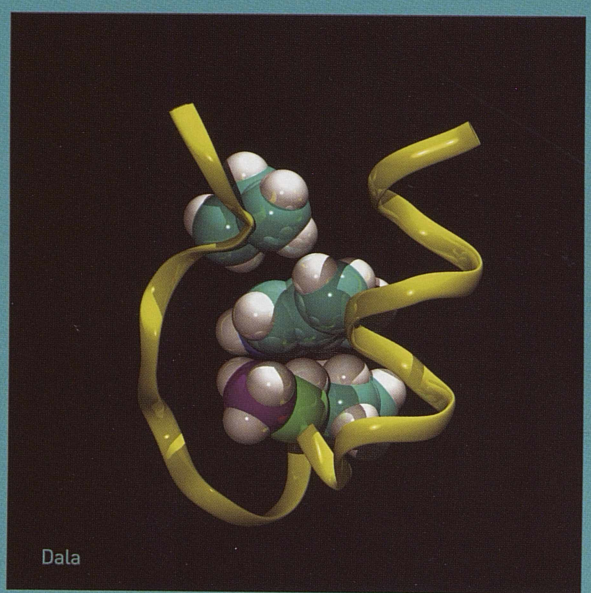
"It provides a model system," says Simmerling, "for close collaboration between experiment and simulation on the same sequence, and until Trp-cage there was nothing like this. The proteins that we could simulate were too small to be stable, and the systems that were stable experimentally were too big to simulate. It's an important meeting place for experiment and theory."

Simmerling is one among a team of computational biochemists who have developed a widely used software package, called AMBER, that employs a method called molecular dynamics (MD) to simulate proteins and DNA. MD calculates the forces that act among all the atoms in the molecule and tracks their movement over time.

Before Andersen's group released their experimental findings of Trp-cage's structure, Simmerling used AMBER to accurately predict it. Starting with only the amino-acid sequence, his simulations arrived at a structure in excellent agreement with the configuration, as determined by NMR methods, that Andersen's group subsequently published. "We demonstrated," says Simmerling, "that MD simulations have come a long way, and are at a point where accurate structure prediction by simulation may soon be routine enough to contribute significantly to our understanding of folding."

This simulation also suggests structural detail that goes beyond the experimentally determined shape. Based on these details, Andersen's group is working to further analyze and refine the structural picture.





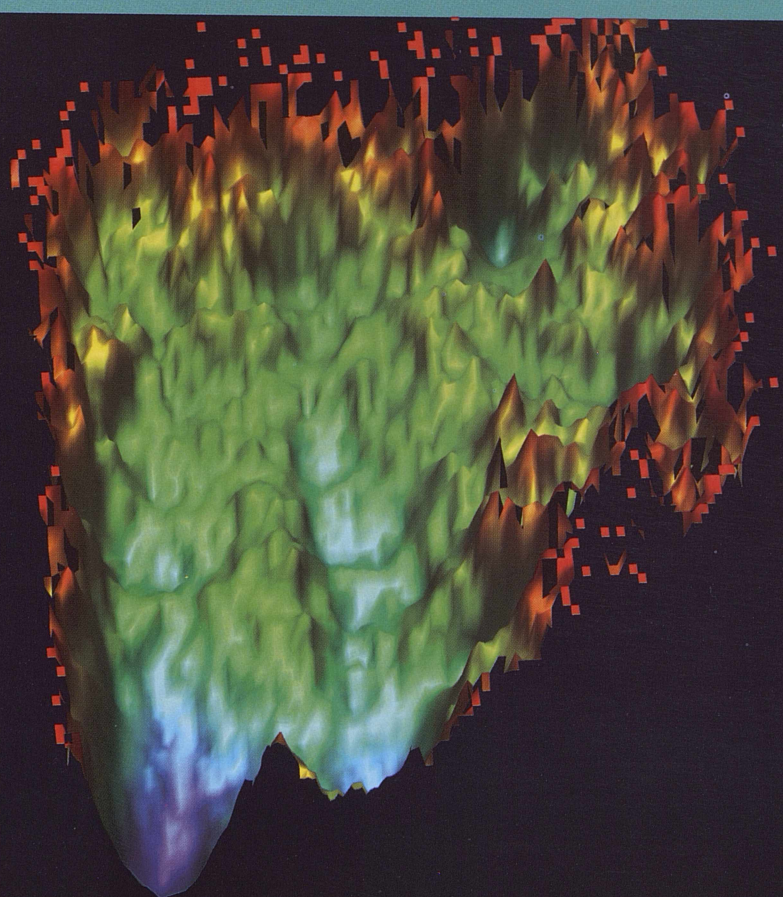
FOLDED, UNFOLDED, REFOLDED

These graphics from the simulation show a simplified representation of Trp-cage with the protein backbone (yellow ribbon) and selected amino acids. In the native, folded structure (Gly), spheres of cyan (carbon) and white (hydrogen) represent side-chains of tryptophan (blue, nitrogen) within a "cage" of two prolines. This characteristic fold gives Trp-cage its name. The small glycine (green) is at a bend in the fold.

In one of many structures (Lala) adopted by Trp-cage after alanine (purple side chain) replaces the glycine, the tryptophan no longer packs into the cage and the protein remains floppy. A slight change to the alanine side-chain permits the tryptophan to pack in the cage, forming a structure (Dala) similar to native Trp-cage.

RIDGES AND VALLEYS OF PROTEIN ENERGY

This graphic represents the free-energy landscape of Trp-cage. Color (increasing from dark blue to red) corresponds to altitude in the landscape. Stable proteins tend to form in structures that correspond to the low-energy valleys.



REPLICA EXCHANGE & ENERGY LANDSCAPES

To predict Trp-cage's shape, Simmerling relied on a well-tested axiom of molecular structure. A molecule tends to move toward being in a shape in which the atoms expend the least possible energy to maintain structure. Although Simmerling's structure-prediction simulation was remarkably successful, proteins in the real world aren't represented solely by their low-energy state.

"There's a native lowest energy structure," explains Simmerling, "and at the same time, there may also be non-native, unfolded structures. We need to know not just the native structure, but we want to know its relative probability. Is it native 30 percent of the time? Or 90 percent? All we can say from the first simulation is this structure is best, but we don't know what that means in terms of real stability."

To produce this information on relative probabilities, Simmerling employed a method called "replica exchange," a relatively recent innovation in MD simulation that he implemented as part of AMBER. In this approach, designed to exploit a massively parallel system such as LeMieux, many separate simulations of a single protein run at the same time on different processors. The processors exchange information to arrive, eventually, at a picture of the protein's "energy landscape"—a map of the relation between possible shapes and their likelihood.

Among amino acids, glycine is the smallest, and is distinctive in having no attached chemical group, called a side chain. For this reason, glycine provides structural flexibility and often appears at a tight turn in a protein's fold. In experiments, changing a Trp-cage glycine to alanine—an amino acid only slightly larger—prevented folding. "We can take something that's over 90 percent folded," says Simmerling, "make one small change, and it won't fold anymore. Alanine has only a small side chain—a single methyl group (CH_3)—and yet it has a huge effect."

Somewhat differently from the experiments, the replica-exchange simulations showed that with alanine the folded structure still exists, but it's highly unstable. Spurred by these simulations, further experiments found that this altered Trp-cage retains some areas of structure.

Based on analysis of these results, the researchers tried another switch. They replaced the alanine with its mirror image, called d-alanine—which flips alanine's side-chain from one side to the other of the protein. With this slight change, simulations showed that Trp-cage regained nearly all of its folded stability. Experiments confirm that this switch restores nearly full stability.

USING 1,000 PROCESSORS FOR REPLICA-EXCHANGE SIMULATIONS MAPS THE ENERGY LANDSCAPE

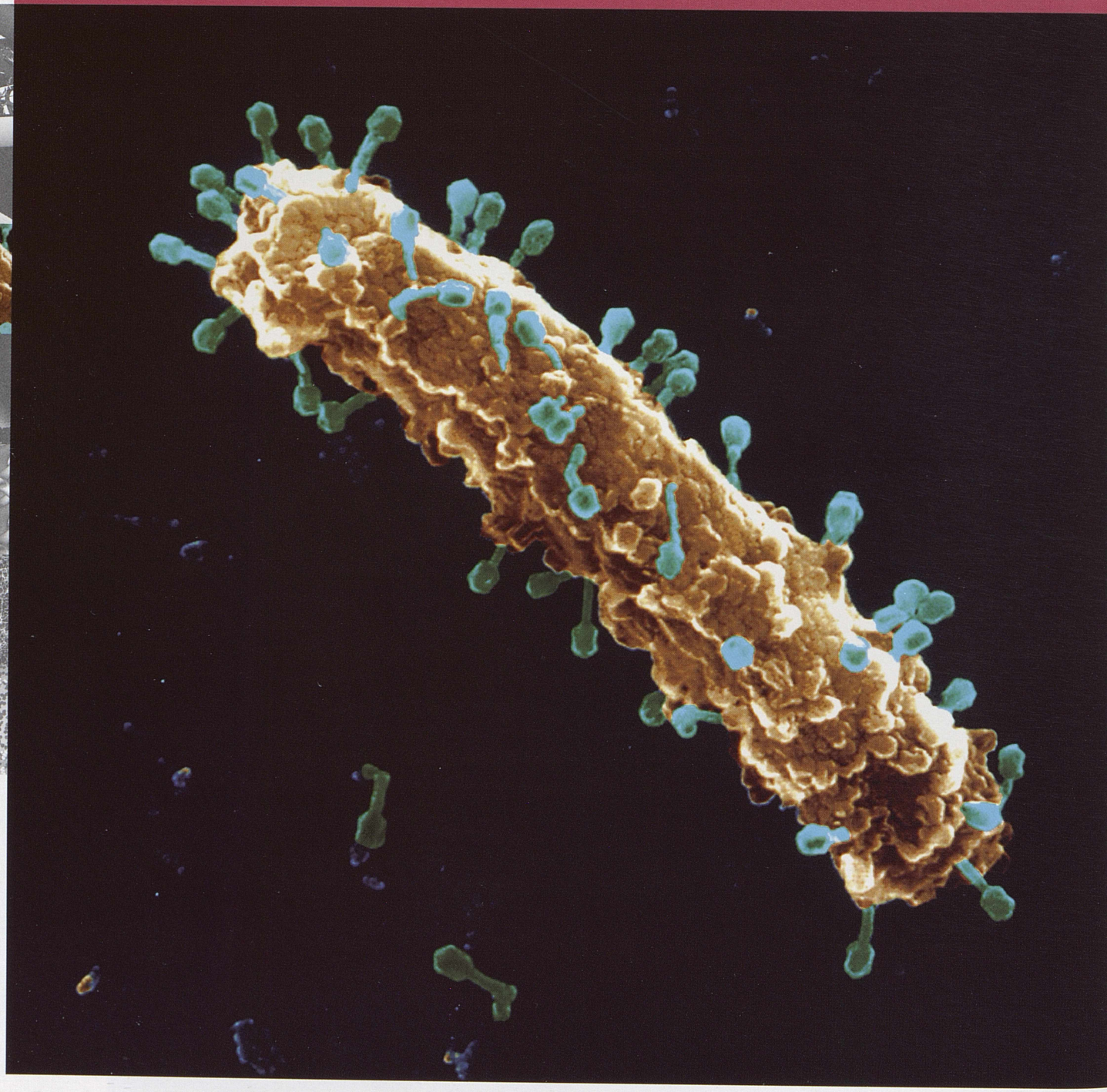
Using LeMieux, Simmerling carried out replica-exchange simulations of Trp-cage, showing that it's in its native, folded state 90 percent of the time, a very stable structure. In further studies, using up to 1,000 LeMieux processors (at 90 percent parallel efficiency), he simulated three altered versions of the protein—each with a single change of amino acid—which the Andersen group also looked at experimentally. One of these changes produced a dramatic result in the stability of the folded structure.

"There's wide interest in glycine," says Simmerling, "and how it's involved in folding. And there hasn't been data, especially with respect to how it compares with d-alanine. We think this study, relying on both experiment and simulation, will be among the first to show that not only do we have this model system, Trp-cage, that's sensitive to change, but also that simulations on computers like LeMieux are helping us to understand and begin to predict the effects that gene mutations will have on these key molecules of life."

MORE INFORMATION:

http://www.psc.edu/science/simmerling.html

THE STORY *of a* PHAGE



PSC BIOINFORMATICS TRAINING LED TO THE FULL SEQUENCING OF A LITTLE UNDERSTOOD BACTERIOPHAGE

Like a mosquito on a summer evening, a bacteriophage is either feasting or in search of its next meal. But a bacteriophage isn't interested in human blood and isn't flying around your backyard. Bacteriophages—aka phages—are microscopic killers of bacteria. Wherever you find bacteria, which is nearly everywhere, you'll find phages.

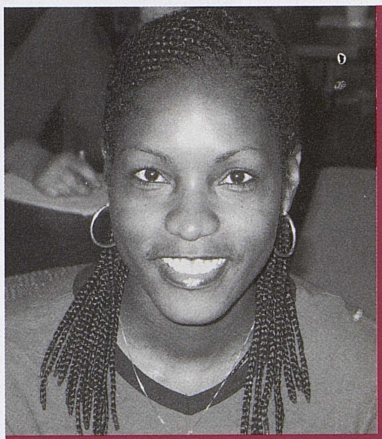
A phage is a virus, and the feasting begins when, like any virus, it attaches to its bacterium host and injects its DNA. The phage DNA hijacks the bacterium's machinery and begins to reproduce itself. Soon, the bacterium is teeming with new phages that burst forth from the bacterium, destroying it. The hunt for a new victim begins.

Not long after Canadian scientist Felix d'Herelle gave a name to viruses that infect bacteria in 1917, he recognized their potential to treat disease. Using sewage, he isolated the dysentery phage and put it in solution. After he and other doctors in a Paris hospital drank a few pints to test it, they administered their phage solution to children dying from dysentery, who were cured the next day.

D'Herelle traveled across Europe and the Soviet Union with his microscopic miracles. When Alexander Fleming stumbled on penicillin in 1928, however, the world had a magic bullet for bacterial infections. Except in a few countries, the use of phages declined into oblivion.

