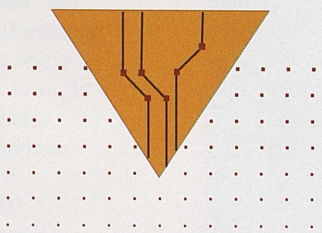
1993

PROJECTS IN SCIENTIFIC COMPUTING

PITTSBURGH SUPERCOMPUTING CENTER

Cover: *A representation of the enzyme Eco RI endonuclease wrapping around DNA. It was created by University of Pittsburgh biologist John Rosenberg from his calculations at the Pittsburgh Supercomputing Center (see p. 10). Rosenberg used RIBBONS (a program developed by Michael Carson, University of North Carolina, Chapel Hill) to generate the image on a Silicon Graphics workstation.*

Projects in Scientific Computing 1993



PITTSBURGH
SUPERCOMPUTING
CENTER

Published July 1993 by the Pittsburgh Supercomputing Center,
4400 Fifth Avenue, Pittsburgh, PA 15213. (412) 268-4960.

Foreword

It has been a year of great achievement for the Pittsburgh Supercomputing Center (PSC) community. With the installation of the first Cray C90 outside government, we have begun a complete revamping of our basic systems. This substantial upgrade of the computational capability available to the national research community is enabling researchers to expand the depth and scope of the problems they can tackle. Indeed, the C90 was requested by a majority of the National Science Foundation (NSF) proposals for grand challenge computational resources.

In its first few months of service, the C90 already has enabled ground-breaking research, some of which is described in this booklet. Molecular dynamics computations are addressing one of the premier grand challenges in molecular biology, the protein-folding problem. Significant new results have been achieved in modeling the growth of semiconductors — research that will eventually make possible even faster, more powerful supercomputers. For the first time, a model for predicting individual thunderstorms has been used in a fully automated mode for operational weather forecasting. And the C90 has brought researchers within striking distance of a realistic, three-dimensional model of blood flow in the heart. This is truly exciting work.

We eagerly await Cray's massively parallel product, the T3D. The very first T3D is due to be shipped to PSC in the summer of 1993. Recognizing that software is the major issue confronting effective use of massively parallel machines, we have launched a vigorous initiative, together with Cray Research, to make commonly used software packages, primarily in chemistry and biology, quickly available on the coupled C90-T3D system. Supported by a new file-serving system, based upon the Andrew File System and Cray's EL mini-supercomputer, we expect that this tightly coupled MPP and vector system quickly will prove itself as a powerful research tool, heralding a new paradigm in scientific computing.

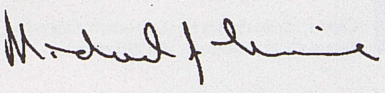
Under the banner of the MetaCenter, we are striving together with our sister NSF Centers to provide the research community with a wide range of computing platforms, to work towards unified user interfaces, and to undertake projects beyond the capability of any single center. PSC is leading joint efforts in developing a national file system and working in the MetaCenter to explore many more issues including mass storage technologies, information services, increased networking capability, applications software for parallel machines, and visualization tools.

In the past year, federal attention in high-performance computing has been focused on developing a national information infrastructure, extending the reach of high-performance computing to new constituencies, such as schools and small industry. In pre-college education, PSC is taking the lead through Common Knowledge: Pittsburgh, a project that will explore the educational benefits of networking the entire Pittsburgh Public Schools system.

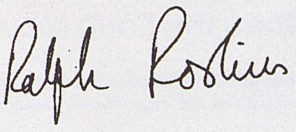
It was particularly gratifying that PSC this year received the prestigious Computerworld Smithsonian award in the area of science, recognizing research that improves the quality of life. This award is tribute to the excellence of the scientific community working at the PSC and to the dedication and expertise of our entire talented staff.

Ralph Roskies (left) and Michael Levine, scientific directors of the Pittsburgh Supercomputing Center, with the CRAY Y-MP C90.

Michael J. Levine, Scientific Director



Ralph Z. Roskies, Scientific Director



Contents Projects in Scientific Computing, 1993

Pittsburgh Supercomputing Center, 1993

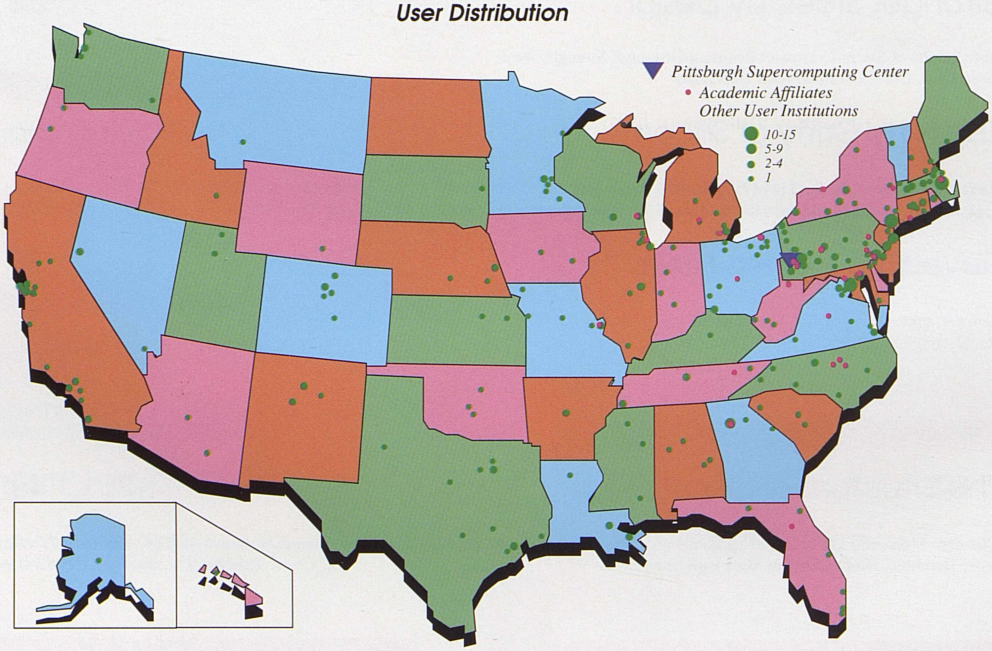
PITTSBURGH
SUPERCOMPUTING
CENTER

Highlights

- The newest, most powerful system from Cray Research the Y-MP C90, is running in production mode, allowing researchers to address problems beyond the capability of prior systems.
- Work is underway preparing for the arrival of Cray's new, massively parallel system, the T3D; the first model to leave the plant is scheduled for August delivery in Pittsburgh.
- SuperCluster, the center's system of 10 linked DEC workstations, is being upgraded to dramatically increase computing power.
- The MetaCenter, a collaboration among the four National Science Foundation supercomputing centers to better support scientific research nationwide, was announced in November 1992, and a number of joint efforts are underway.
- The 1993 Computerworld Smithsonian Award for science recognized the center's efforts to bring high-performance computing to bear on research that improves the quality of human life.

Pittsburgh Supercomputing Center

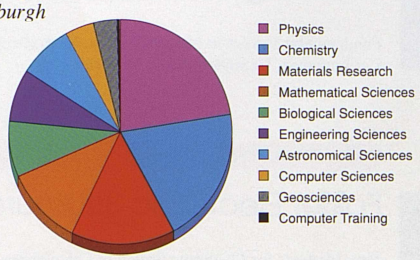
User Distribution



Established in 1986 with a grant from the National Science Foundation (NSF) supplemented by the Commonwealth of Pennsylvania, the Pittsburgh Supercomputing Center is a joint project of Carnegie Mellon University and the University of Pittsburgh together with Westinghouse Electric Corp.

To date, more than 5,200 scientists and engineers at more than 400 universities and research centers (red and green dots) in 49 states and the District of Columbia have used the center's computing resources to advance their research. This work has resulted in more than 1,500 papers in professional science and engineering journals.

Researchers connect to the Center via regional electronic networks that feed into NSFnet, a high-speed pathway that links NSF supercomputing centers. Twenty-eight universities (red dots) are Pittsburgh Supercomputing Center Academic Affiliates (see back page). Representatives from these campuses form the Center's main advisory body.



CRAY C90 Usage by Discipline

Arrival of the C90

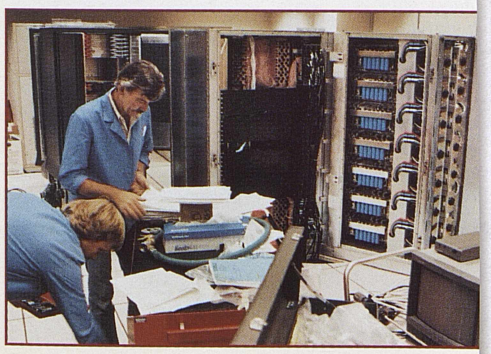
1993 finds the Pittsburgh Supercomputin
positioned to address the grand challenges o
science and engineering. A mighty, new ma
Y-MP C90, arrived in October 1992 and qu
the center's ability to serve its users. Anothe
Cray's much anticipated massively parallel
soon to arrive.

Marvin Zalevsky, managing director, and
executive director. "Our work with Cray has
where we are one of the premiere installatio
general scientific computing," says Clayton,
expand when we become the first T3D instal

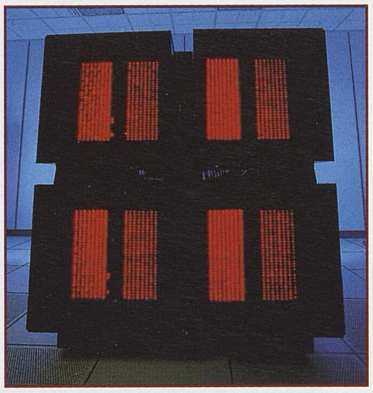
The C90 is the newest, most powerful s
Research, and Pittsburgh was the first non-
the United States to receive it. "With this n
doubled the full scientific capability of the
ing centers' program," says co-scientific d
Levine. Scientists moved quickly to take a
resource, and by March over 30 proposals
research — large-scale problems with sign
impact, such as global change modeling a
proteins and DNA — sought time on the C

Even as it was being tested, the C90 pro
scientific results. From early December to
handful of experienced computational scie
machine as "friendly users" — to help sha
running real research code. This opportuni
significant progress for several projects, in
ductor modeling (p. 40), complex liquids (
structure (p. 14) and heart modeling (p. 22

The C90 was delivered Oct. 23, 1992, and
a team of Cray Research engineers worked
round-the-clock to bring the system online.