Today, with fast evolving, antibiotic-resistant bacteria, phages are back in the spotlight. With bioinformatics tools, researchers are seeking to understand them at the level of DNA and genes.

In recent work with PSC computational resources, and an important boost from PSC training, Aleisha Dobbins of Howard University and coworkers at the Pittsburgh Bacteriophage Institute, reported the complete genome analysis of a well-known but little understood phage. Her results—reported in the *Journal of Bacteriology* (April 2004)—reveal the entire DNA sequence and identify the genes of the SP6 bacteriophage.



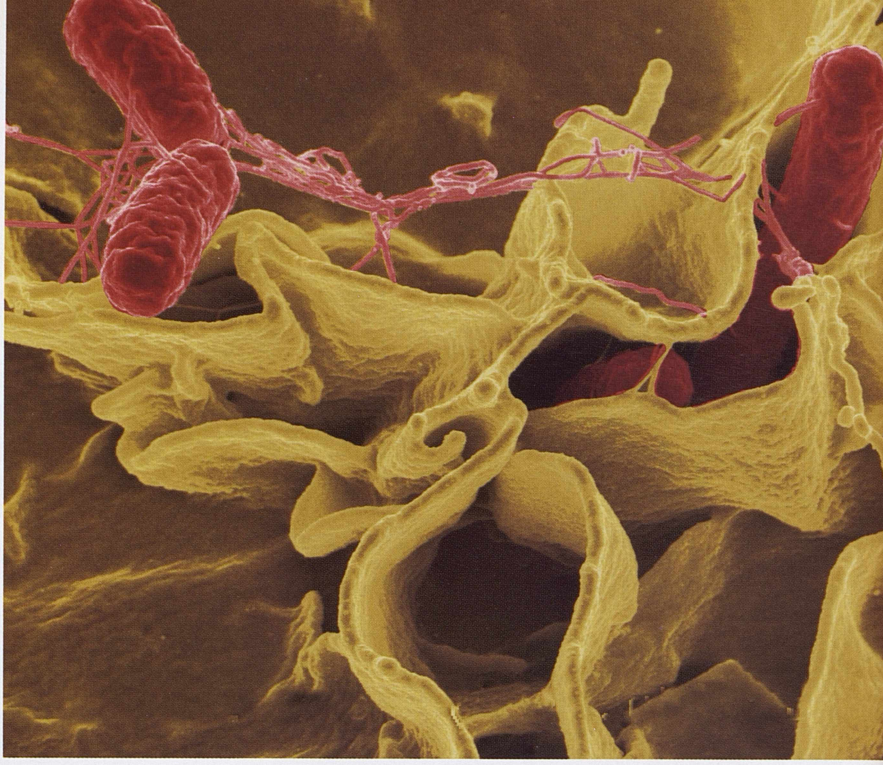
Aleisha Dobbins, Howard University

POPULOUS PARASITES

With phages, the sheer numbers are almost scary—10 million populate a milliliter of seawater, about 50 drops. Phages comprise the majority of organisms on the planet, and through the recycling of carbon in the oceans may be responsible for up to a quarter of the planet's energy turnover.

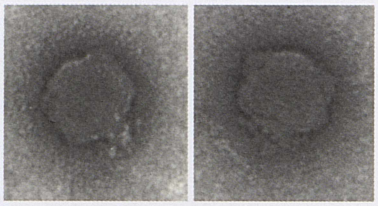
"When people hear the word 'virus,' they think trouble," says Dobbins. "But phages kill bacteria and have no effect on humans. They can be used in addition to antibiotics. With interest in phage therapy resurfacing, it's important to do sequence analysis and map the genes of more phages."

← Electron microscope image of T2 bacteriophages invading an E. coli bacterium.



An electron microscope image of the Salmonella bacterium. (Rocky Mountain Laboratories, NIAID, NIH)

Electron micrograph of two SP6 virus particles, the roughly hexagonal-shaped head is about 50 nanometers in diameter.



Research in phages is also part of the effort to defeat human viral disease. Having all the genes they need to replicate themselves, phages are similar to viruses that invade humans, but easier to study because their hosts are bacteria, not humans. If researchers learn how phages assemble their protein houses, called capsids, they may develop the means to dismantle them. Without the capsid shells, both phages and human viruses are harmless bits of floating DNA.

The SP6 phage in particular attracts attention because of its host. Phages are picky parasites. Each one invades a particular bacterium, and SP6 goes after Salmonella, the nasty bacteria that dwell in raw meat and cause food poisoning. While SP6 hasn't been yet been used to treat food poisoning, it is widely used in biotechnology.

SP6's RNA polymerase, an enzyme that transcribes DNA into RNA, is commonly used in genetic technology to modify and clone the DNA sequences of bacteria. Despite wide use, SP6 had not been sequenced and most of its genes had not been identified before Dobbins' work.

GETTING ANSWERS

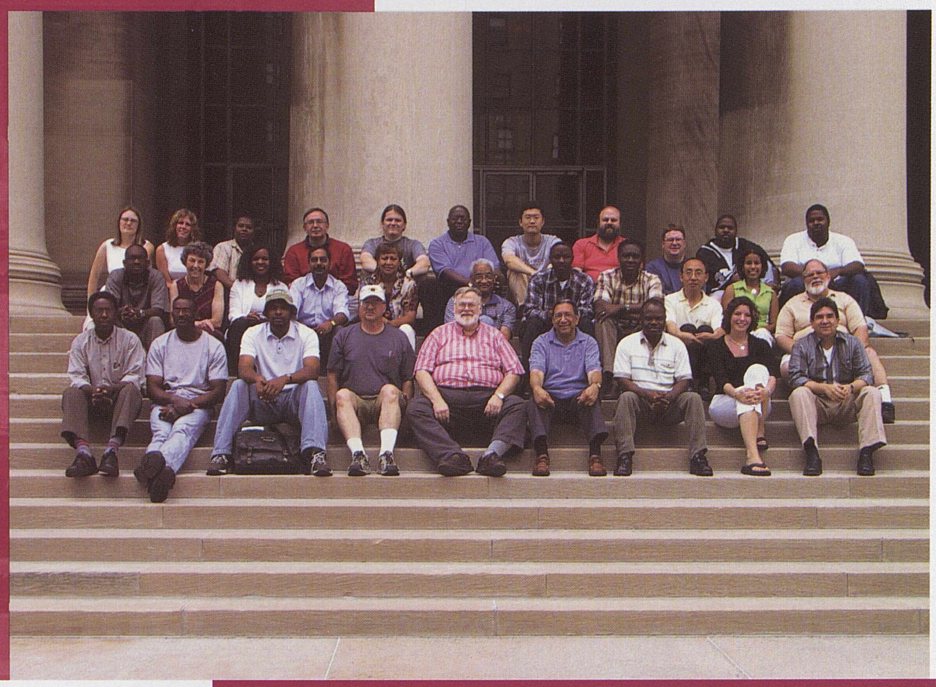
As a Ph.D student working on her dissertation, Dobbins planned to focus on one of SP6's genes. Her plans became more ambitious, however, when she went to a PSC bioinformatics workshop, led by PSC scientist and sequence-analysis expert Hugh Nicholas. Through the workshop, Dobbins gained the ability to tackle the much larger project of the entire SP6 genome.

"Through this workshop," says Dobbins, "I gained knowledge of the bioinformatics tools I needed to sequence the genome. And I learned how to use software to identify the termination sequences."

Nicholas introduced Dobbins to researchers at the Pittsburgh Bacteriophage Institute (PBI) at the University of Pittsburgh. Dobbins and PBI co-directors Roger Hendrix and Graham Hatfull decided that rather than examining one gene, it made sense to sequence and examine the entire genome. This would allow them to compare SP6 with other well-known phages and, perhaps, to draw conclusions about SP6's evolution, information that relates directly to the ability of bacteria to evolve and defeat antibiotics.

Phages and bacteria evolve in conjunction, bound together in the race to outwit one another and survive. Through this evolutionary drama, phages introduce new genes into the bacteria population. "Most human pathogens are as toxic as they are because of genes that were brought in by phages," says PBI's Hendrix. "There's a lot of interest in what this population looks like and by comparing them to each other we can start to see how the population evolved up to where it is now."

At PBI, Dobbins sequenced SP6's entire genome of over 40,000 nucleotides, the building blocks of DNA. She also identified some of the genes and their order. With training from Nicholas and the PSC workshop, she used Tourney, PSC's sequence-analysis computer, to identify the terminator sequences—regions of the genome that signal RNA polymerases to stop transcribing and disconnect. With Tourney, Dobbins also compared SP6 sequences with databases of known phage gene sequences and thereby identified SP6's genes.



From July 12 to 23, 2004, PSC hosted 19 faculty and staff from nine universities for its two-week course, "Developing Bioinformatics Programs." PSC scientists Hugh Nicholas (1st row center) and David Deerfield (2nd row right) led the course. Five interns from three universities stayed at PSC for five weeks to continue work on their research projects.

Dobbins identified SP6 as being part of the T7 phage family, which includes T7 and T3, two of the most well researched phages—a family relation that had been suspected, but not verified. Because of their similarities, Dobbins used a template of the T7 RNA polymerase, which is also used in genetic technology, to build a model of SP6's polymerase, the gene she hoped to examine in her original research plan.

Phages in the T7 family presumably evolved from the same ancestor as SP6, and have many similarities in sequence and gene placement. But there are many family mysteries. Through comparative analysis, Dobbins found that one sequence of genes appearing in the same place in most phages in the T7 family was in a much different place in SP6. The group is not only in a different place, but reversed in order. As with any research, answers spark new questions.

Dobbins' work will feed an ongoing discussion about phage evolution. While some believe that they evolved from a common ancestor millions of years ago, others argue that similar structures arose independently, or diverged more recently.

"The evolution of bacteriophages has not totally been traced," said Dobbins. "We don't know how they have evolved or what kind of effect this phage had on bacteria. We completed the sequence and found that it had 52 genes, and of the 52, 64 percent are unique to SP6. A lot of additional work needs to be done to identify the function of those genes." **CK**

MORE INFORMATION:

http://www.psc.edu/science/dobbins.html

TOOLS FOR THE JOB

The tools of bioinformatics, which marry information science and statistics with the life sciences, allow researchers to understand biological systems like never before. But researchers need to learn about these new and rapidly improving tools. PSC's "Developing Bioinformatics Programs" workshop introduces faculty and graduate students from minority-serving institutions to the computational, mathematical, and biological issues of bioinformatics.

"Bioinformatics computer programs in general involve implementing a mathematical model and comparing the model with the data to see how they relate to each other," says PSC scientist Hugh Nicholas. "The most common bioinformatics task involves taking a sequence from a biologist's laboratory and comparing it to all sequences in the database looking for relationships according to a mathematical model of sequence evolution."

The two-week workshop trains faculty who plan to establish an introductory bioinformatics course at their home institution, and graduate students, such as Aleisha Dobbins, to use bioinformatics tools to complete a research project. The course is sponsored by a grant from the Minority Access to Research Careers Branch of the Branch of Division of Minority Opportunity in Research of the National Institute of General Medical Sciences. It grew out of a bioinformatics workshop originally developed through support from NIH's National Center for Research Resources, which also supports Tourney, PSC's sequence-analysis computer, used during the workshop and by university students in courses developed through the workshop. Tourney is available to support bioinformatics course work at any U.S. academic institution.

PHAGE RESEARCH MAY HELP TO DEFEAT HUMAN VIRAL DISEASE

WEATHER_FORECASTING

re-TWISTERED TWISTER

THE MOST REALISTIC TORNADO SIMULATION EVER DONE WILL HELP REDUCE TORNADO FALSE ALARMS

Oklahoma, where the wind comes sweepin' down the plain. When singing and dancing cowboys deliver this song, you might think wind in Oklahoma is a rejuvenating experience no one should miss. Oklahomans know, however, that springtime on the plains is not a musical comedy, not when cumulus clouds roll up in dark, stormy fists, and conditions are right—or wrong—for the wind to curl into tight spirals of energy with high-speed updraft.

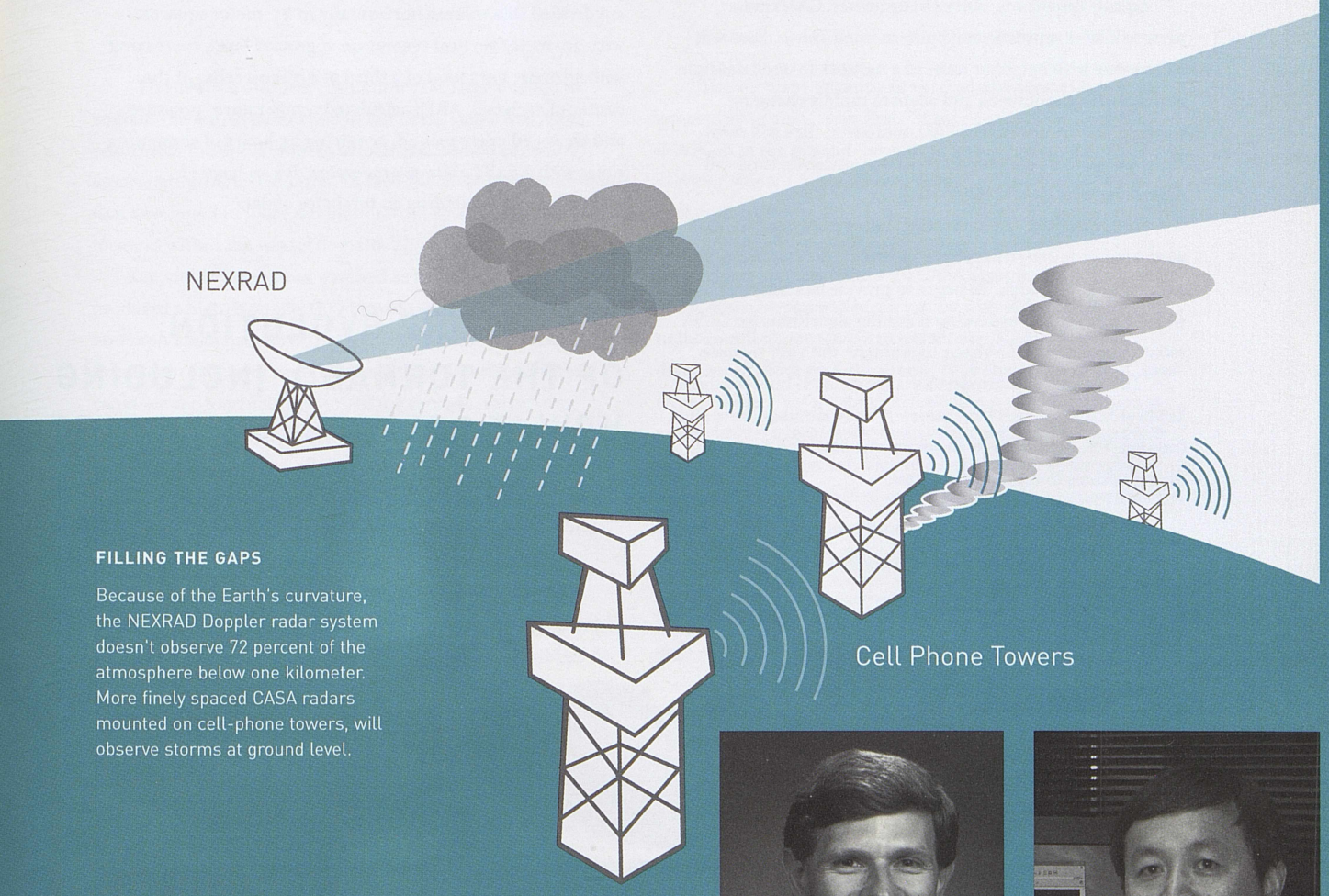
Severe storms spawn about 800 tornadoes a year in the United States, most of them on the Great Plains. The yearly toll in property loss and economic cost runs to billions of dollars, and that doesn't count an average of 1,500 injuries and 80 deaths.

Better forecasts would reduce loss and save lives, but tornados, by their nature, are notoriously hard to predict. Forecasters can identify storms with tornadic potential, but warning time with current technology is seldom more than half an hour, and the warnings are imprecise. A major problem, furthermore, is forecasts that cry wolf.

"We've been getting much better at detecting tornados and their precursor cyclones," says Kelvin Droegemeier, director of the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma, "but three out of four tornado warnings is a false alarm." Droegemeier and his CAPS colleagues have developed a storm-scale prediction capability, the Advanced Regional Prediction System (ARPS), that has set milestones in accurate forecasting of severe storms. But tornados remain a persistent challenge, and a big part of the problem is forecast data inadequate to the task.

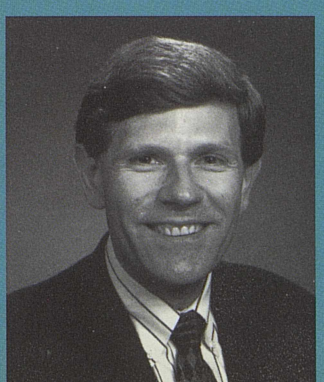
Help is on the way. To prepare for a new, more comprehensive approach to the data-gathering stage of storm forecasting, Droegemeier's colleague, Ming Xue, used LeMieux, PSC's terascale system, to carry out the largest tornado simulation ever done. With 2,048 LeMieux processors and the ARPS model, he successfully reproduced a 1977 storm and the high-intensity tornado it spawned.

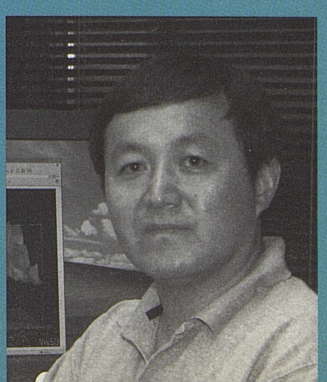
The results—which capture the tornado's vortex structure, with wind speed of 260 miles per hour—represent a watershed in tornado research. "This is the highest resolution simulation ever done," says Droegemeier, "of an entire thunderstorm and its tornado."



FILLING THE GAPS

Because of the Earth's curvature, the NEXRAD Doppler radar system doesn't observe 72 percent of the atmosphere below one kilometer. More finely spaced CASA radars mounted on cell-phone towers, will observe storms at ground level.





Kelvin Droegemeier (left) and **Ming Xue**, University of Oklahoma

GROUND LEVEL DATA

Warnings for no-show tornados may be inevitable with storm forecasting, and in large part people accept that it's much better to err on the side of caution. Nevertheless, false alarms force businesses to close and people to stay home from work or flee the area. One false-alarm study in Wichita, Kansas found an economic cost of a million dollars for one evening.

A major factor in the high false-alarm rate has been the inherent limitations of NEXRAD, the national system of Doppler radar that feeds atmospheric data to weather-forecasting computer models. One hundred forty-two NEXRAD units cover the United States, but their conical beams focus at long-range, and because of Earth's curvature, they look over the heads of tornados.

"NEXRAD sees maybe two kilometers above the ground," says Droegemeier. "They're missing a good chunk of the storm, especially near the ground, where tornados occur. That's why we have this high false-alarm rate."

As a solution to this problem, CAPS participates in a project called CASA (Center for Collaborative Adaptive Sensing of the Atmosphere). An NSF engineering research center, led by the University of Massachusetts, Amherst, CASA will deploy small, inexpensive Doppler radar units on cell phone towers. These units will aim at the ground and fill-in what NEXRAD doesn't see.

Equally important, notes Droegemeier, CASA radar units will have sophisticated built-in intelligence. Each will collaborate with neighbor units in a network to track multiple atmospheric phenomena and adapt to rapidly changing weather. In contrast to NEXRAD, moreover, they will meet the needs of multiple end users, who at any given time will determine what data is collected.

First, however, the smart CASA detectors need to learn what kind of lower atmosphere information is associated with the origin of tornados. To begin to provide this information, which CAPS will use to develop scanning algorithms for CASA radar, Xue turned to LeMieux to simulate the 1977 tornado.

"The only way we can develop algorithms to hunt for these kinds of features," says Droegemeier, "is to simulate them and then say 'OK, let's suppose these are real data' and then apply the algorithms to the simulated data."

TORNADO ON THE GRID

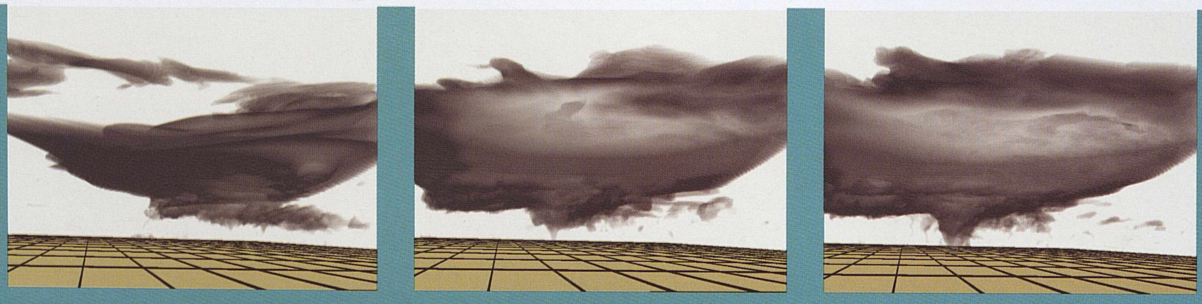
Along with the difficult problem of developing algorithms for CASA, Xue's simulation also aims at better understanding of tornado dynamics. Why does one supercell thunderstorm—an intense storm with rotating updraft—produce a tornado while another one doesn't? "We don't really know," says Droegemeier, "if there is a uniform dynamical mechanism or multiple mechanisms. Storms are much different from one another, and there's no reason to believe that all tornados form in the same way."

Computing limitations have hampered previous attempts to simulate tornados, with grid resolution held to a relatively coarse 100 meters, too porous to reliably capture a forming tornado—which have a median path width of only 46 meters. Other simulations have increased resolution locally—superimposing a finer mesh over a small area of the coarse grid. While somewhat successful, localized grids tend to skew the results, forcing the simulated tornado to form where it doesn't in reality.

With LeMieux, Xue wasn't bound by these limits. "I am basically able to run a uniform high-resolution grid that covers the entire storm system," says Xue, "as well as being able to resolve the very small tornado in a system. And there is no uncertainty about where the tornado forms."

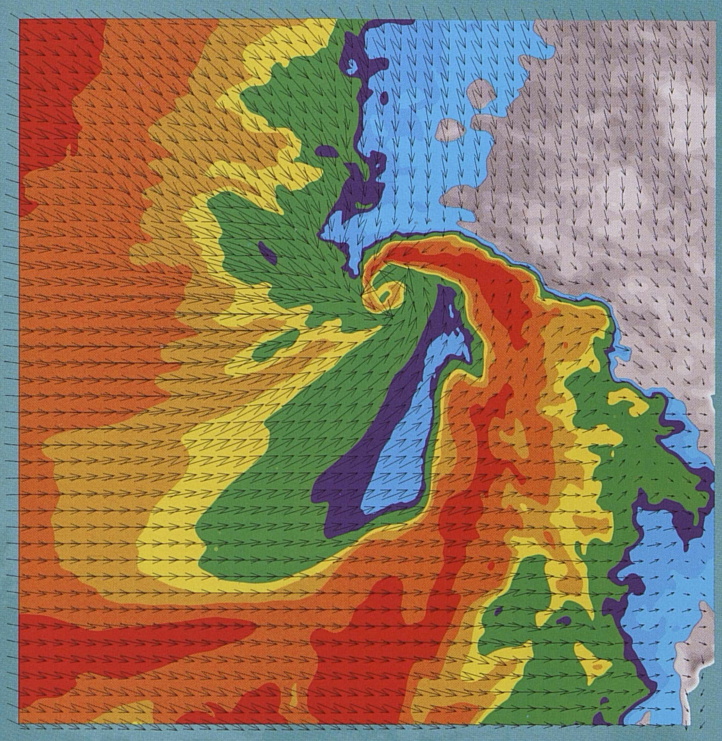
To capture the entire storm, Xue laid out an area 50 kilometers on each side to an altitude of 16 kilometers. He subdivided this volume horizontally in 25-meter squares, with 20-meter vertical resolution at ground level, increasing with altitude, for a total of a third of a billion cells. At the center of each cell, ARPS calculated temperature, pressure and air speed every second, requiring 24 hours of computing time, with 2,048 LeMieux processors, for an hour of simulated storm, yielding 20 terabytes of data.

THE ENTIRE EVOLUTION OF THE TORNADO, INCLUDING HOW IT COLLAPSED, APPEARS VERY REALISTIC



SIMULATED TWISTER

This sequence of frames (left to right) from an animation, created by PSC visualization specialist Greg Foss, show the simulated tornado as it moves within the grid volume and evolves. (Darkness of gray shading corresponds to water concentration.)



THE SPIRAL HOOK

This image shows horizontal wind-speed and direction (vectors) and radar reflectivity, which corresponds to rain-water concentration (increasing from light blue through green and yellow to red), in a horizontal plane (7 x 7 kilometers) 10 meters above ground. The spiral hook of reflectivity at the center is characteristic of tornado structure.

"The use of a uniform resolution grid large enough to contain the entire parent storm is a first," notes Xue, "and eliminates the uncertainties of artificial human control associated with nested grids. In fact, the most intense tornado that developed in these simulations did so at an unexpected location within the model domain."

Xue and Droegemeier watched as the simulation produced a high-intensity (F_5) tornado with wind speeds over 260 mph and a pressure drop of 80 millibars—the most intense tornado ever seen in simulation. Significantly, they could see the internal structure of the tornado, its twisting vortex and unstable wave patterns, and the coalescence of multiple vortices. "Basically," says Xue, "the entire evolution of the tornado, including how it collapsed, appears very realistic."

From where does a tornado get its whirling rotation? It's one of the big questions of tornado dynamics, and Xue's simulation provides a more detailed picture than has been available until now. Using the massive 3D data from ARPS, he calculated and tracked the path of air parcels in the storm. Air entering the storm, he found, undergoes an initial downdraft and flows along the ground before it feeds into the tornado vortex.

"The air turns abruptly in vertical directions in a very concentrated area," says Xue. "It carries vorticity from the direction of the ground-level flow, which becomes vertical vorticity when it turns upward abruptly. This has not been well understood."

Droegemeier expects to incorporate findings from Xue's simulations into algorithms for CASA Doppler radar by spring 2006, when he plans to begin testing a set of four of these units on cell phone towers in Oklahoma. Each will cover a diameter of about 60 kilometers. "With these new radars," says Droegemeier, "we are going to have fine-scale information to couple lower-atmosphere physics with wind and temperature structures at the ground. This will be another leap forward in numerical weather prediction. We believe CASA radars hold the promise of significantly reducing tornado false alarms—from the current 75 percent to 25 percent." (TP, MS)

MORE INFORMATION:

http://www.psc.edu/science/droegemeier2004.html

KETCHUP ON THE GRID WITH JOYSTICKS

MORE THAN 6,000 PROCESSORS AND 17 TERAFLOPS OF COMPUTING AT SIX DIFFERENT FACILITIES ON TWO CONTINENTS

Ketchup. Anticipation. You want the tasty tomato essence to flow freely from the mouth of the bottle. Turn it upside down and wait. Nothing happens. Shake. Shake again.

This classic ketchup flow problem is one among many applications of research by an Anglo-American team of scientists interested in little understood phenomena that occur in liquid mixtures, like ketchup, when they act like solids. They developed an innovative and powerful computational approach that allows them—with big help from the most advanced computing resources available—to simulate these mixtures, including what happens when you apply force. Shake again, harder, and glop, the ketchup remembers its liquid nature and deposits about six times as much tomatoey essence as you need.