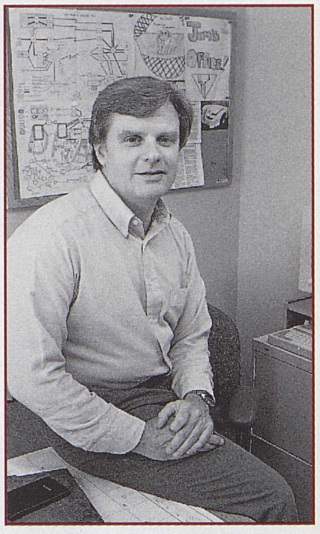
aff.



stem: T3D

new machine under control than preparation
atest technology system. Cray's first
rocessing (MPP) system, code-named the
or 1993 availability. The first available
Pittsburgh in August. The center's staff will
ray to test the new system and already has
rams to take advantage of Cray's much
he MPP field.
eeds now, some people say, is an MPP
ed with a powerful conventional supercom-
l lead the way. "Most computationally
roblems ," says co-scientific director Ralph
of which are suited to MPP, while other
verful scalar or vector machines like the
PP, many processors sit and wait, wasting
arallel parts of the job are run. With the
ndle this processing. We expect it to be a
."

with the T3D builds directly on its
coupling the CRAY Y-MP with the
CM-2. In 1991, as part of Carnegie
igabit
he center
k Systems
ork Technolo-
systems with
nsfer link.
for the first
nputing
the MPP and
f a heteroge-
ng system.

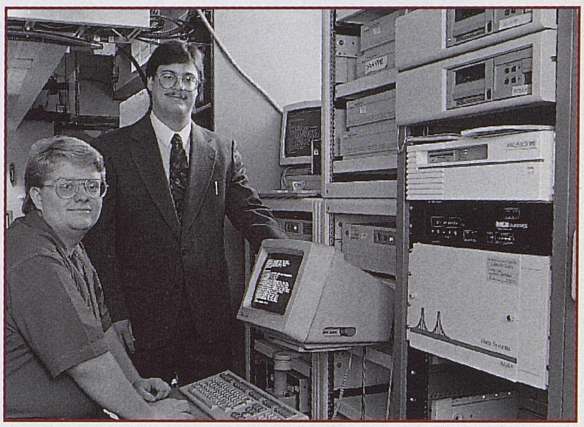


nacci, manager,
tions and
wo initiatives
tivity in
s group this
izing network
e of the center's new data migration facility and
proving network connections among the four
omputing centers. "The key in both cases," says
"is optimizing new technology for a high-
e computing environment."

SuperCluster Phase Two

June 1, 1993: SuperCluster celebrated its first birthday. The center's linked system of 10 DEC DS 5000 workstations has been in production for a year and has proved itself a valuable alternative to the C90 for many jobs. "This high-performance cluster has meant added versatility and expanded access to advanced computing for our users," says Levine. "A number of significant projects — including protein-structure refinement and large-scale air pollution modeling — have made productive use of this resource."

That's only the beginning, says research systems programmer Rob Pennington, who leads the SuperCluster project. SuperCluster Phase Two, faster and better than before, is on the way. Ten of DEC's newest workstations, driven by the powerful Alpha AXP chip, will replace the DS 5000s, and they will be linked with a new switched interconnect that will increase communication among the processors in the cluster from 10 to 100 times, promoting greater use of parallel processing. "Phase Two will be in place by the end of summer," says Pennington, "and it will provide a dramatic increase in computing power."



Jon Boone (seated) and Steve Cunningham of the center's communications group in the "machine room," where the center's networking equipment is maintained in a controlled environment.

The MetaCenter Unveiled

In November 1992, the four NSF supercomputing centers announced they were collaborating to form a unified national resource for computational science and engineering. The aim ultimately is to make it possible to distribute scientific computing among systems at any of the four centers. "The MetaCenter," says Levine, "will work toward giving scientists and engineers the ability to move portions of their problems directly to appropriate computer architectures without regard for where the computers are located."

MetaCenter, n.: a coalescence of intellectual and physical resources unlimited by geographical constraint; a synthesis of individual centers that by combining resources creates a new resource greater than the sum of its parts.

Staff members working on the Andrew File System include (standing) Carolyn Councill, Bill Zumach, Doug Balog, (seated) Vidya Dinamani and Kathy Benninger. Not pictured are Jon Goldick , Chris Kirby, William Brown and Chris Maher, manager, scientific support.

Within the MetaCenter, individual centers have coordinated acquisitions, so that researchers can have access to large MPP machines from the various manufacturers entering this volatile new market. Thus, the major MPP thrust at Pittsburgh will be the T3D, extending work begun with the Connection Machines, CM-2 and CM-5. The Connection Machine CM-5 available here in late 1992 is no longer available. The CM-5 is now a central focus of the National Center for Supercomputing Applications, which is continuing work with Pittsburgh users whose projects are best targeted for this system. This kind of cooperation to make diverse computing resources available to the U.S. research community is a major benefit of the MetaCenter initiative.

Pittsburgh hosted the second NSF MetaCenter workshop on March 22-23, 1993, at which nearly 60 people from the four centers and NSF gathered in working groups. To help continue this cooperative work, videoteleconferencing equipment linked all four centers in early 1993. Working group meetings and frequent center directors' meetings make productive use of this equipment.

One of the most important steps toward achieving MetaCenter computing is a national file system, so that researchers can easily move data between machines thousands of miles apart without regard to where the data is stored. Pittsburgh has taken the lead in this project, which will be based on the Andrew File System (AFS), a system designed to make all files appear as part of the user's local file system regardless of where they physically reside. AFS was developed at Carnegie Mellon, with support from IBM (which also supports the center's AFS effort). A team of Pittsburgh Supercomputing Center programmers and scientists adapted AFS for use in a supercomputing environment, and it is now also in use at the other NSF centers.

Biomedical Supercomputing

In recognition of the center's efforts to bring high-performance computing to bear on research that improves the quality of life, the Pittsburgh Supercomputing Center won the 1993 Computerworld Smithsonian Award for science. This prestigious international award is given for uses of information technology that benefit society.

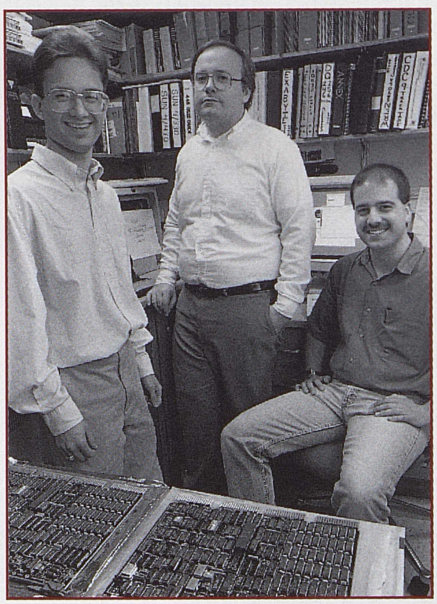
The award cited DNA research led by University of Pittsburgh biologist John Rosenberg. The collaboration that got this work off the ground originated at one of the center's early biomedical workshops. With funding from the National Center for Research Resources' Biomedical Research Technology Program of the National Institutes of Health, the center's biomedical initiative encourages the use of supercomputing in biomedical research through the development of software and through educational workshops.

Jim Kasdorf (left), director of supercomputing at Westinghouse Electric Corp., joined Ralph Roskies and Michael Levine at a June 7 ceremony in Washington, D.C. where it was announced that the Pittsburgh Supercomputing Center won the 1993 Computerworld Smithsonian Award for science. "This is a tribute to our staff," says Roskies. "We've done an outstanding job at enabling other people to get their work done. You don't get a lot of attention for this. A lot of details go into making everything mesh so that researchers experience things as going smoothly. Our staff has done a glorious job."

In August 1992, for the second year in a row, the biomedical staff mounted a week-long workshop, funded by the Human Genome project, on sequence analysis of nucleic acid and proteins. Twenty-two researchers learned to use high-speed techniques for matching gene sequences against existing databases, an important part of the ongoing work to map the human genome.

In November 1992, the center's first ever workshop on neural computing, co-sponsored with Yale University, drew 50 highly regarded researchers from around the United States. The workshop focused on both artificial intelligence-based neural network simulation and biological models of nervous system function. It identified current and future needs for high-performance computing in all areas of computational neuroscience.

In early 1993, the W.M. Keck Foundation awarded funding to the center to establish a graduate and postdoctoral training program in computational biology in conjunction with the University of Pittsburgh and Carnegie Mellon University. An interdisciplinary group of more than three dozen faculty at both campuses — in biology, chemistry, physics, mathematics and computer science — is participating. The center will help set direction for the program and provide high-performance computing facilities and training. Hugh Nicholas, biomedical scientific specialist, will teach a course in sequence analysis at the University of Pittsburgh as part of the program.



Ken Goodwin (right), coordinator, Ed Berger (center) and Jeff Semke of the center's hardware group, which also includes (not pictured) Steve Petko.

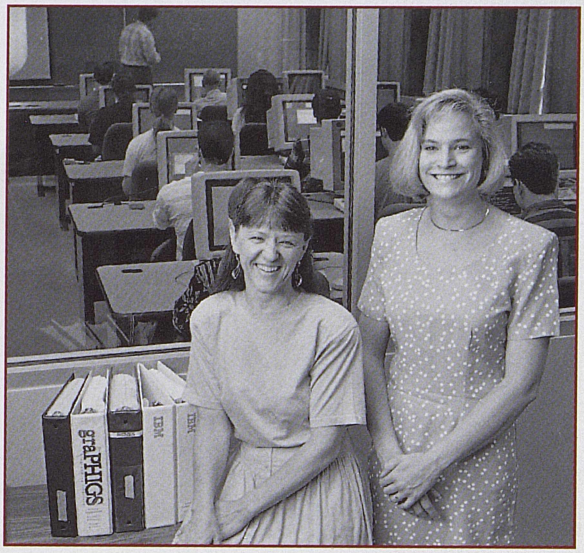
Susan Lambl, allocations coordinator. Lambl oversees grants of time on the center's computing systems, handling account administration for over 2,000 active users. This year her group began using a new database to help manage this information.