In work dubbed the TeraGyroid Project, this Boston-U.K. team pushed research in their field to where it hasn't gone before, and they've done it using the Grid, a much talked-about but still new style of scientific computing that links resources without regard to location. During SC2003 in Phoenix, TeraGyroid tied together more than 6,000 processors and 17 teraflops of computing at six different facilities on two continents. With this ability to flex the young muscles of Grid technology—including the National Science Foundation's TeraGrid—they've arrived at important scientific results that couldn't have been accomplished by now except for their Grid-based approach.

A PITTSBURGH CLASSIC

The Heinz octagonal, fluted, long-necked bottle with the keystone label (and the pickle logo) carries worldwide visual identity as the real thing in ketchup. In 1876, H.J. Heinz took a stand on quality—a Pittsburgh trait—and introduced Heinz ketchup in a clear glass jar, so consumers could see what they were buying. If the ketchup sticks, Heinz recommends a firm smack to the embossed "57" on the neck.

"Because of the Grid," says Bruce Boghosian of Tufts University, "we made progress in three months that would have taken more than a year by conventional methods." Boghosian, a professor of mathematics, leads the stateside contingent of the TeraGyroid effort.

"The Grid creates a tremendously powerful environment," says University of London chemist Peter Coveney, who leads TeraGyroid on the U.K. side. Using the Grid and a sophisticated approach called computational steering, the scientists "steer" their simulations in real time, so that the heavy-duty computing can focus where the dynamics are most interesting. As a result, says Coveney, "our productivity skyrocketed."

The TeraGyroid group's Grid sophistication received a 2004 ISC Award, the major supercomputing award in Europe, for "Integrated Data and Information Management." At SC2003 in Phoenix last November, TeraGyroid was recognized as the "Most Innovative Data-Intensive Application," which understates what is probably the most impressive feat yet of wide-based Grid computing. Linking the TeraGrid with the U.K. E-Science Grid via dedicated trans-Atlantic fiber, the intercontinental team relied on U.K. resources at Daresbury Lab and Manchester along with computing, storage and visualization facilities at four TeraGrid sites: PSC, NCSA, SDSC and Argonne.

Their simulations rely on an approach called the lattice-Boltzmann (LB) method, and—with PSC's LeMieux—they performed the largest LB simulation to date. These November computations were prepared by two months of work at U.S. and U.K. facilities and continued with follow-up computations into early February 2004, work which in sum yielded three terabytes of useful data.

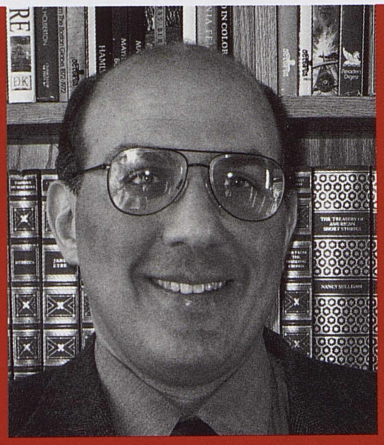
To gather and collate this quantity of data from multiple sites was itself a large task. Their analysis points to new understanding of the somewhat bizarre liquid crystalline materials they are studying, and a strange shape known to mathematicians as a gyroid.

AMPHIPHILES & GYROIDS

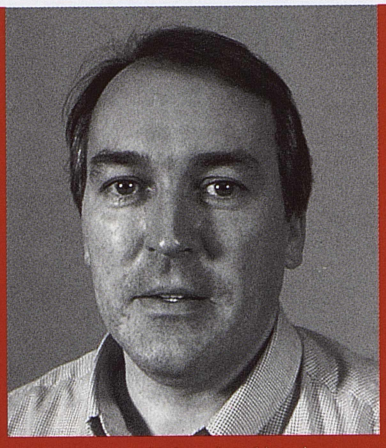
We've all noticed the line between oil and vinegar in a bottle of salad dressing. In chemical terms, the two liquids are immiscible. Add a third substance called an amphiphile to this kind of mix and interesting things happen. An amphiphile, by definition, is a chemical species that goes both ways. In vinegar and oil, for instance, one end of the molecule likes vinegar, the other oil.

Naturally, the amphiphile migrates to the interface between the two fluids, where both ends are happy. And if you add more and more amphiphiles, eventually there's no space at the interface. If you then keep adding amphiphiles, they work hard chemically to create more interface area, which they do by twisting and rippling the surface until, eventually, it breaks apart into oil-vinegar droplets of various contorted shapes that maximize the interface area.

Some of these contorted shapes—called mesophases— have features that resemble solid materials, hence the stuck-in-the-bottle consistency of ketchup. They also lead to many industrial applications. Soap, detergent and shampoo, for instance, contain surfactants, another name for amphiphiles. Sophisticated chemical use of amphiphilic fluids has led to high-gloss liquid waxes and environmentally-friendly dry-cleaning solvents.



Bruce Boghosian, Tufts University



Peter Coveney, University of London

ABOUT 30 ORGANIZATIONS PLAYED A PART AND PROBABLY MORE THAN 100 INDIVIDUALS

One mesophase in particular, a unique shape called a gyroid, hence TeraGyroid, is the focus of the team's recent work. This convoluted and, at the same time, regularly patterned shape appears widely in biological systems. The endoplasmic reticulum, for instance, the organelle that manufactures proteins inside the cells of plants and animals, is gyroidal in structure, and gyroids have important applications in controlled drug release and biosensors.

One of the ways that a gyroid resembles the lattice like, crystalline structure of a solid is that defects occur, disruptions to the regular pattern. These defects have a large effect on the mechanical properties of the fluid, such as rigidity—as when ketchup is stubborn. The objective of the TeraGyroid simulations at SC2003 was to look at these defects for the first time. To address questions such as: Under what conditions do defects appear? How do they change when force is applied?

"Dislocations in liquid crystals aren't well studied," says Boghosian, "how they develop at various temperatures, how they tend to move, how they behave under shear."

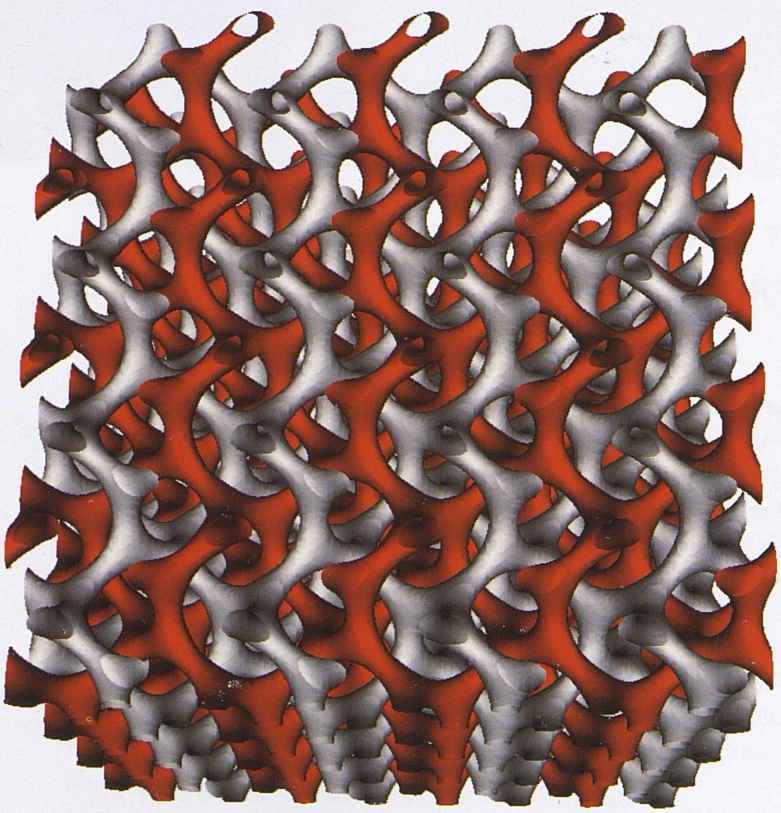
JOYSTICKING ON THE TERAGRID

In the middle ground between modeling each atom of a molecule and the high ground of computational fluid dynamics, the lattice-Boltzmann model maps fluids as particles on a 3D lattice and follows their mass and momentum changes over time. Since the 1990s, the LB method has proven itself to be both highly efficient on parallel systems (linear scaling on all platforms on which it runs) and an accurate method for modeling the dynamical properties of mixed fluids.

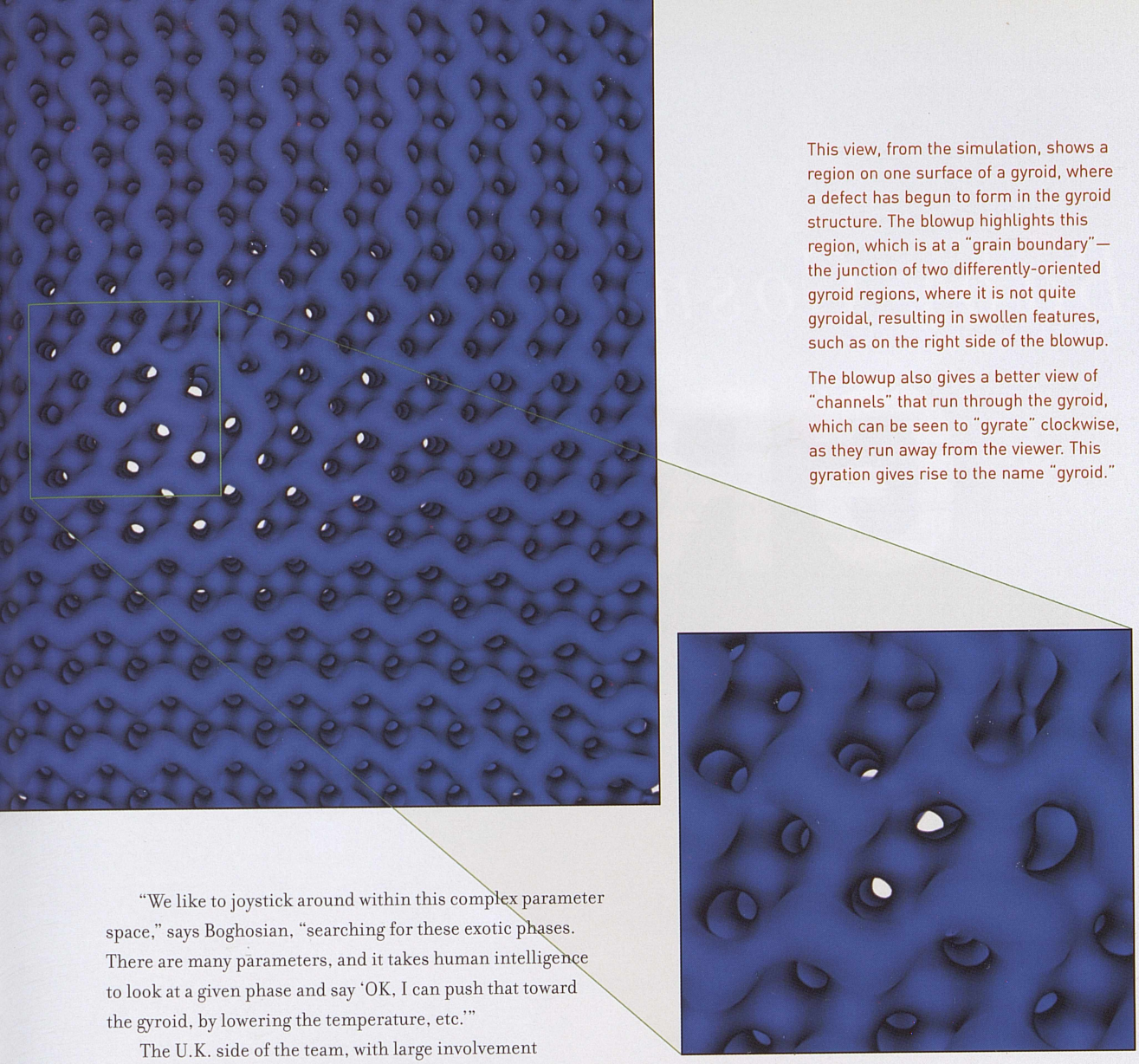
In the late 90s, Boghosian and Coveney developed the first LB code capable of modeling "ternary amphiphilic fluids" such as oil, water and surfactant. With this code, called LB3D, Coveney and Ph.D student Nelido González-Segredo in 2003 for the first time "captured" the self-assembly of a gyroid mesophase in an amphiphilic fluid, a result that caught the attention of other scientists and laid the groundwork for the TeraGyroid Project.

To go beyond being able to see a gyroid to looking at gyroid defects as they form and change demanded both a much grander scale of simulation and—to use the resources efficiently—the sophistication of computational steering.

The gyroid is elusive, a mesophase that forms only under a narrow and not yet well understood set of conditions. The LB model includes many parameters—temperature, relative concentrations, how much the two immiscible fluids "dislike" each other, how strongly the surfactant interacts with each of them, and others—all of which affect whether a gyroid forms and the nature of its defects. With computational steering, the scientists are able to visualize a simulation in progress, make decisions to change how it evolves, and zero in on what they want to see.



One view, called the wishbone view, of a gyroid structure, which divides space into two interpenetrating regions or labyrinths.



This view, from the simulation, shows a region on one surface of a gyroid, where a defect has begun to form in the gyroid structure. The blowup highlights this region, which is at a "grain boundary"—the junction of two differently-oriented gyroid regions, where it is not quite gyroidal, resulting in swollen features, such as on the right side of the blowup.

The blowup also gives a better view of "channels" that run through the gyroid, which can be seen to "gyrate" clockwise, as they run away from the viewer. This gyration gives rise to the name "gyroid."

"We like to joystick around within this complex parameter space," says Boghosian, "searching for these exotic phases. There are many parameters, and it takes human intelligence to look at a given phase and say 'OK, I can push that toward the gyroid, by lowering the temperature, etc.'"

The U.K. side of the team, with large involvement of Ph.D. student Jonathan Chin, carried most of the load in Grid-enabling LB3D. Remote researchers at Tufts (grad student Lucas Finn) and at British Telecommunications, Ipswich, U.K. collaboratively steered the simulation. Coveney emphasizes that the vision of the Grid—"transparent" access to resources without regard to location—is not yet reality, and it was a large task to coordinate this ambitious project. "About 30 organizations played a part," he says, "and probably more than 100 individuals. As one example, we had to negotiate agreements between six different certificate authorities to use all these intercontinental resources."

Once the steering team zeroed in on a defect phase, the researchers enlarged the simulation—from two million lattice sites (128^3) on smaller runs to more than a billion sites ($1{,}024^3$)—using LeMieux. "The idea of Grid computing," says Coveney, "is to be able to look at problems that haven't been accessible before. To capture these defects, you have to simulate extremely large amounts of material, and this takes you into the terascale domain. LeMieux is the only machine we have access to that could accommodate these extremely large simulations."

The results show—among other things—that ternary amphiphilic fluids are non-Newtonian. In other words, they behave like ketchup, not like water. The viscosity depends on the shear rate. At rest, the fluid reacts to external force like a solid. Shake it vigorously, and it flows like a liquid. "We've been able to confirm," says Coveney, "that our gyroid behaves like a shear-thinning, non-Newtonian fluid."

"Being able to do these simulations," says Boghosian, "allows us to study the dynamics of the fluid in more detail than is currently possible by experiment." As this research develops better understanding of these phenomena, it will undoubtedly lead to useful products and technologies impossible to anticipate —who knows, maybe someday a tasty ketchup that flows from the bottle smoothly and easily. **MS**

MORE INFORMATION:

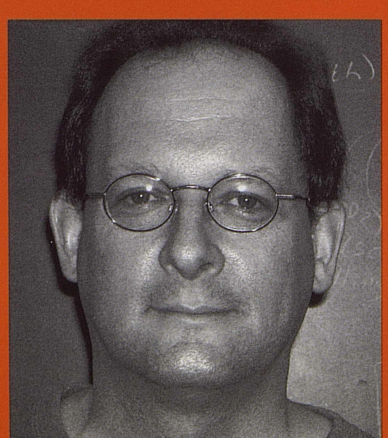
http://www.psc.edu/science/teragyroid.html

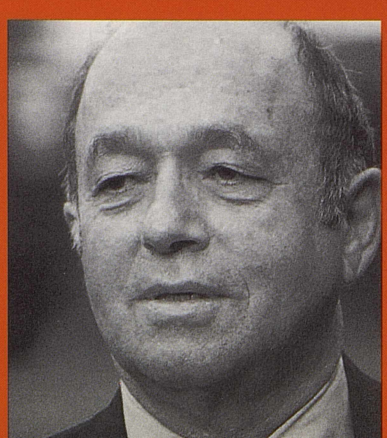
EVOLUTION_AND_STRUCTURE_OF_THE_UNIVERSE

Baby Cosmos GROWS

THE LARGEST SIMULATION OF THE UNIVERSE EVER DONE SHOWS THE STRUCTURE OF DARK MATTER

Paul Bode (left) and **Jeremiah Ostriker**,
Princeton University





UP

The universe may have begun with a bang, but the images that reach us from 379,000 years after that singular instant 13 billion years ago present a fairly mundane picture. The most notable characteristic is uniformity. Over immense distances, the temperature of the unimaginably hot matter spread evenly through the early universe fluctuated by mere thousandths of a degree. Yet those tiny fluctuations generated the diverse splendor of the galaxies, nebulas, stars and planets we see today.

Two years ago, a satellite—the Wilkinson Microwave Anisotropy Probe (WMAP)—captured the first light that escaped from that hot, uniform early time, providing astronomers with a baby picture of the universe. With each passing month, sky surveys and x-ray observatories add more details to fill in the gaps between then and now. And these observations are only the beginning. In the coming decade, a new wave of missions promises deeper, sharper views into early periods of structure formation.

It's an exciting time for astrophysicists, with one question uppermost in their minds. How well will the new information match up with theories about formation and evolution of the universe? If gravity is the primary force sculpting the heavens, as theories predict, then what structural features should astronomers expect to find, if they look in the right places, within the huge forest of emerging data?

Large-scale computational simulations play an indispensable role bridging the gap between theory and reality in our burgeoning knowledge of the cosmos. To help narrow that gap, astrophysicists Paul Bode and Jeremiah Ostriker of Princeton University used LeMieux to carry out the largest simulation of the universe to date. Starting with the baby picture from WMAP, and depicting the universe with unprecedented detail, they harnessed LeMieux's parallel-processing power to evolve the baby cosmos forward to the present.

Unlike most simulations of cosmic structure, which start with a section of the universe and look only at the end result, Bode and Ostriker assembled a photo album with which to view the universe as it grows up. Designed to facilitate comparison with observations, their album presents the universe as an Earth observer sees it. "We end up producing a virtual night sky," says Ostriker, "which anyone can then study in a computer." With analysis still underway, they've already turned up hints of some as-yet unconfirmed characteristics of the early universe.

COMPUTING IN THE DARK

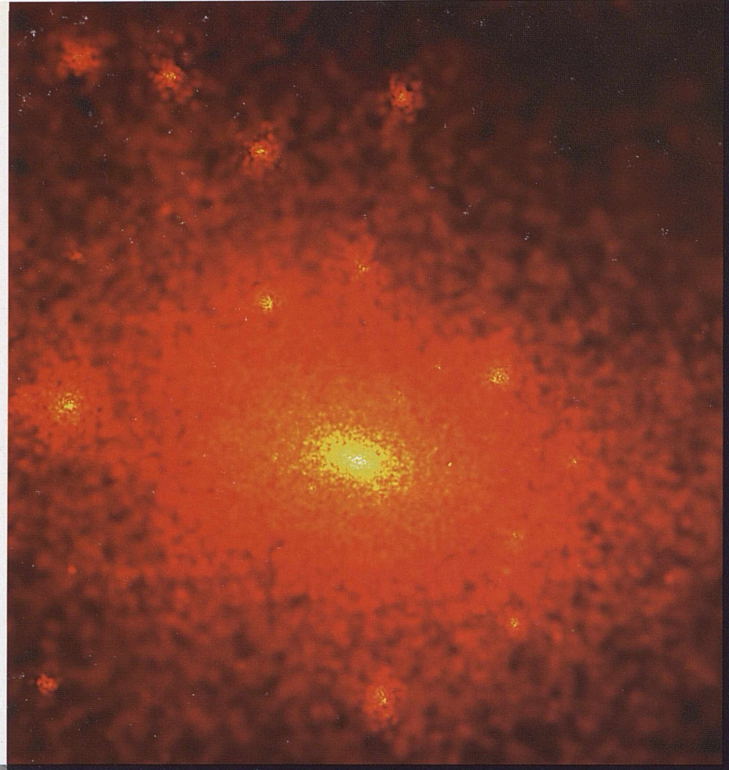
With big help from LeMieux, Bode and Ostriker populated their universe with two billion virtual particles—each the size of several galaxies—twice as much granular detail as the most ambitious similar simulations. As a concession to computational economy, however, their simulation takes place in the dark. The virtual universe contains no flowing gases and igniting stars.

All two billion particles represent dark matter—a mysterious type of mass we cannot see. These particles, which attract each other, are also interacting with a still more inscrutable, gravity-less component that makes up about 73 percent of the energy and mass balance of the universe, so-called dark energy, which scientists theorize tries to push space and everything in it apart.

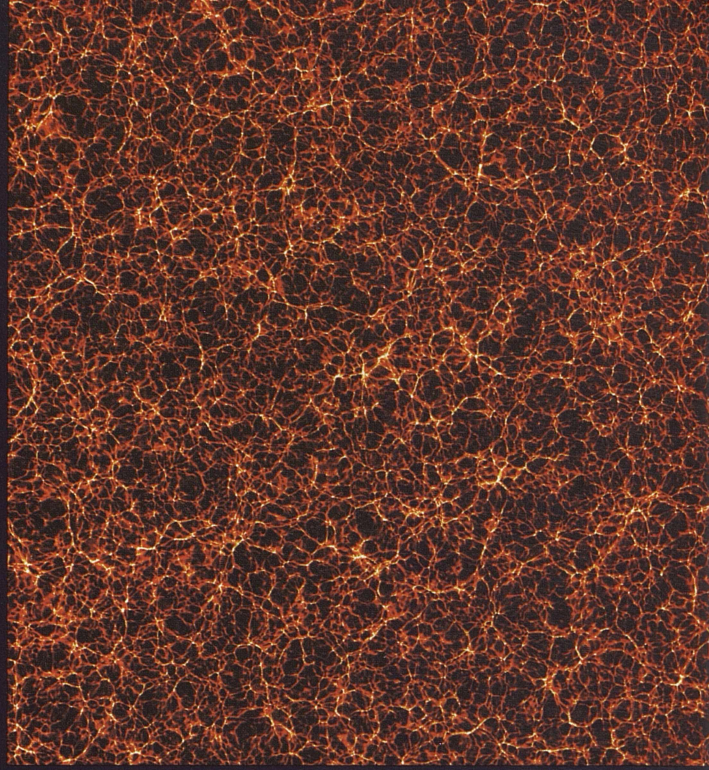
Tracking the interactions of two billion particles over 13 billion years to build a virtual model of the universe presents a large computational challenge. "It's just at the edge of what you can do," says Ostriker, "that's why you need the biggest supercomputers."

LeMieux's combination of number-crunching power and storage capacity provided the combination needed to compute the position of the particles and store their arrangement through time. "It's the whole package, really," says Bode, "lots of processors and lots of memory, lots of disc storage as well."

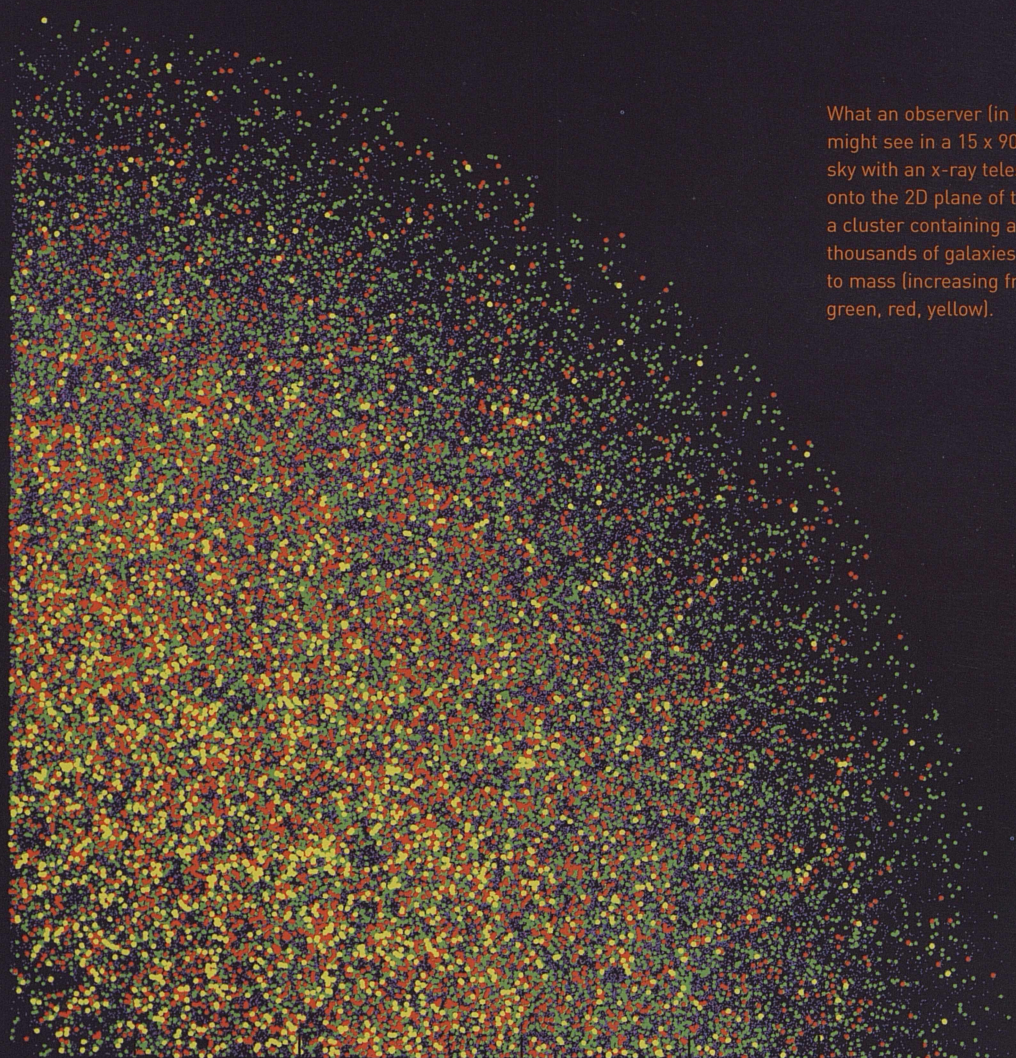
Even with these computing resources, however, modeling the gravitational landscape shaped by two billion dark-matter particles depends on software ingenuity. Gravity acts over long distances, and every particle shapes the gravity that acts on every other particle. To take advantage of parallel processing, particles are parceled out to different processors, and the need to calculate the force exerted by the particles at one processor on particles elsewhere can create an intra-processor traffic jam of messages.



Closeup of a large dark-matter halo, about six million light years on each side. Brightness corresponds to density.



This thin-slice snapshot through the simulation volume, about 3 million light years thick by 4.5 billion light years on each side, shows the filamental structure of dark-matter clusters. Brightness corresponds to density.



What an observer (in lower-left corner) might see in a 15 x 90 degree wedge of the sky with an x-ray telescope (as projected onto the 2D plane of the page). Each dot is a cluster containing anywhere from 100 to thousands of galaxies, color corresponding to mass (increasing from violet to blue, green, red, yellow).

LIKE WATER IN A RIVER VALLEY, DARK MATTER POOLS INTO FILAMENTS

"You have to figure out a way to avoid spending all your time passing messages around," says Bode. The solution, first developed by former Princeton graduate student Guohong Xu, and continually modified and refined by Bode, splits the force affecting each particle into two parts, a long-range part that comprises the effect of all particles and a short-range part that accounts for the gyrations of the particle's neighbors.