Education and Training of Users

A central part of the center's mission is to help train scientists and engineers in how they can use high-performance computing to advance their research. During the past year, the center sponsored 21 workshops with over 710 participants. A series of seminars held bi-weekly through the academic year drew over 1,000 attendees. Workshops include introductory level courses and advanced training that apprises researchers about state-of-the-art applications in their field.

Pittsburgh Supercomputing Center Workshops
(from July 1992 to June 1993)

Summer Institute on Supercomputing
Advanced Image Processing in Electron Microscopy
Computational Neuroscience Research
Nucleic Acid and Protein Sequence Analysis
Supercomputing Techniques: Connection Machine
Supercomputing Techniques: Cray
Introduction to GAUSSIAN
Cray Advanced Users
Supercomputing Techniques: Parallel Processing



Alice Conniff (left), education and training coordinator, and Jamie Christner, education and training assistant, outside the Computer Training Center, where participants in the center's workshops receive hands-on training. "Through our introductory workshops," says Conniff, "we take people with little exposure to supercomputing and turn them into confident users."

Corporate Affiliations

"Computational science and engineering is transforming the research environments of major American corporations," says John Butler, who along with Martin Straut works at making high-performance computing available to industry. ALCOA, USX, Chevron, Westinghouse and DuPont, among other corporations, have taken advantage of the Pittsburgh Supercomputing Center's resources to boost their productivity.

The center offers much more than hardware and software, emphasizes Straut. "Computational science is not just hardware. Access to our workshops and consulting staff gives companies the know-how and support they need to get started with new technology. Interaction between our staff and corporate product teams broadens our awareness of the problems corporations face and enhances their ability to compete in a global market." (MS)

John Butler (left) and Martin Straut, corporate liaisons.

Information at the Pittsburgh Supercomputing Center

Proposals for Computing Time
Susan Lambl

Biomedical Initiative
Nancy Blankenstein

Workshops & Summer Institute
Alice Conniff

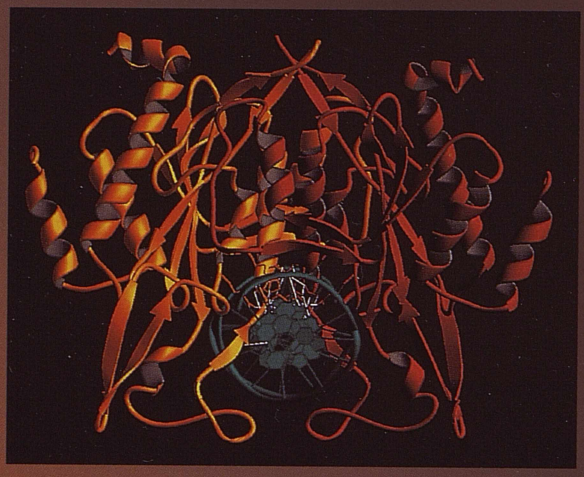
High School Initiative
Casey Porto

Corporate Affiliates Program
John Butler & Martin Straut

Newsletter and Documentation
Vivian Benton

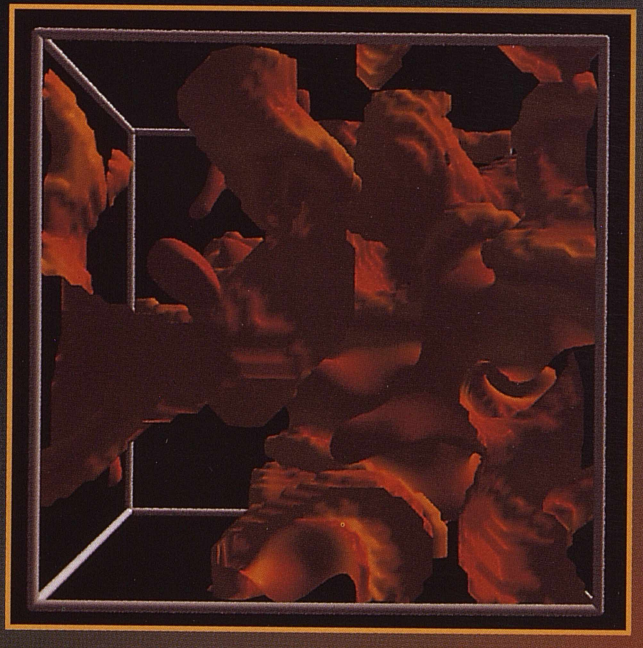
These Pittsburgh Supercomputing Center staff members can be contacted by telephone (412-268-4960) or electronic mail.
Internet: *lastname*@psc.edu.
Bitnet: *lastname*@cpwpsca.

Scientific Visualization



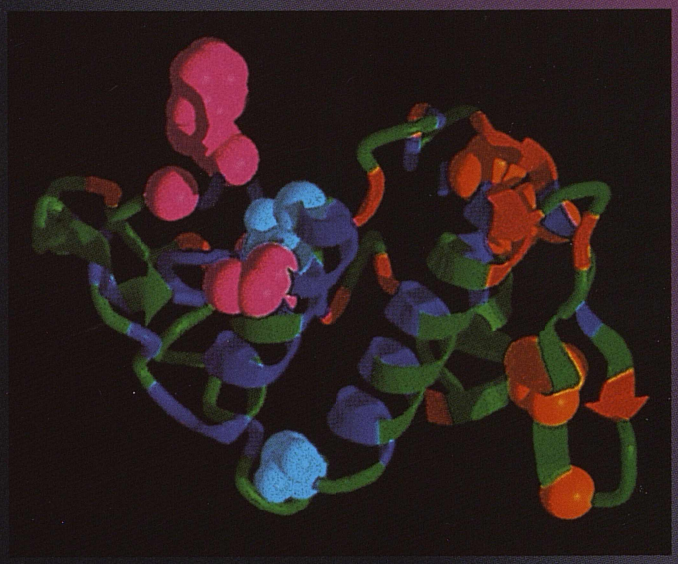
Kinky DNA

The 1993 Computerworld Smithsonian Award for science, won by the Pittsburgh Supercomputing Center, recognized research by biologist John Rosenberg of the University of Pittsburgh on the protein-DNA interaction depicted in these graphics. This ribbon model from Rosenberg's calculations shows top, front and side views of the restriction enzyme Eco RI endonuclease wrapping around DNA (blue). The enzyme is composed of two symmetrical subunits (red and yellow) that correspond to the twofold symmetry of DNA's double helix. Each subunit kinks one of the DNA backbones (which can be seen in the side view) and like scissors snips the DNA at the kink. Because of its ability to cut DNA at this precise location, Eco RI has become one of the biotechnology industry's most important tools for cloning DNA.



Vortex Tube Dynamics

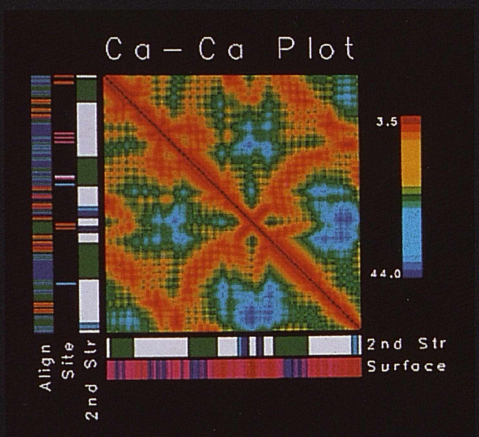
This three-dimensional graphic, showing vortex filaments threading through a cube, comes from simulations of the complex Ginzburg-Landau (CGL) equation, which can describe the behavior of fields in various hydrodynamic, optical and chemical systems. The colored surfaces represent low amplitudes from the CGL waveform, with associated phases colored from red to yellow. Joel Welling of the Pittsburgh Supercomputing Center graphics staff produced the graphic from calculations by Greg Huber of Boston University on the Connection Machine, CM-2.



Sequence Analysis of Rattlesnake Protein

Graphics tools developed by the Pittsburgh Supercomputing Center biomedical group provide an integrated view of information obtained from MaxSegs, a program for sequence analysis of proteins and nucleic acids. The ribbon diagram represents three-dimensional structure of the protein phospholipase A-2 from the western diamondback rattlesnake. Color coding shows how the amino-acid sequence varies in comparison to the same protein in other poisonous snakes, a few mammals and the honeybee. Red indicates the most variable amino acids and blue least variable. Amino acids that do not vary are the most important in maintaining three-dimensional structure and biochemical function of the protein.

The color-coded square, a C (carbon) alpha to C alpha plot, gives a more abstract representation of the same information. The square is a matrix of the amino-acid sequence colored according to distance (in angstroms) between amino acids in the three-dimensional structure. Red indicates the closest and blue the farthest apart. The linear sidebars represent four kinds of supplementary information. "2nd Str" (secondary structure) shows regular, repeating helices (green), beta-sheets (blue) and extended regions (dark blue). "Site" highlights amino acids important for specific functions of the molecule. "Align" shows the same amino-acid variability that is color-coded in the three-dimensional structure. "Surface" indicates whether an amino acid is on the surface of the three-dimensional structure (red) or buried in the protein (blue).



Embryonic Development

These images show results from a model of how differentiating cells interpret positional information. Michael Levin of Tufts University developed his own software for the model, which uses a modified Julia set algorithm and runs on the Connection Machine CM-2. Anjana Kar of the Pittsburgh Supercomputing Center graphics staff produced these 35 mm slides from Levin's data.

Reaching Out to Schools

The High School Initiative in Computational Science

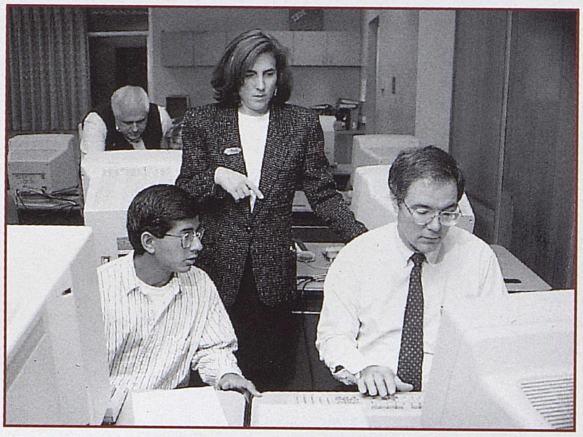
"It has changed the way I teach physics," says Ivan Ober, teacher at Mt. Lebanon High School, south of Pittsburgh. "The skills I've been taught, the data I've stored and the workstation we were given will benefit our school for years." Ober is talking about his experience in the Pittsburgh Supercomputing Center's High School Initiative in Computational Science.

Through this program, the center's staff and supercomputer become active partners with high school teachers in making science something to be excited about. The aim is to bridge the gap between textbook learning and hands-on science by letting students become involved with computational science projects of their own choosing.

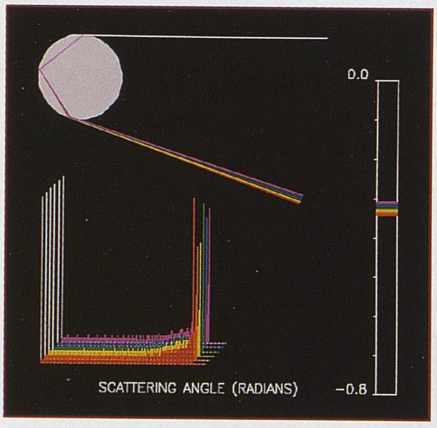
Now beginning its third year, the program each year selects proposals from 10 schools. Teacher-student teams then attend an intense five-day workshop at the center, where they build basic skills and refine their approach to the project outlined in their proposal. The center provides network access from each school to continue work during the school year, and a consultant from the center's staff works directly with the project teams, visiting each school periodically. Each team also works closely with a mentor drawn from university faculty and industrial researchers in the Pittsburgh area. Electronic mail facilitates regular communication between the school project teams and their mentors.

Through the first two years more than 65 teachers and 325 students have taken part in the program, enriching their skills and gaining confidence in their ability to accomplish complex tasks. Probably the most valuable product of this unique program is the enthusiasm it generates among students. Mt. Lebanon teacher Ober says his students often spent an hour or more after school working on their projects, and this experience is the norm for the program. "It was the students' excitement and zest for learning," says Ober, "that made the project a success."

The High School Initiative is funded by the National Science Foundation (NSF) Education and Human Resources Directorate. A generous grant from Digital Equipment Corp. for the first two years awarded a DS5000 workstation to each school.



Casey Porto (standing), education liaison, who directs the High School Initiative, with teacher Edmund Escudero and student Rahul Rathod of Summit Country Day School, Cincinnati, Ohio, during a workshop at the Pittsburgh Supercomputing Center.



How do Rainbows Form?

A team of 21 students and three teachers at The Ellis School in Pittsburgh used supercomputing to investigate the physics of light and color in their project: The Formation of Rainbows. Lead teacher Mark Walker designed his proposal around a project that could involve students from all grade levels and diverse interests.

The team created computer code that models white light reflecting from a spherical raindrop. The program calculates angles of refraction inside the raindrop for six wavelengths of light. Though Walker wrote most of the algorithms, students modified the code to accommodate different sets of data. Students also created a 204-element color table for the color display in their final product, a thousand-frame animation.

Teacher Mark Walker and student Parke Weigman of The Ellis School discuss their project with mentor Robert Schumacher, physics professor at Carnegie Mellon University.

"If we're going to strengthen our economy and create jobs, we must move these advanced technologies from the laboratories into the marketplace — into the factories where cars are manufactured, into the hospitals where surgery is performed and into the schools where children are being educated."

— Senator Albert Gore, chairman of the Senate Commerce Subcommittee on Science, Technology and Space. July 1992.

Common Knowledge: Pittsburgh

The Pittsburgh Supercomputing Center is providing the technical know-how for Common Knowledge: Pittsburgh, an innovative program to make computer networking an effective component of learning in each classroom of the Pittsburgh Public Schools. The National Science Foundation (NSF) awarded $1.989 million to support the project for the first two years. It will be a pilot program nationally for using computer networking in kindergarten through 12th grade education.

One objective of Common Knowledge is to provide access to the network from every classroom in the Pittsburgh Public Schools in the next five years. Teachers and students will be able to communicate electronically within their schools and with other schools in the district. Through access to NSFnet, they will also be able to use electronic resources nationwide, including the catalogs of all major libraries, research databases and on-line educational materials.

"Networking in the schools is a powerful resource to overcome isolation and to provide equal access to highest quality educational materials," said Robert Carlitz, University of Pittsburgh physics professor who is project director for Common Knowledge.

Another goal will be to integrate networking into the school curriculum. Curriculum supervisors and teachers in the school district have been involved in planning Common Knowledge, and projects already underway use networking not only in science and math, but in nearly all disciplines, including innovative projects in writing, foreign language study and environmental science.

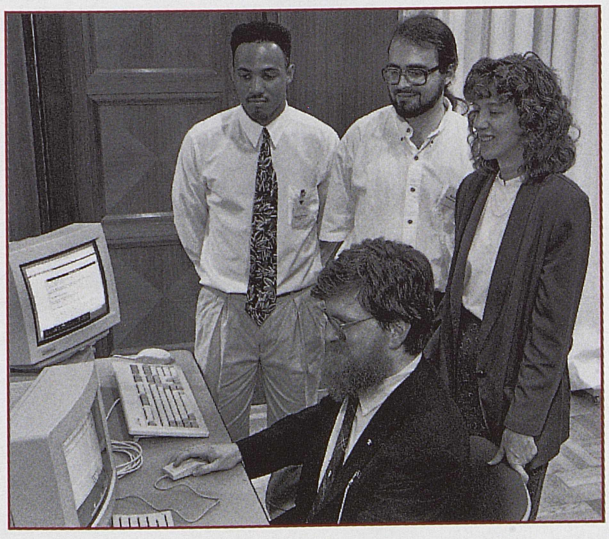
The foreign language project, for instance, will allow students to correspond electronically with their peers in other countries, and a program at Westinghouse High School makes it possible for science and math students to communicate with mentors on the research staff at Westinghouse Electric Corp. "This project has captured the imagination and enthusiasm of many of our administrators and teachers, who see it as a way to extend instruction beyond the four walls of the classroom," said Superintendent of Pittsburgh Schools Louise Brennen.

Computer equipment for Common Knowledge is being provided with the cooperation of Apple Computer, Inc., Digital Equipment Corp., Sun Microsystems, Inc. and Telebit Corporation. (MS)

"Through our high school outreach program, we have seen how computing technology can excite and motivate students and enrich their educational experience," said Ralph Roskies, the center's co-scientific director, speaking at the June 14 reception officially announcing Common Knowledge. "When the Pittsburgh Supercomputing Center was established we were sure it would affect the community in ways we couldn't foresee. This project is one example of how high-tech infrastructure contributes to improving the quality of life."

"Networking technology has proven its value at forging collaborations among scientists widely dispersed geographically," added co-scientific director Michael Levine. "The national investment in this infrastructure can be leveraged through projects like Common Knowledge to reach into the schools nationwide, where it will be a powerful educational tool — an exciting way to acquaint young minds with computer technology and through it, science."

Pittsburgh Supercomputing Center staff who provide the technical expertise for Common Knowledge: (l to r) Maurice Freeman, student intern, Gene Hastings (seated), project engineer, Kevin Sullivan, project administrator, and Wendy Huntoon, project manager.



How Proteins Get in Shape

Molecular Dynamics Simulations of Protein-Folding Intermediates
Charles L. Brooks III, Carnegie Mellon University

The Protein-Folding Problem

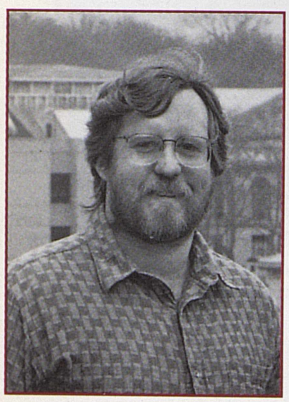
A droopy, strung-out chain of amino acids that isn't much good for anything — that's what rolls off the assembly line of the molecular factory inside a cell when a protein is created. All the pieces are there, and they're in the right sequence. Yet the new protein is unfit for duty. It isn't in shape.

To do its job, whatever it may be among the thousands of life-sustaining jobs proteins do, this dangly chain forged from hundreds of amino acids must fold into just the right three-dimensional configuration. It happens within seconds, a long time in protein biochemistry, and the result is a complex bundle of twists and turns with clefts and notches precisely sculpted to allow the protein to attach and release other molecules. Function follows from form, and when the form is right, the protein goes to work.

For biomedical scientists, this phenomenon poses the kind of mystery science thrives on. How is it that a particular sequence of amino acids uniquely determines the right shape, out of almost unlimited possibilities, so that the protein can perform its predestined biological role? It's called the protein-folding problem, and solving it is no mere intellectual exercise. In theory, knowing the biological laws that govern protein folding could make it possible to create new proteins made to order to cure diseases and abate maladies from indigestion to arthritis.

The Supercomputer as Laboratory

"The relationship between protein sequence and three-dimensional structure is one of the primary unsolved problems in biology today," says Charles Brooks, who is attacking this grand challenge problem in several related projects at the Pittsburgh Supercomputing Center. Brooks is a leader in the development of a computational method called molecular dynamics (MD). He helped develop the CHARMM package of MD software, used by a number of researchers working in protein and DNA structure analysis, and he has been applying MD in protein research at the Pittsburgh Supercomputing Center since it opened its doors in 1986.