The software implementing this algorithm, called Tree-Particle-Mesh, made efficient use (90 percent scaling) of 420 LeMieux processors, and with five days of computing built the virtual dark-matter universe.

COLD DARK WATER IN THE VALLEYS

Much more than cold interstellar dust, black holes, and dark, dead stars, the exact nature of most dark matter is unknown. Scientists suspect, however, that dark matter makes up about 24 percent of the universe's mass and energy and exerts gravitational force. Luminous matter contributes only 3 percent, meaning that the gravitational landscape of the universe is defined largely by dark matter. In the Cold Dark Matter theory, which Bode and Ostriker implemented on LeMieux, this means that dozens, hundreds, sometimes even thousands of galaxies cluster in clumps of dark matter, called halos.

"If we can track all of the dark matter," says Ostriker, "then we have a good picture of the structure within which the galaxies find themselves. We take what we think is the right model of cosmology, we put in the initial ingredients — which are basically the fluctuations that have been seen by the WMAP satellite—then we turn the crank on the computer and allow gravity to act with these little ripples. We find dark matter accumulating into halos and more massive halos. And they have substructure and merge and do all sorts of wonderful things."

Many of the photos from this virtual album will provide key points for comparison with observations. Because galaxies are packed inside dark matter and carried along by the speed of the dark-matter halo surrounding them, for instance, it's possible to compare with observational data on galaxy velocity. Ostriker and one of his students are cataloging the speed of dark-matter clusters from the simulation to see how this velocity distribution compares with the speed of galaxies astronomers are cataloging from observations.

With a working assumption that galaxy structures are influenced by the dark matter that envelops them, Bode has tracked the evolving shape of the largest clusters of dark matter in the simulation. In early periods of structure formation, Bode found that clusters were more aligned and elongated than expected, supporting the idea that matter pooled into strung-out filaments, much as water migrates to and flows down the center of a river valley. This effect is more striking than expected, says Bode, and as observations of large galaxy clusters at earlier times come in, it will be interesting to see how well the simulation matches up.

Bode is also looking forward to comparing the number of giant clusters of dark matter in the simulation with the number of galaxy swarms in the real universe. If the simulation's mass density—a key theoretical parameter that describes how closely mass was packed in the beginning—is larger than in the real universe, the simulation clusters would come together faster and form larger clusters than in the universe. If the simulation is off in the other direction, it will have fewer giant clusters than the real universe.

Bode and Ostriker are also using the gravitational potential of the dark matter distribution to calculate the temperature of gas within dark-matter clusters. "It's an imperfect connection," says Ostriker, "but right now it's the best tool we have. Until now we didn't even have that option because we couldn't make simulations of anything on a big enough scale to compare to the real universe. We could only do little pieces." With dark matter particles and LeMieux, it was possible to do a much larger section of the universe, one of the largest volumes of space ever simulated. "These simulations enable you to look all the way back through space to the beginning."

What if the simulation doesn't match up with observations? That's the beauty of computational simulations, says Ostriker. They make it possible to systematically test and adjust theory. "We can then do another simulation, with a different cosmology. We'll increase dark-matter content, or we'll change the dark-energy content, because in fact we don't know these quantities very well." As observations become more detailed and simulations more accurate in representing theory, science will move step-by-step, says Ostriker, toward knowing what initial features went into creating the universe. **KG**

MORE INFORMATION:

http://www.psc.edu/science/bode.html

PSC SCIENTISTS HAVE FOUND A WAY TO DELIVER HDTV-QUALITY STEREO VIEWING AT REASONABLE COST

SEEING DOUBLE *at the* MOVIES

If you're a baby boomer, you may remember watching horror movies with "3D glasses." This novelty approach to creating a visual sensation of depth was one of the first movie forays into stereo viewing. Since then, with stereo goggles for computer gaming and—at the high-end of quality and expense—with supercomputing applications in science and engineering, the technology has improved. But you still can't go to the movies and see good 3D, say PSC scientists, who have come up with an approach that may fill the void in theaters, meeting rooms and at home.

The key is low cost and high performance, says Joel Stiles, a medical doctor, physiologist and neuroscientist, who along with colleague Stuart Pomerantz, a programming expert, created the new system, called the PSC Stereo Animation System (PSC-SAS). "If you make good stereo content and have a good display system," says Stiles, "stereo viewing works fabulously. But it's tremendously underutilized, because few people have put all the pieces together in an optimum way. Our system provides theater-quality stereo viewing of complex animations at extremely high performance-cost ratio."

With PSC-SAS, images display on a screen, as in normal movie viewing, and the viewer wears light, comfortable glasses, like polarized sunglasses, making the system easily adaptable to in-theater or home viewing. Many stereo-viewing systems, in contrast, rely on special goggles, connected to a computer that can be used by only one person at a time while looking at a computer monitor.

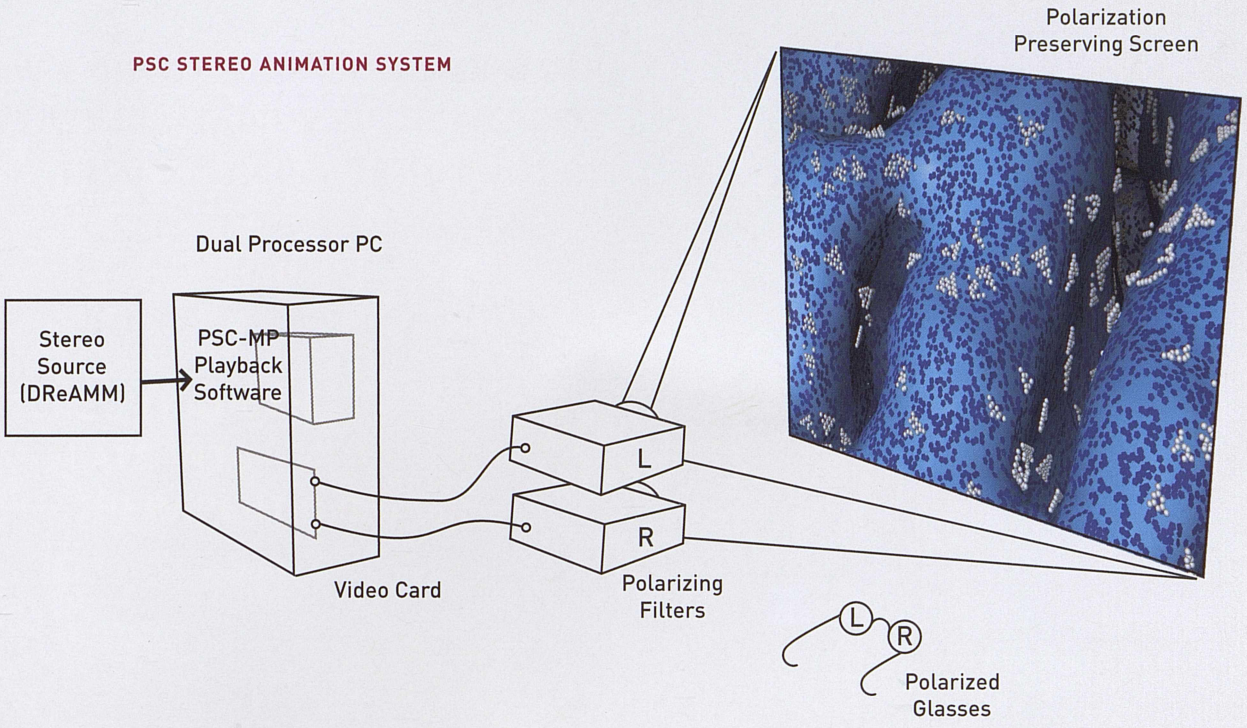
To date, PSC-SAS has been used mainly in scientific settings, for 3D display of dynamic data, compiled as movies, from computational simulations of biomolecules, cellular physiology and other applications, where being able to see depth enhances the ability to understand and analyze complex phenomena. The reaction among scientists, say Stiles and Pomerantz, has been remarkable.

"Many scientists and other professionals who have seen various types of stereo display," says Stiles, "are jaded. They think it sort of works, but it's basically a toy. When they see our system, they say 'wow' and they're reaching out to touch what they see."

UPDATING AN OLD TRICK

All stereo-viewing systems achieve the effect of depth by displaying a slightly different image to each eye. PSC-SAS is distinctive in providing high-quality content in movie form with "passive" stereo display technology. Most scientific systems and computer games are "active" stereo, relying on goggles that are, in effect, shutter glasses, electronically switching between the right and left eye at a rate faster than the wearer perceives.

Active stereo can provide high-quality effects, but it has several disadvantages. Prominently, it may present a health and safety problem. In some people, the rapid on-off flashing seen by each eye may become uncomfortable, and flashing lights can sometimes trigger an epileptic reaction. Active stereo, furthermore, is prohibitive for theater viewing, because of the cost and inconvenience of providing computerized goggles to everyone in the theater. A more recent technology—"glasses-free" stereo on computer monitors—offers limited resolution and requires the viewer's head to remain in a particular location.



With PSC-SAS, dual projectors display a right and left-eye image on the screen simultaneously, overlaid on each other—based on the well understood phenomenon of polarized light—so that one image is polarized at a 90 degree angle to the other. Polarized glasses allow the left and right eye to perceive the two distinct images separately. With this approach, many viewers at one time can see stereo depth.

"One lens is polarized in one direction," says Pomerantz, "and the other in the opposite direction. As long as the filters on the projectors match the filters on the glasses, you can deliver one image to the right eye and another to the left. It's an old trick."

PSC-SAS implements the old trick with stereo-movie content created by software called DReAMM, developed by Stiles, coupled to playback software called PSC-MP, developed by Pomerantz. To accommodate the high resolution of scientific images, PSC-SAS relies on sophisticated compression techniques that reduce file size, but only to a degree that the eye can't detect. PSC-MP delivers the polarized images to the dual projectors in synchrony at high realism. It decodes and transmits data at 100 megabits per second, 20 times faster than DVD data rates, for high-definition quality at 30 frames per second.

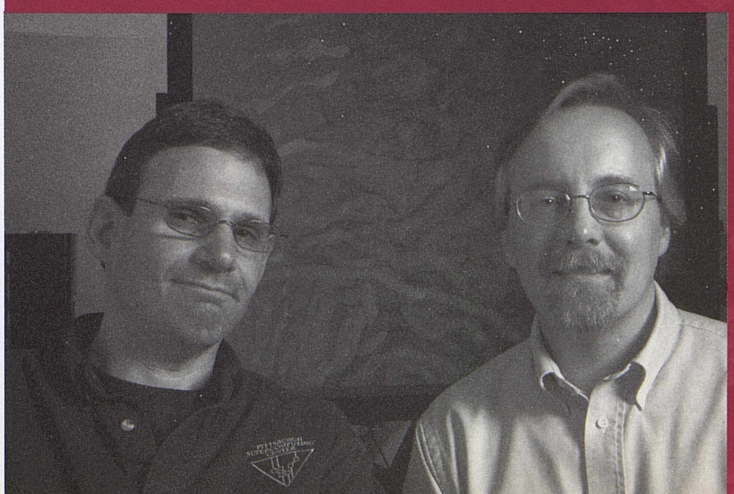
The result—vivid color and sharp, unpixelated images without uncomfortable, unsafe goggles—also comes at reasonable cost. Stiles estimates a total expense of $12,000 for the hardware components of PSC-SAS, available off-the-shelf, easily within the range of today's home theater market. A non-depolarizing screen, two computer projectors, a dual-processor PC, and a pair of polarized glasses—bring your own popcorn. **MS**

PSC-SAS is available for licensing and commercialization through the Carnegie Mellon University Innovation Transfer Center.

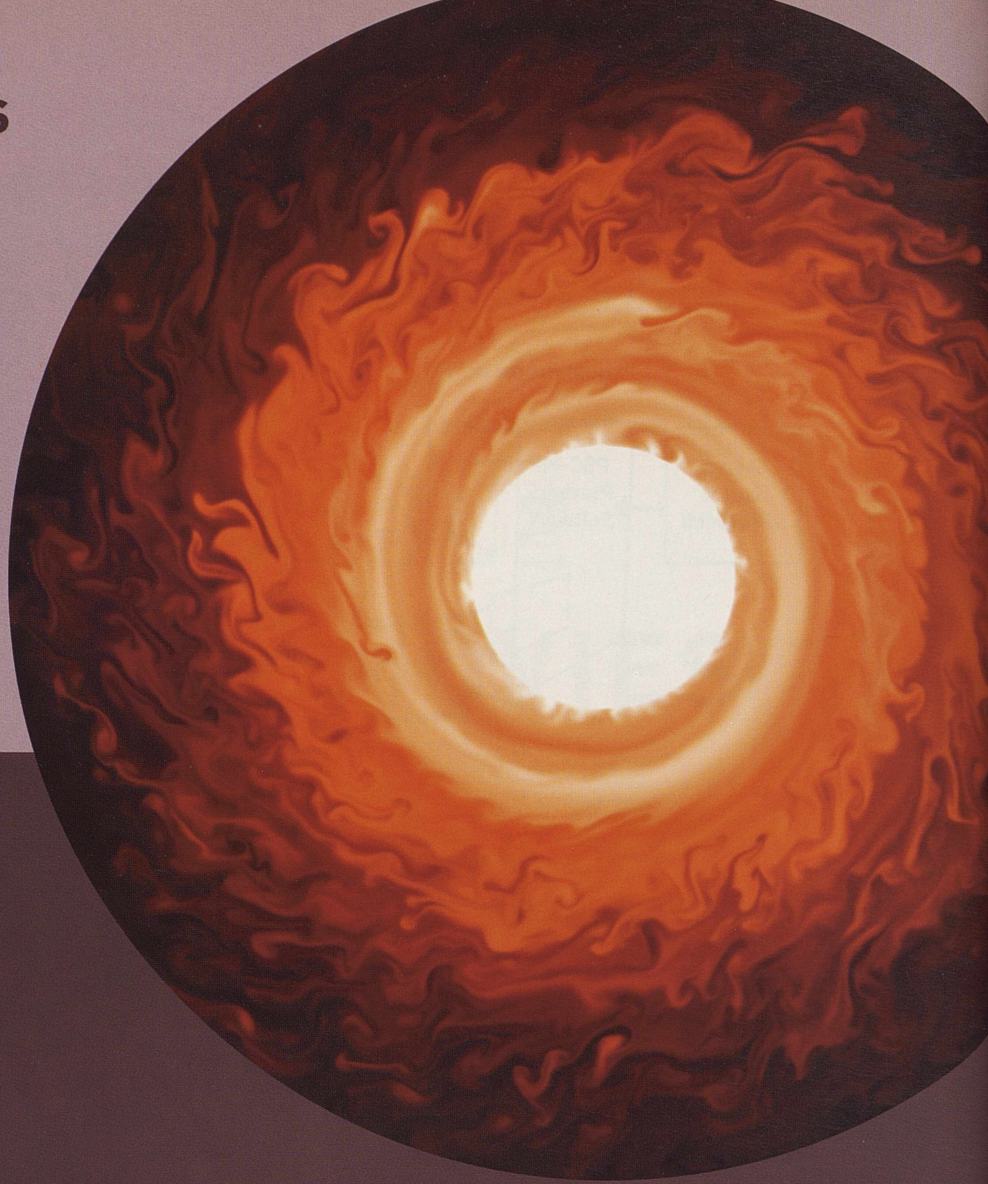
MORE INFORMATION:

http://www.psc.edu/science/stereo2004.html

Joel Stiles (right) and **Stuart Pomerantz,**
Pittsburgh Supercomputing Center



In Progress



CONVECTION IN GIANT PLANETS

WATCHING THE TURBULENT ROILING OF FLUIDS IN A PLANET'S INTERIOR

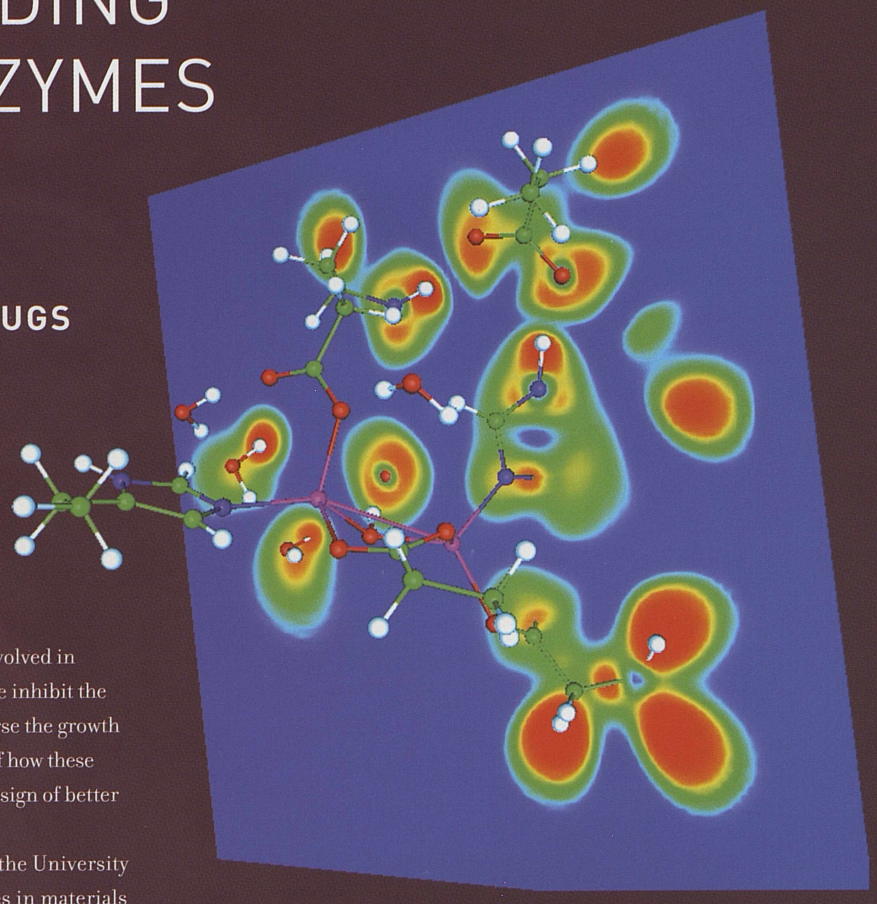
Scientists know that much as Earth generates its magnetic field from the geodynamo— turbulent, roiling motion of fluids in the planet's interior, giant planets like Jupiter and Saturn also have a magnetic field produced by turbulent fluid convection in the interior. For these giant planets, furthermore, this interior fluid motion is also related to zones of differential rotation on the planet's surface. Although scientists have studied these phenomena for decades, we still lack detailed understanding.

Gary Glatzmaier of the University of California, Santa Cruz developed the first computational model of Earth's geodynamo that evolves on its own—self consistently. In 1995, running at PSC, this model produced the first simulation of convection and magnetic-field generation in the Earth's fluid core, and the first simulated magnetic-field reversal, a phenomenon that has happened many times over Earth's geological history.

In recent work with LeMieux, Glatzmaier and graduate student Martha Evonuk simulated rotating convection in Jupiter's interior. This snapshot from their simulation represents temperature on a slice through the planet (brightness corresponding to temperature). It shows small-scale turbulent convection within a large-scale spiral pattern. "Longitudinal flow near the outer boundary is oppositely directed to that near the inner boundary," notes Glatzmaier, "so the spiral continues to wind up."

UNDERSTANDING METALLOENZYMES

PICTURES THAT CAN TRANSLATE INTO CHEMOTHERAPEUTIC DRUGS

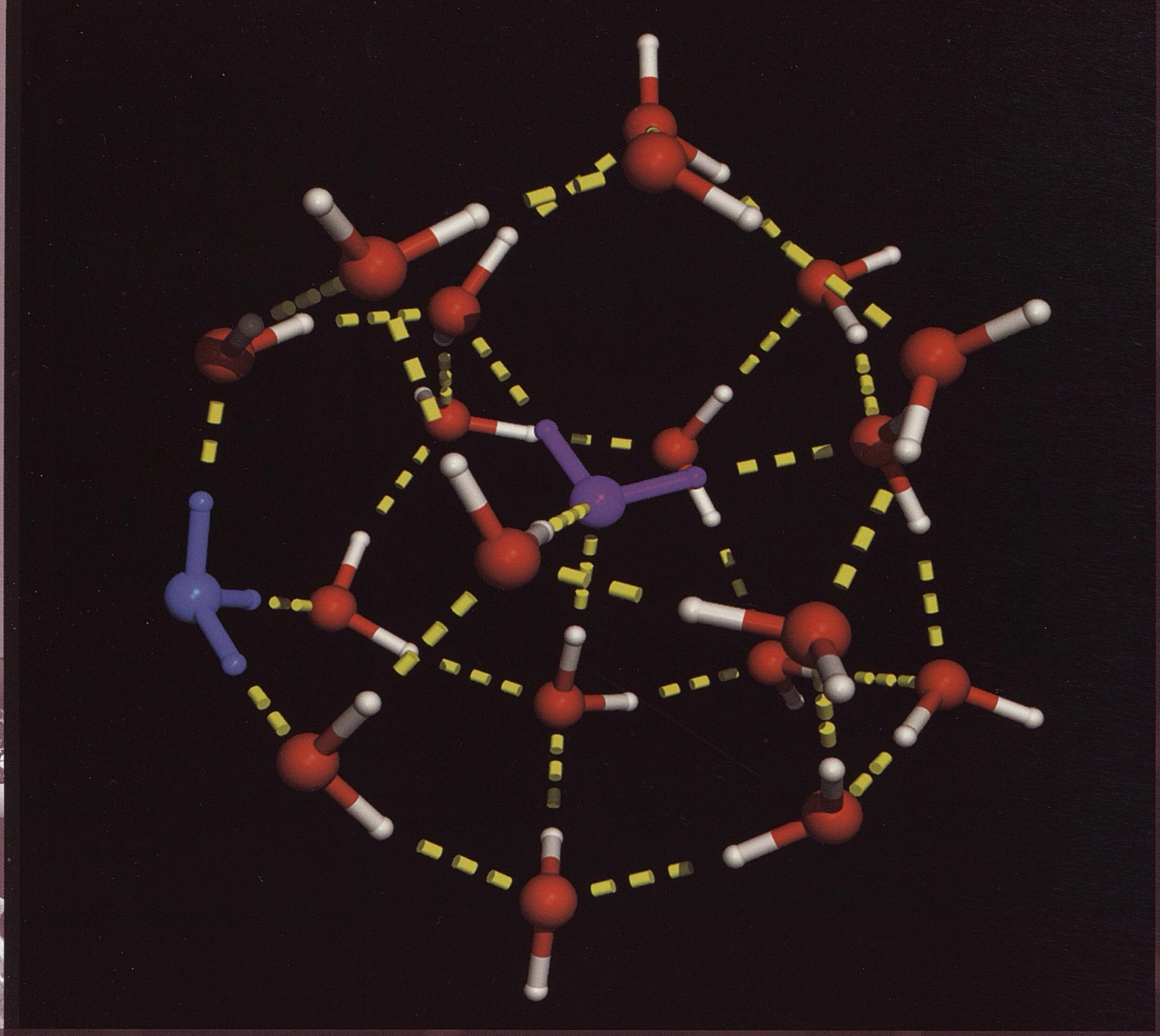


A wide variety of proteins in humans and other mammals use metal ions within the body, such as zinc or manganese, to catalyze biochemical reactions. These metalloenzymes perform many different roles, including DNA repair, hormone regulation and tissue repair. Some of them are involved in carcinogenesis, and some cancer drugs now in use inhibit the action of these enzymes as a way to retard or reverse the growth of cancerous cells. More detailed understanding of how these metalloenzymes function can translate into the design of better chemotherapeutic drugs.

The research group led by Michael Klein at the University of Pennsylvania carries out wide-ranging studies in materials science and biochemistry, and metalloenzymes have been the focus of recent PSC computations with LeMieux. The active sites of these enzymes—where the catalytic reaction takes place—contain one or more metal ions, held in place by nearby amino acids. During a reaction, many changes take place in the electronic structure within and near the active site, which can be explored only with quantum-mechanical simulations. These demanding calculations have become possible only recently, with the advent of terascale-class systems such as LeMieux.

This graphic depicts simulation results from an enzyme called Aeromonas proteolytica aminopeptidase (AAP). Involved in the final stages of protein manufacture, AAP is representative of other similar metalloenzymes. The ball-and-stick model represents its active-site structure (H-white, O-red, N-blue, C-green, Zn-lavender). The catalytic action at the bimetal core (two Zincs) depends on the exact positioning of amino acids closest to the metals (first shell amino acids) and their hydrogen-bonding pattern with nearby (second shell) amino acids.

The cross-sectional plane maps electron density as color contours (purple increasing through blue, green, yellow, red). This plane showcases the strong hydrogen bond between a first and second shell amino acid (upper right center, contours that touch) thought to be important in the fine-tuning of events at the core. Electronic data from these calculations is of great value in experimental efforts to design new therapeutic drugs.



WATER'S MAGIC NUMBER

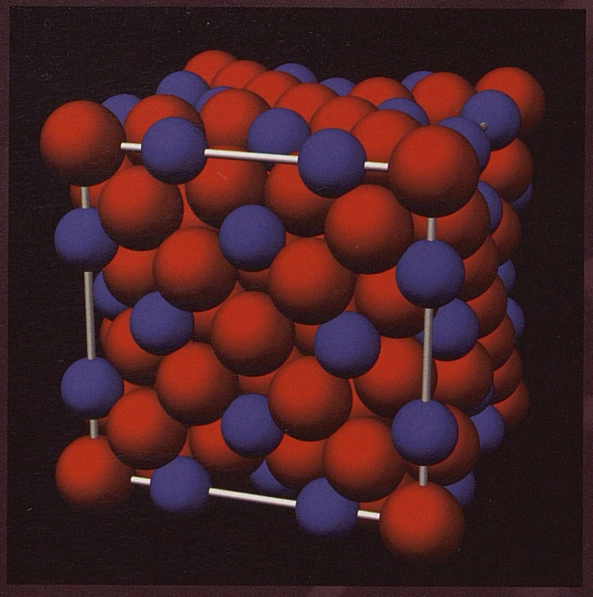
SOLVING ONE OF H_2O'S ENDURING MYSTERIES

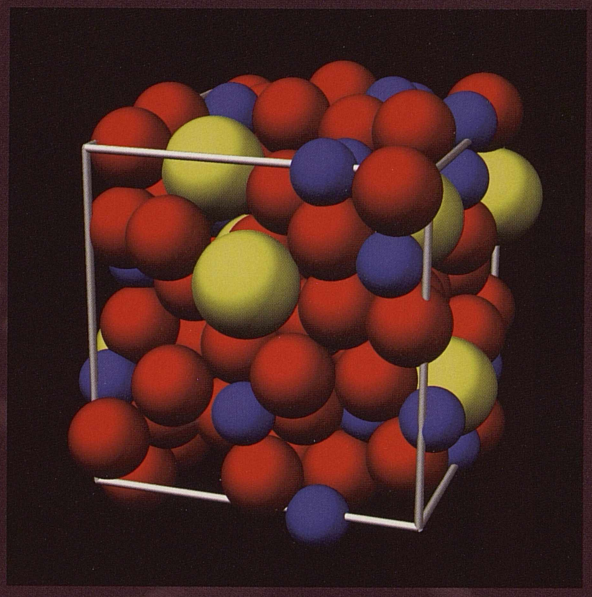
Water. Where would we be without it? You might think science would know all there is to know about this ubiquitous substance, without which our planet couldn't sustain life. But it's not so simple. School children know the basic formula, but H_2O holds many mysteries, not least among them how multiple H_2Os join together to form molecular clusters.

One of the more enduring mysteries has been the structure of water's "magic number cluster." Mass spectrometry shows that a cluster of 21 water molecules—with one extra proton (H+)—is much more stable than clusters with either 20 or 22 water molecules. "There's something imparting special stability," says Ken Jordan, professor and chair of the University of Pittsburgh Department of Chemistry, "and that's often associated with a special geometrical arrangement."