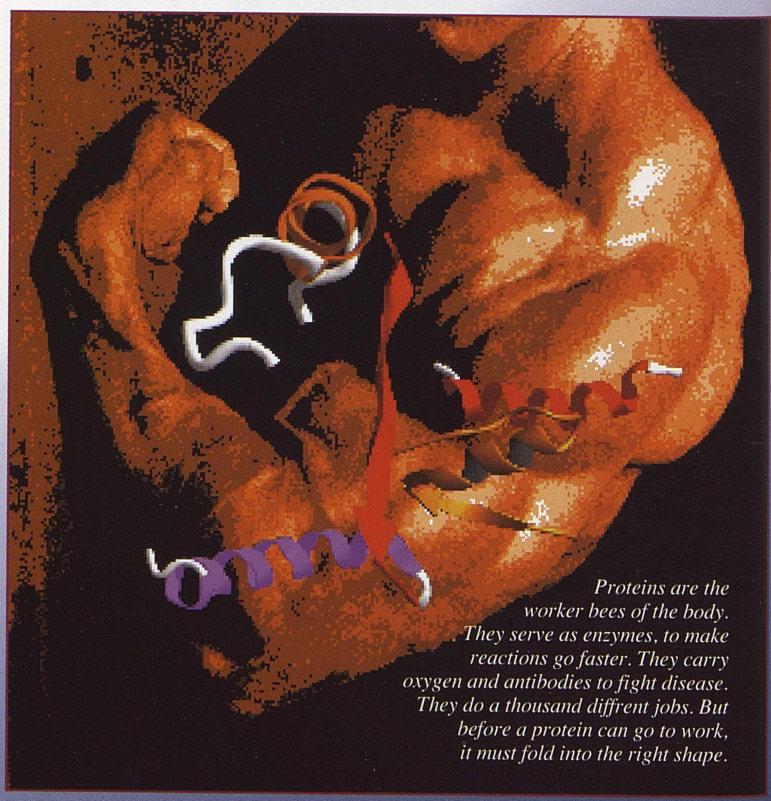


Charles L. Brooks III

As the name implies, MD is a way of tracking how molecular structure changes over time. In their natural environment inside living cells, proteins are flexible and constantly in motion, almost vibrating — changing from one picosecond (a trillionth of a second) to the next. MD computes the forces acting on each atom at each instantaneous change of time (using Newton's equations of motion) and gives detailed information about these infinitesimal movements. To get realistic results, it is important to consider not only the protein itself but also the water molecules that surround it and influence its shape. This means that thousands of atoms need to be included, and a single MD simulation can take hundreds of supercomputing hours.

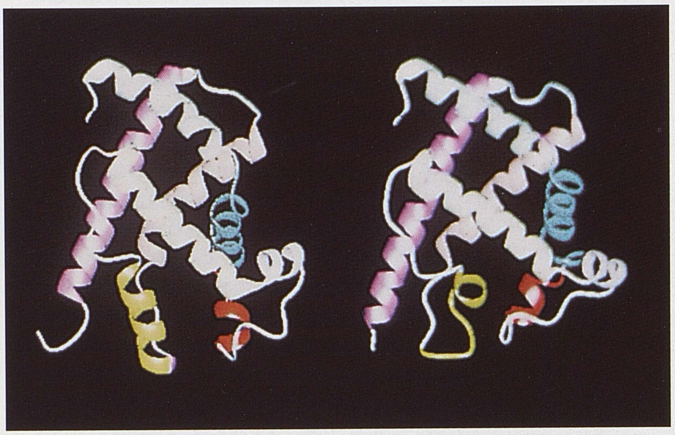
As supercomputing hardware and software have evolved over the last few years, Brooks has begun using MD to simulate processes involved in how proteins unfold, which is believed to mirror the rules of folding. Because some of these events occur on such short time scales, laboratory methods have been unable to study them, and Brooks' computations exploring early stages of unfolding are in many ways like experiments. Using the supercomputer as a laboratory, he gathers data about a fundamental process that isn't well understood.

"We're exploring phenomena that haven't been explored before," says Brooks. "There's not a well defined model in which we provide input and get the answer. We don't know in detail what we're looking for, so we're coming as close to doing experiments as you can without being in the laboratory." Brooks was able to conduct a number of these experiments during the "friendly user" period of testing on Pittsburgh's new CRAY C90. He focused on apomyoglobin, a partially unfolded form of myoglobin, an oxygen-carrying protein that occurs widely in muscle tissue.



Proteins are the worker bees of the body. They serve as enzymes, to make reactions go faster. They carry oxygen and antibodies to fight disease. They do a thousand diffrent jobs. But before a protein can go to work, it must fold into the right shape.

"The relationship between protein sequence and three-dimensional structure is one of the primary unsolved problems in biology today."



An Unfolding Protein

These images, from simulations by Charles Brooks, show apomyoglobin unfolding in response to changes in pH. The ribbon diagrams indicate the main features of the protein's secondary structure, with adjacent helical segments in different colors. On the left is apomyoglobin at neutral pH and body temperature (pH 7, 310° K). On the right is apomyoglobin at decreased pH (near pH 4), showing unfolding in the F helix (yellow).

Exploring the Protein-Folding Pathway

Brooks' objective in his research on apomyoglobin is to extend and complement findings from recent nuclear magnetic resonance (NMR) experiments. Using sophisticated methods of NMR spectroscopy, these experiments give a rough picture of what parts of a protein are folded or not under different conditions. In particular, the NMR experiments indicate that as apomyoglobin is subjected to a progressively more acidic environment it unfolds. It seems to unfold, however, in a particular way — not in a continuous movement from folded to unfolded states but following a specific pathway, as Brooks puts it, with waystations along the way.

The experiments indicate that there are intermediate, stable states of the protein. Apomyoglobin's pathway ends at about pH 2, where it completely unfolds. But in a less acidic environment around pH 4, it appears to assume a partially folded, compact structure that Brooks refers to as a "molten globule." Part of his objective is to provide more detail about this and other intermediate structures. "Experimentally, people have begun to characterize these protein-folding intermediates as either metastable or transient structures. But there is a large lack of detailed structural models."

Brooks simulated apomyoglobin under five different pH conditions. The number of water molecules included ranged from 3,800 to 6,600 as the protein partially unfolds (requiring more water to surround it), and the smallest simulation included about 14,400 atoms in total. Each of these massive computations tracked the protein for at least 1.5 nanoseconds, with an instantaneous snapshot every two femtoseconds (a millionth of a billionth of a second) — resulting in at least 750,000 individual "frames" of data for each simulation.

These calculations would not have been feasible, says Brooks, without the "friendly user" opportunity on the C90. "I would not have been able to explore this system in this detailed way on conventional machines, and in this case 'conventional machines' extends even to the C90 under normal conditions of a large number of multiple users."

Brooks and his colleagues are analyzing this huge collection of data using interactive graphics and other tools, a task that will continue for months and perhaps years. Preliminary results have been encouraging. "The most obvious thing we see," says Brooks, "is strong, direct correlation at nanosecond time-scales between the motion of the three-dimensional structure and the NMR data. This was the first benchmark comparison we wanted to do — to give us confidence we're on the right track." (MS)

Reference:

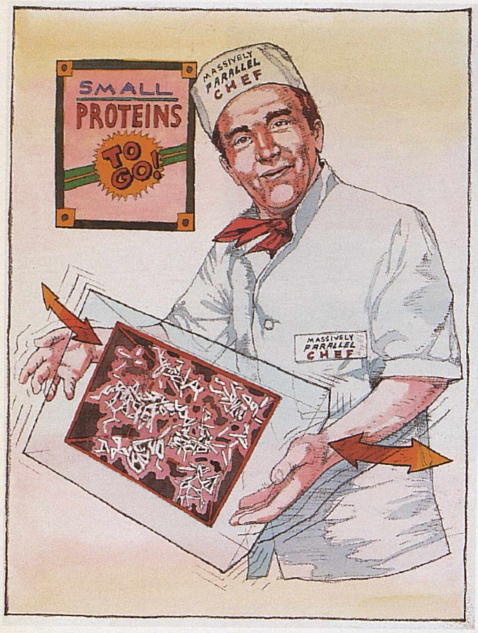
Charles L. Brooks III, "Characterization of 'Native' Apomyoglobin by Molecular Dynamics Simulation," *Journal of Molecular Biology* **227**, 375 (1992).

This research is supported by the National Science Foundation and the National Institutes of Health.

Shake and Bake

Parallel Solutions to the Phase Problem in X-Ray Crystallography

Herbert A. Hauptman and Charles Weeks, Medical Foundation of Buffalo, Inc.
Russ Miller, State University of New York at Buffalo



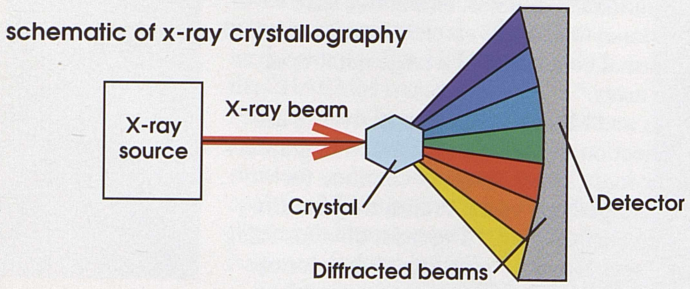
Automated Molecular Structure

A team of researchers led by Nobel prize-winning biophysicist Herbert Hauptman may be on the verge of cracking one of the premier problems in contemporary science — the phase problem of x-ray crystallography. Using massively parallel computing and an algorithm affectionately known as "Shake-and-Bake," they have achieved results that hold promise of cutting the time required to determine the structure of molecules that now take months and years to hours.

By throwing the bulk of the process onto a supercomputer, the Shake-and-Bake numerical recipe offers the possibility of an almost automatic path to solving unknown structures for medium-sized molecules. Structures in this range include small proteins and many antibiotics. "The central importance of this," says Hauptman, "is that it would enable us to design more effective drugs, either for the prevention or treatment of disease, and to do it in a rational, intelligent way, rather than by trial-and-error."

The Phase Problem & the Direct Method

Shake-and-Bake addresses a problem that has held up progress in molecular biology for decades. X-ray crystallography, the primary tool (and until recent years the only tool) for determining the structure of molecules, works by passing an x-ray beam through a crystallized molecule. The pattern of the diffracted x-rays, recorded as an array of spots on a photographic plate, depends on the arrangement of atoms in the molecule.



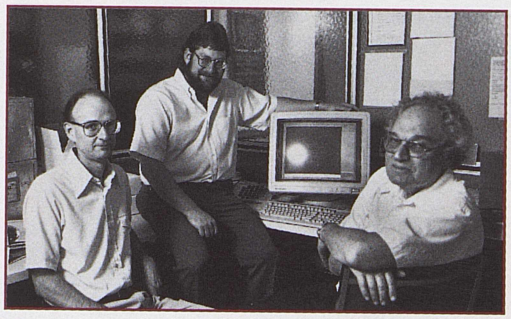
The problem is that the diffraction pattern registers only the intensity of the waves. To work back to the molecular structure, it's necessary to also know the relative timing when each wave hits, called "phase data" — the position of the wave crests and troughs relative to each other. Without the phase data, it often takes years of painstaking trial-and-error work relying on informed intuition to hit on the right structure.

Hauptman and Jerome Karle won a Nobel prize in 1985 for their work in the late 1940s and early 1950s on "the direct method" — a mathematical approach that makes it possible to glean phase data from the diffraction intensities. The key insight behind the direct method derives from the "atomicity" of molecules — i.e., atoms are small, discrete points relative to the spaces between them — which limits the possible relationships between phase and intensity to a range of probabilities.

The direct method has made it relatively routine to determine structure for molecules of 100 or fewer atoms. As molecular size increases, however, the probability relationships become weaker, and the direct method breaks down for molecules larger than about 150 non-hydrogen atoms. Another method, based on inserting heavy metal atoms into the crystal structure, has worked reasonably well for very large molecules — more than about 600 atoms. But for molecules in-between, ranging from about 200 to 500 atoms, structure determination remains a hit-and-miss proposition that takes months or years if it succeeds at all.

Massively Parallel Shaking and Baking

The aim of Shake-and-Bake is to extend the direct method to larger molecules. It relies on a reformulated mathematical approach, devised by Hauptman in 1988, that involves minimizing the value of a function involving thousands of phases. To make the scheme workable, Charles Weeks of the Medical Foundation of Buffalo automated the solution process and computer scientist Russ Miller of SUNY Buffalo headed the effort to develop algorithms for massively parallel computers. "We need an enormous amount of computational power," says Miller, "and the only way to get it is through massively parallel computing."



Charles Weeks (left), Russ Miller and Herbert Hauptman.

The process begins with a trial molecular structure. "You essentially toss the atoms into a shoebox," says Miller, "with enough knowledge to make sure the structure makes chemical sense." The computer calculates a set of phases corresponding to these atomic positions and then randomly perturbs them, the shake part of the process, to arrive at a lower value in accord with Hauptman's formula. Then a new set of atomic positions is calculated, and the method continues, back and forth between atomic positions and phases as many as 200 times for a single trial structure.

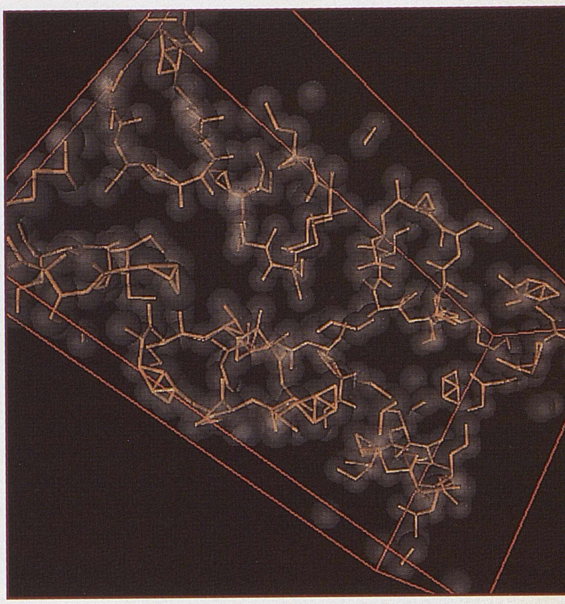
"The central importance of this is that it would enable us to design more effective drugs, either for the prevention or treatment of disease."

When enough trial structures have been shaked and baked, the right answer will emerge as a significantly low value for Hauptman's formula, at least that is the theory, and so far the researchers are, in Hauptman's words, "very much encouraged."

Encouraging Results

Since the beginning of 1992, Hauptman's team has worked on the Connection Machine CM-2 (and CM-5 that was on site most of 1992) at the Pittsburgh Supercomputing Center and a CM-5 at Thinking Machines in Cambridge, Mass. They have applied their procedure to 15 molecules ranging in size between 28 and 500 atoms. Most of these structures were already solved, and the calculations were done to see if the new method worked. In all cases, Shake-and-Bake cooked up the right structure, usually in a matter of hours computing time.

A dramatic success came in July 1992, when for the first time Shake-and-Bake solved a previously unknown structure. The molecule was ternatin, a 110-atom compound similar to cyclosporin, the immune system suppressant used in transplant surgery. Russian scientists had worked for years without success on ternatin, and David Langs, a colleague of Hauptman's at the Medical Foundation of Buffalo, tried half a million possibilities over the previous two years using traditional direct methods. After picking a set of trial structures and assigning parameters, Miller fired up Shake-and-Bake before going to bed one night and in the morning had a value that looked right. Within a few hours, Langs worked out the details of the structure.

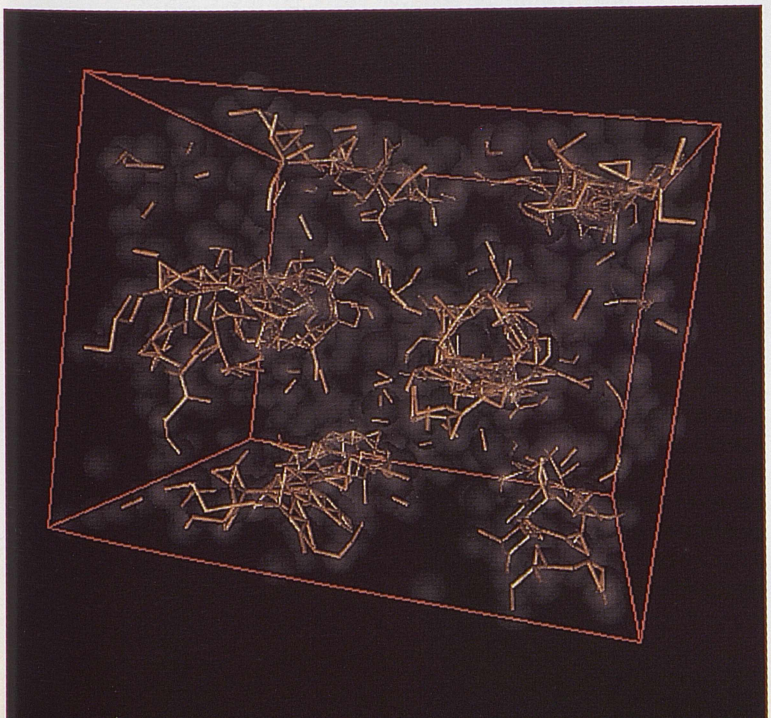


Unit-cell structure of ternatin as determined by Shake-and-Bake.

Another encouraging success came this year with crambin, a 400-atom protein — the largest structure ever solved by direct methods. Shake-and-Bake cooked up the right answer in 10 hours of CM-5 computing. "We have definitely strengthened traditional direct methods," says Hauptman, but he withholds final judgment. "We have yet to solve an unknown structure with 300 to 400 atoms. Once we have done 10 or 15 of those without failure, then we'll be in the position to say this works."

In the meantime, Miller is working on tightening his algorithm and cleaning up the code. "We're trying to get a handle on how many cycles we need for different size structures, and we're developing a clean, user-friendly version so crystallographers around the world will be able to use this." (MS)

The unit-cell molecular structure of gramicidin A, an antibiotic composed of 317 atoms, as determined by Shake-and-Bake. Gramicidin A acts as a membrane pore in cells, which is why the structure forms a tube along its central axis. To determine this structure originally (accomplished by David Langs of the Medical Foundation of Buffalo) required about 10 years using traditional direct methods. With Shake-and-Bake it took several months. Semi-transparent spheres represent atoms and the short lines represent possible chemical bonds. To arrive at a final, refined structure, the Shake-and-Bake solution is fed into a crystallographic refinement program.



Reference:

R. Miller, G. T. DeTitta, R. Jones, D. A. Langs, C. M. Weeks, H. A. Hauptman, "On the Application of the Minimal Principle to Solve Unknown Structures," *Science* **259**, 1430 (1993).

Gramicidin A and ternatin graphics courtesy of Robert Jones, Thinking Machines Corp.

This research is supported by the National Institutes of Health and the National Science Foundation.

Shaking Hands with DNA

The Refinement of *trp*-Repressor Structure by Simulated Annealing and Restrained Molecular Dynamics
Oleg Jardetzky and Daqing Zhao, Stanford University

"Learn to love the questions themselves, like locked rooms, like tappings behind dark glass."

— Rainer Maria Rilke

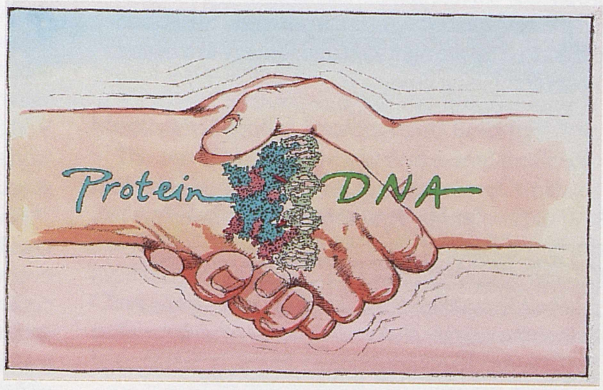
Proteins and Progress

When Watson and Crick worked out the three-dimensional structure of DNA in 1953, their achievement opened doors onto a realm of scientific problems that until then had been glimpsed as if looking through a keyhole. Their work helped to spawn the science of molecular biology, and its primary task since then has been to determine the three-dimensional structure of thousands of very large, complex molecules, predominantly proteins, that govern the intricate web of biochemical reactions we call life.