Studies over the past 30 years have postulated a dodecahedron, a cage of 20 water molecules, with an H_2O in the middle. But where's the extra proton? Does it go with the central H_2O or on the cluster surface? This question has persisted. Using LeMieux, Jordan did calculations to complement laboratory teams at Yale and the University of Georgia. Their collaborative findings appear to settle the question.

As depicted here, water's magic number cluster is a dodecahedron with the proton on the surface. The molecules are bound together by hydrogen bonds (dotted lines). One H_2O (purple) is in the center of the cage, and the excess proton is associated with an H_3O+ ion on the surface (blue). This finding, reported in *Science* (May 21), has stirred much debate. "There's still a question about how fast the proton can move around the surface," says Jordan. "It's not a finished story."





RECIPES FOR METALLIC GLASS

RUSTPROOF STEEL THREE TIMES AS STRONG AS CONVENTIONAL STEEL

A new kind of steel that's actually a glass is three times as strong as normal steel, resists rust and is nonmagnetic. It may someday be used in cars and buildings, and its magnetic properties offer the prospect of submarines and ships with hulls safe from mines that rely on magnetic detonation.

Research teams at the University of Virginia and Oak Ridge National Laboratory this summer announced such a steel—called a metallic glass or amorphous metal. While their work relied on costly, time-consuming trial-and-error methods of melting and casting alloys, Carnegie Mellon physicist Michael Widom has collaborated with PSC physicist Yang Wang to create a series of recipes for mixing and cooking elements to create many other metallic glasses.

As depicted in the two contrasting graphics, in a normal metal, atoms of two different elements (red and blue) arrange themselves in a regular pattern, a crystalline lattice structure. Widom's computations show, however, that adding a small amount of a large atom, yttrium (yellow), disrupts the crystalline lattice, allowing the irregular metallic glass structure to form.

With support from the Defense Advanced Research Projects Agency, Widom and Wang have used computational methods to investigate the composition and properties of various metallic glasses. Widom's computations confirmed the structure of the new amorphous steel and produced recipes for 2,000 other iron-based alloys.

PIPELINES TO THE STARS

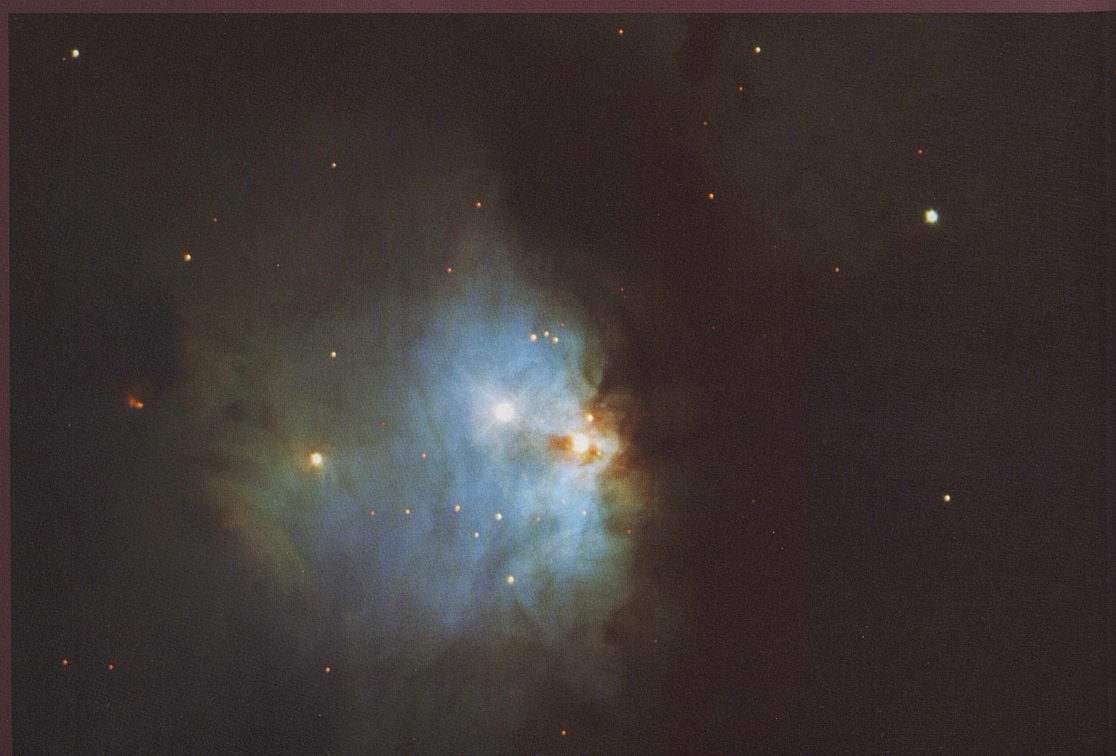
HOW TO DRAW INSIGHT FROM THE SKY AROUND US

Astronomy faces a data avalanche. Breakthroughs in telescope, detector, and computer technology have allowed astronomical surveys to produce terabytes upon terabytes of data about the sky around us. This vast data spans the range of the electromagnetic spectrum—from x-rays through ultraviolet and visible wavelengths to the cosmic microwave background. The next decade will bring several new surveys, with even more data. How can astronomers and astrophysicists keep up with analyzing and drawing insight from this massive trove of information?

The National Virtual Observatory (NVO) is a major new initiative in astrophysics, sponsored through an NSF Information Technology Research project, together with related international efforts. The objective is to interlace this data across the electromagnetic spectrum and to provide tools to explore and extract from it usefully. Only by collating data from multiple sources can science realize the full potential of this information.

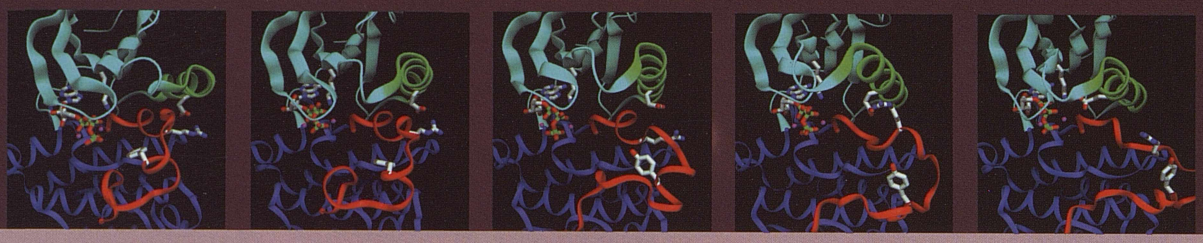
PSC astrophysicist Jeff Gardner collaborates with Andrew Connolly, University of Pittsburgh, Roy Williams, Caltech, and others on a major initiative to combine NVO with the NSF TeraGrid. Still in preliminary stages, this project aims to create an NVO Testbed, through which the massive data can be made available to supercomputing. Ultimately, the researchers expect to test representative astrophysics applications with NVO data and to encourage new ways to enlist supercomputing to extract knowledge from this invaluable data.

M78 Nebula in the Orion constellation, from the Sloan Digital Sky Survey. Hot young stars in the nebula's center illuminate the surrounding gas. Further out, dark clouds of dust prevent much of the scattered light from reaching us, creating a complex pattern of light and shadow. (SDSS Collaboration)



SIGNALS FOR CELL GROWTH

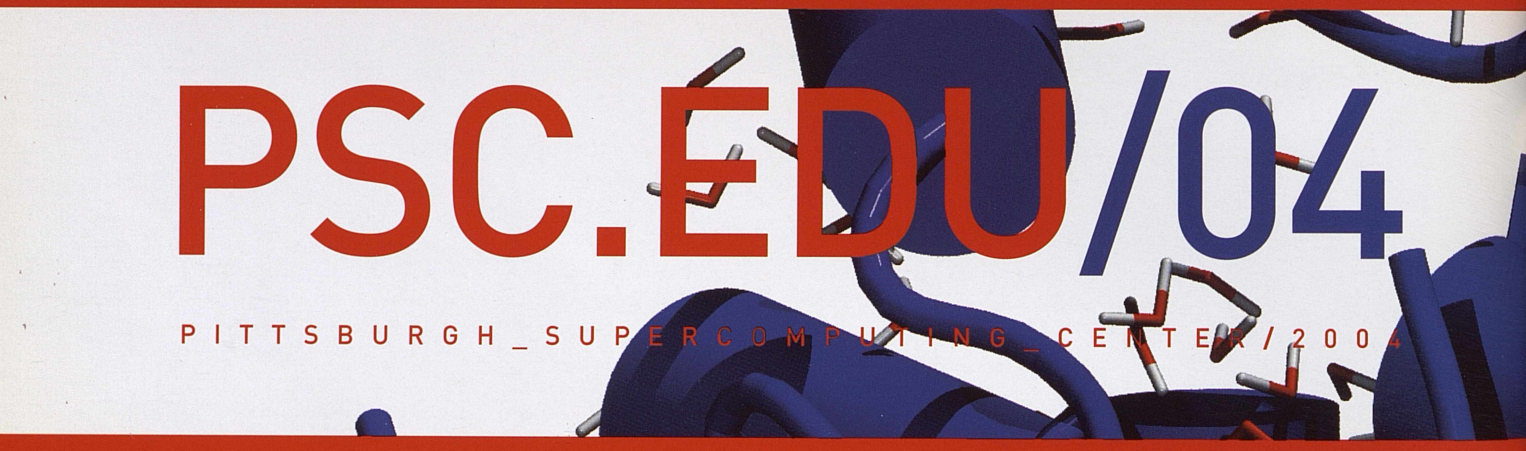
THINGS GO AWRY WHEN THIS PROTEIN IS OVERACTIVE



A family of proteins known as Src tyrosine kinases plays a key role in receiving and passing on a flurry of biochemical messages that regulate the growth of cells. They are active or inactive depending on a biochemical switching mechanism that changes the protein structure. Things can go badly awry, however, when this protein becomes overactive. Mutated versions of Src that send an unwavering signal for cell growth are often found in human cancers.

Scientists are searching for drugs to stop these rogue kinase proteins, and with that underlying objective, biophysicists Benoit Roux and Nilesh Banavali of Weill Medical College of Cornell University are using LeMieux to study Src. While the structure is known, this static picture doesn't show how flexible regions of the protein move and change as the protein switches from the inactive to active state.

Roux's research team uses molecular dynamics, a computational approach that tracks the protein's movement atom-by-atom over time, and he is looking in particular at the changes associated with opening the activation loop of Src. In this sequence of frames from a simulation of the Src kinase domain over a period of 14 nanoseconds, the activation loop (red)—with an associated tyrosine amino acid—moves from a closed (inactive) to open (active) state.



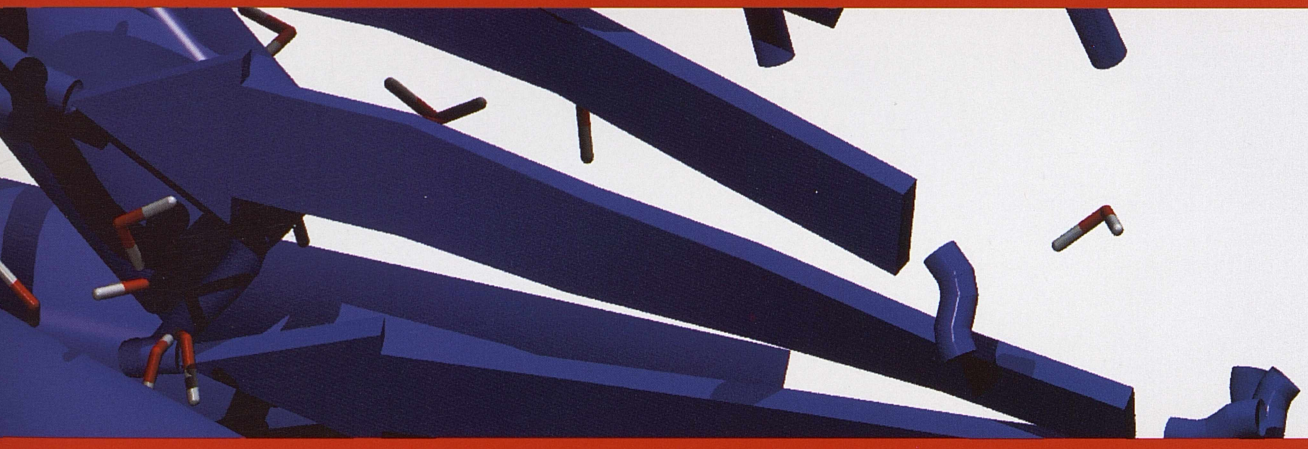
PSC.EDU/04

PITTSBURGH_SUPERCOMPUTING_CENTER/2004

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 U. S. Department of Defense
The U. S. Department of Energy
Cisco Systems, Inc.
Cray Inc.
Intel Corp.
Silicon Graphics, Inc.
Hewlett-Packard Company
The Buhl Foundation
The Grable Foundation



PSC_2004

EDITOR/WRITER: Michael Schneider, PSC
CONTRIBUTING WRITERS: Katie Greene, Celeste Kimbrough, Tim Palucka
GRAPHICS RESEARCH, PHOTOGRAPHY DIRECTION, COPY EDITING: Sean Fulton, PSC
TRANSCRIBING: Jeff Davison & Robert Gardner, PSC
PHOTOGRAPHY: Photography & Graphic Services at Mellon Institute
GRAPHICS: Thanks to Greg Foss, PSC, for tornado and black liquor graphics, and for his help in general. Thanks to David Deerfield, PSC, for photos. Thanks also to Carlos Simmerling, Cray Inc., Pittsburgh Public Schools and to Sean Fulton for re-renderings.
DESIGN: Wall-to-Wall Studios, Inc.
PRINTING: Schiff Printing

♻ PRINTED ON SAPPI MCCOY GLOSS PAPER, A PREMIUM SHEET WITH 10% POST-CONSUMER WASTE FIBER, WITH VEGETABLE-BASED INKS.

PITTSBURGH SUPERCOMPUTING CENTER
MELLON INSTITUTE BUILDING
4400 FIFTH AVENUE
PITTSBURGH, PENNSYLVANIA 15213