Progress has been difficult, often involving years of work to solve the puzzles posed by a single protein. In the late 50s, Max Perutz and John Kendrew at Cambridge University won a Nobel prize for working out the structure of hemoglobin and its close cousin myoglobin — the first proteins solved. Since then researchers have solved hundreds more, slowly and steadily building an invaluable body of knowledge.

The ultimate goal of this work is to attack disease at its molecular roots. When relationships between molecular structure and biological function are understood well enough, it could be possible, for instance, to design proteins to lower blood pressure or halt the growth of malignant cells. These are rich possibilities for improving the quality of human life, yet the primary task today remains much as it was in the 1950s — to determine the three-dimensional structure of proteins.

Composed of 3,400 atoms, *trp* repressor is the largest protein yet solved by NMR methods. Jardetzky's group also did computations on *trp* repressor bound to the segment of DNA it interacts with, a complex composed of approximately 4,700 atoms. Their work yielded an accurate structure of this complex, by far the largest studied by NMR, and as a result, they have extended the viability of NMR methods.



Protein Structure from NMR

Oleg Jardetzky, director of the Stanford Magnetic Resonance Laboratory, and his colleagues have pioneered the development of nuclear magnetic resonance (NMR) as a tool in protein structure determination. For the past five years, Jardetzky and his coworkers have aimed at using NMR to solve the structure of a large protein, *trp* repressor, that is important in understanding how the synthesis of proteins from DNA is controlled. Computations at the Pittsburgh Supercomputing Center during the past year overcame the final stumbling blocks.

"The supercomputer is absolutely essential," says Jardetzky. "You have to do an extensive series of calculations, basically model building — exploring the conformational space available to the molecule within the constraints imposed by the NMR data." Jardetzky's group relied on code it developed, called PROTEAN, to place the atoms in roughly the right positions. For final refinement of the structure, they used X-PLOR (developed by Axel Brünger of Yale University) and CHARMM (Martin Karplus, Harvard) — molecular dynamics calculations that track how the atoms change positions over time.

Oleg Jardetzky, director of Stanford Magnetic Resonance Laboratory.

For years, x-ray crystallography has been the only method available for structure determination. Within the last five years, however, NMR has begun to take hold as a second technique that is complementary to crystallography and in some ways offers advantages — primarily that it doesn't require a crystallized form of the molecule. The protein is scanned in solution, its natural environment, and therefore — unlike a crystallized molecule — is free to move as it would inside the cell of a living organism.

By comparing their results to a previously obtained crystallographic structure of *trp* repressor, Jardetzky's group confirmed the basic structural features. The NMR results also showed, however, that flexibility in a particular region of the protein plays an important role in its biological function.

"The supercomputer is absolutely essential."

Key to a Good Fit: Flexibility

Trp is the abbreviation for tryptophan, one of the 20 amino acids essential for formation of proteins, and *trp* repressor is what its name suggests — a molecule that represses synthesis of tryptophan. *Trp* repressor has been isolated from a common bacteria, *E. Coli*, and is not active in humans. Nevertheless, it is an important model for understanding how biosynthesis is regulated.

"It is the smallest complete regulatory system so far available," says Jardetzky, "and at this stage of knowledge, it is of great interest as a prototype of how these mechanisms work." When two molecules of tryptophan bind to the protein, it in turn binds to a DNA segment that controls tryptophan production. "It binds DNA in the presence of tryptophan and dissociates if the tryptophan level is low," explains Jardetzky. "In this way, tryptophan regulates its own synthesis."

A major finding of Jardetzky's calculations is that the region of the protein that binds with DNA is flexible and mobile. "What seems to be happening is that when the molecule binds, it doesn't dock as a rigid body. It adjusts the shape of the binding domain to the DNA, and the DNA also adjusts to create the best mutual fit. It is like getting a handshake. You cannot do it with rigid hands. You need some flexibility at the point of contact."

Calculations on the *trp* repressor/DNA complex, furthermore, appear to have resolved a question about whether the protein actually contacts the DNA. The crystallographic structure suggested that interaction with DNA was mediated by water molecules. Other researchers, however, believed that the crystallized protein didn't reflect the reality of the natural environment, and several published studies concluded that direct contact was possible. "We have direct contact between the protein and DNA," says Jardetzky. "We do not see water-mediated contacts, and we expect that these results will settle the matter." (MS)

Reference:

D. Zhao, C. H. Arrowsmith, X. Jia & O. Jardetzky, "The Refined Solution Structures of the *E. coli trp* Holo- and Aporepressor," *Journal of Molecular Biology* **229**, 735-46 (1993).

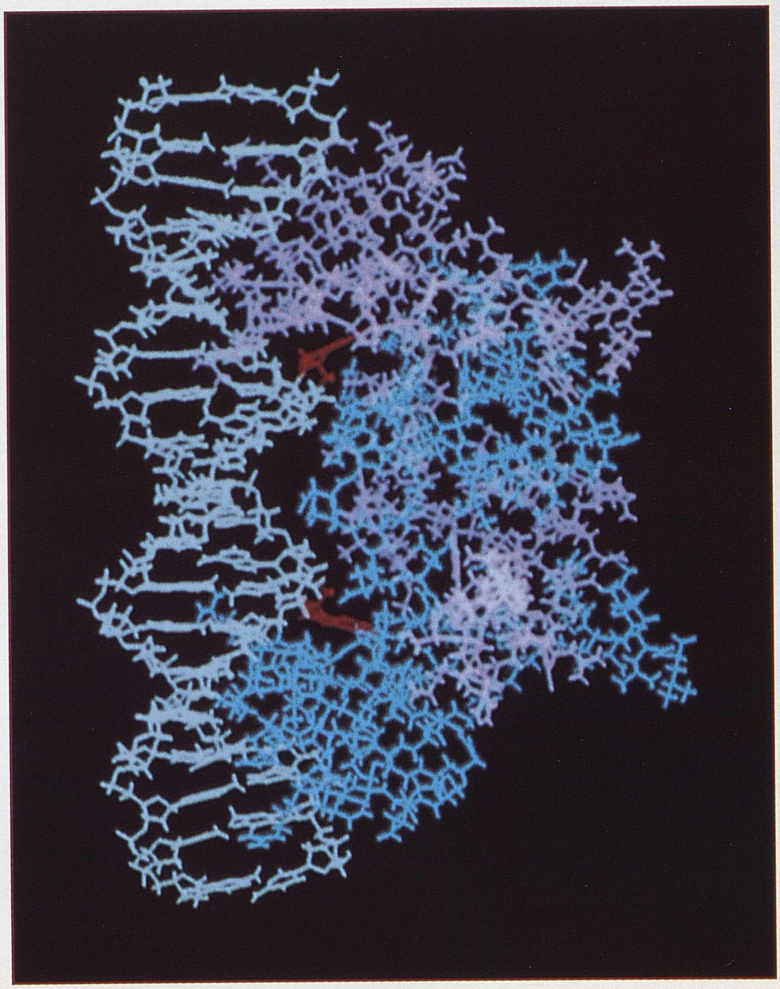
This research is supported by the National Institutes of Health and the National Science Foundation.

Trp Repressor and DNA "Shaking Hands"

This graphic represents the structure of trp *repressor bound to DNA as determined from NMR data obtained by Oleg Jardetzky and colleagues at Stanford Magnetic Resonance Laboratory. Structurally,* trp *repressor is a symmetric dimer, a molecule formed of two identical subunits (light blue and pink) intertwined with each other. The red atoms represent the two tryptophan ligands, molecules that bind with the represssor before it binds with DNA (pale blue).*

Calculations at the Pittsburgh Supercomputing Center (conducted by Daqing Zhao) showed that the tryptophan ligands stabilize the otherwise more flexible DNA-binding region of the protein and hold it at an optimal configuration for binding with DNA.

The repressor regions directly above and below the tryptophan are the protein's "reading heads." They make direct contact with the DNA, wedging into DNA's "major groove" — the opening that forms between DNA's two helical backbones. The DNA, in turn, bends at each end to wrap around the protein.



Fighting the Flu

Computer Predicted RNA Secondary Structure for Influenza Virus Genes
M. Louise Herlocher, University of Michigan and St. Jude Children's Research Hospital

Flu Shot Blues

Each year, thousands of Americans receive flu shots, hoping to avoid the sniffles, stuffy head, fever, body aches and general misery of the flu. But the shots don't always work. Government agencies sometimes don't include the right viral strains in the vaccines. The elderly often don't develop an adequate immune response from the shot. And children, who usually give the flu to mom and dad, don't even get inoculated because three or four shots are needed to achieve immunity.

One potential solution is to develop a vaccine from live virus, which would have more potency than the "killed virus" vaccines currently used. M. Louise Herlocher, a researcher at the University of Michigan, used the CRAY Y-MP at the Pittsburgh Supercomputing Center, to elucidate the nooks and crannies of the complexly folded structure of a live flu virus, and her results for the first time indicate that these folds may be a factor in how the virus reproduces inside human cells.

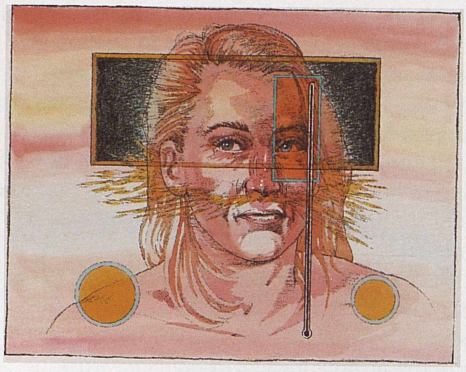
Chilling the Flu Virus

"You get better protection from a live virus vaccine," explains Herlocher, who recently completed a two-year postdoctoral fellowship at St. Jude Children's Research Hospital in Memphis, Tenn. "With a killed vaccine, which can't reproduce in the body and can't infect the host, you get a circulating immune defense. But with the live vaccine, you get a localized response in the nasal passage, which is important because that's where the virus enters the body. For instance, one dose of live vaccine, administered by nose drops, is sufficient to immunize children. So it's always been a goal to develop a live virus that would be safe."

M. Louise Herlocher

About 30 years ago, Herlocher's mentor at Michigan, John Maassab, took a flu virus that thrived at body temperature and forced it to reproduce at successively lower temperatures. The result was a cold-adapted virus that can't grow in the lungs, but can grow in the much cooler nasal passage, where a localized immune response can destroy invading viral particles.

During the last 20 years, about 15,000 people at six U.S. research hospitals have received this vaccine, and the results have been favorable. "People don't get sick from it, and it doesn't pass to family members," Herlocher says. Most important, when the vaccine is given in conjunction with the killed vaccine in the elderly, it greatly improves their immune response.



As compared to months and years of lab work, the CRAY requires only 30 minutes to determine the RNA fold of a single gene.

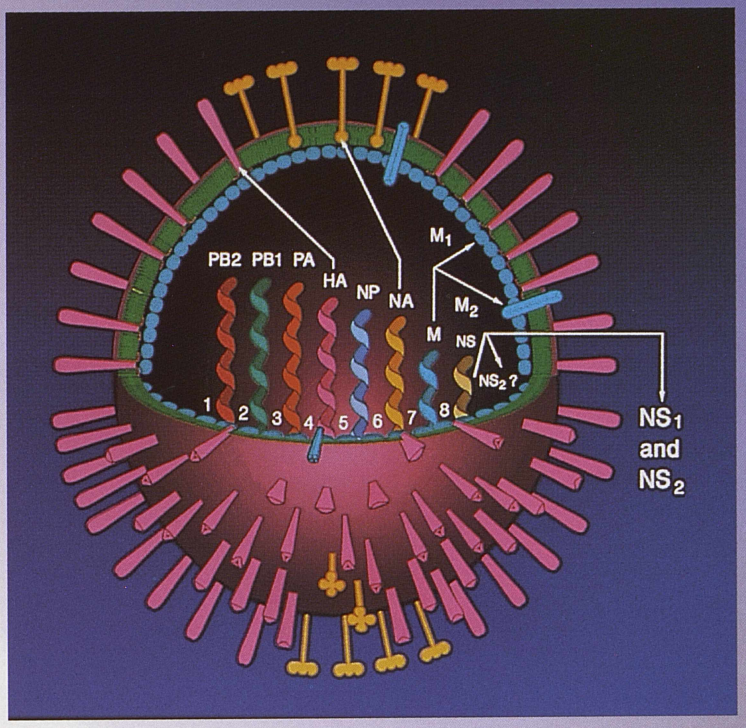
Exploring a Family Relationship

Herlocher has sought to find what features of the structure of this cold-adapted virus could explain why it lives happily only in nasal climes while its parent thrives in the moist, warm folds of the lungs. In particular, her work focuses on the eight strands of genetic material — composed of ribonucleic acid (RNA) — packed within the flu particle's protein and lipid outer shell. These eight genes are comprised of long strands of thousands of chemical bases, and the sequence of these bases determines what amino acids will be synthesized to form the viral proteins. The long RNA strands flop around and become looped, folded structures, and the shape of these complex folds may play a role in the activity of the virus.

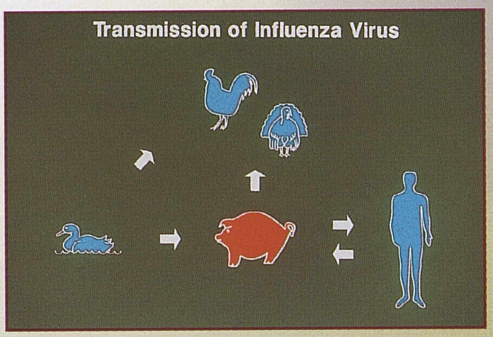
"We're looking at whether we can attach function to folding patterns," she says. "For instance, we want to determine how each viral particle can package its genetic material and how the RNA fold might accomplish this function. In addition, the RNA folding may affect the transport of genetic material in and out of a host cell's nucleus where the viral RNA is reproduced." It may turn out, Herlocher says, that the cold-adapted virus loses its virulence because it's transported or packaged poorly at the higher temperature of the lungs.

In Robert Webster's laboratory at St. Jude, Herlocher did a sequence analysis of the genes in the cold-adapted virus and its parent, comparing all the chemical bases (10,398) with the corresponding bases in the other. She found only four differences or mutations in the cold-adapted virus. She then did detailed RNA-folding calculations at Pittsburgh. Using the CRAY Y-MP and a program that predicts RNA folding developed by Michael Zuker of the National Research Council of Canada, Herlocher did the folding simulation for all eight genes of the cold-adapted virus and its parent at body temperature (37° C). Results showed different folding in two genes, and one of them (PB2) is known to control temperature sensitivity. Though this gene contained two of the four mutations, it synthesizes the same amino acids as the parent. This suggests, Herlocher says, that the RNA folds rather than purely the amino-acid sequence have a significant role in cold adaptation.

Herlocher is now doing laboratory analysis to confirm the RNA-folding predicted by her calculations. This work required 30 hours of CRAY Y-MP time and, says Herlocher, wouldn't have been practical without supercomputing. "It takes 12 hours to do a half gene on a VAX, and this system doesn't have enough memory to do a full gene." Laboratory methods are also limited. "It's impossible to determine every possible RNA fold in the laboratory," and as compared to months and years of lab work, the CRAY requires only 30 minutes to determine the RNA fold of a single gene. (SE)



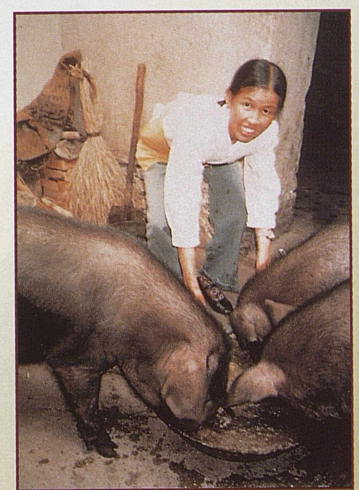
This representation of the flu virus shows a cutaway of the viral shell, coated with protein (blue) inside a lipid bilayer, to reveal the eight genes. Pink and yellow spikes represent hemagglutinin and neuraminidase respectively. These components control the virus' ability to attach to and release from a cell, and they constantly change — making a new vaccine necessary each year.



Reference:

M. Louise Herlocher , Hunein F. Maassab and Robert G. Webster, "Molecular and Biological Changes in the Cold-Adapted Master Strain A/AA/6/60 (H2N2) Influenza Virus," *Proceedings of the National Academy of Sciences,* (in press).

Research at University of Michigan supported by the university and the National Institute of Allergy and Infectious Diseases. Research at St. Jude supported by the hospital's Cancer Center Support Core Group, American Lebanese Syrian Associated Charities and the National Institute of Allergy and Infectious Diseases. Research at the Pittsburgh Supercomputing Center supported by the Biomedical Research Technology Program at the National Institutes of Health.



Blame it on the Ducks and Pigs

Current thinking is that flu evolves each year in southern China and spreads around the globe. "It's the way they farm, out in the villages," says M. Louise Herlocher. "They keep the pigs in the house, and they raise the ducks right next to the pigs, so there's a lot of opportunity for close interaction of the three species. The theory is that influenza goes from the avians to the pigs, from the pigs to the humans."

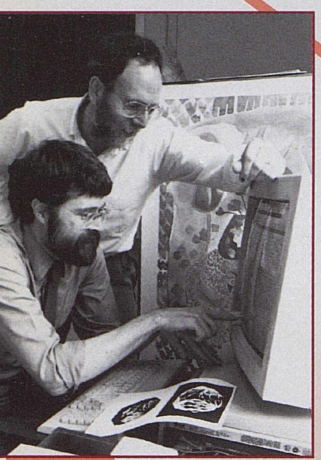
In support of that theory, researchers at St. Jude Children's Research Hospital have isolated avian and swine flu genes in humans and avian flu genes in pigs, but they have not demonstrated the presence of swine or human flu genes in birds.

At the Pittsburgh Supercomputing Center, Herlocher is using the Zuker fold analysis program to determine the RNA folding of the smallest (890 bases) flu gene in the three species. It's known as the non-structural (NS) gene, and its function is not well understood. To accomplish this, Herlocher is using 70 genetic sequences — avian, swine and human — from a database at St. Jude.

"This may or may not tell us anything about host adaptation," Herlocher says, but the folding may highlight which areas of the NS gene's strand are the same in all three species, information that could in turn help pinpoint which parts of the RNA play an active role in disease-producing functions of the virus.

Heart Throb

Cardiac Fluid Dynamics and the Immersed Boundary Method
Charles S. Peskin and David M. McQueen, New York University



Charles Peskin (standing) and David McQueen

Happy Birthday

It's not everyday new life comes into being as a result of calculations on a supercomputer. But it's not much of an exaggeration to say that's what happened when Charles Peskin and David McQueen cranked up their heart model for the first time on the Pittsburgh Supercomputing Center's spanking new CRAY C90.

"It worked!" exclaimed Peskin in February as he watched the visualized results of a "monster calculation" made possible by the C90. As the model heart contracted, twin streams of red and blue particles emerged from the left and right ventricle into the aorta and pulmonary artery. "To our surprise and delight," says Peskin, "the right side of the heart, which previously wasn't ejecting, is now ejecting too."

Peskin and McQueen took advantage of the "friendly user" period of testing on the C90 to put their model through its paces. In a computation that took nearly a week of C90 single-processor computing and 50 million words of memory (one fifth of the system's capacity), the main objective was to fix a problem that restricted flow through the aortic valve. By increasing the number of points in the mesh-like grid used to do the calculations, thereby creating a more realistic match between the detail of the computation and actual heart anatomy, the researchers expected to solve this problem, and they were right.

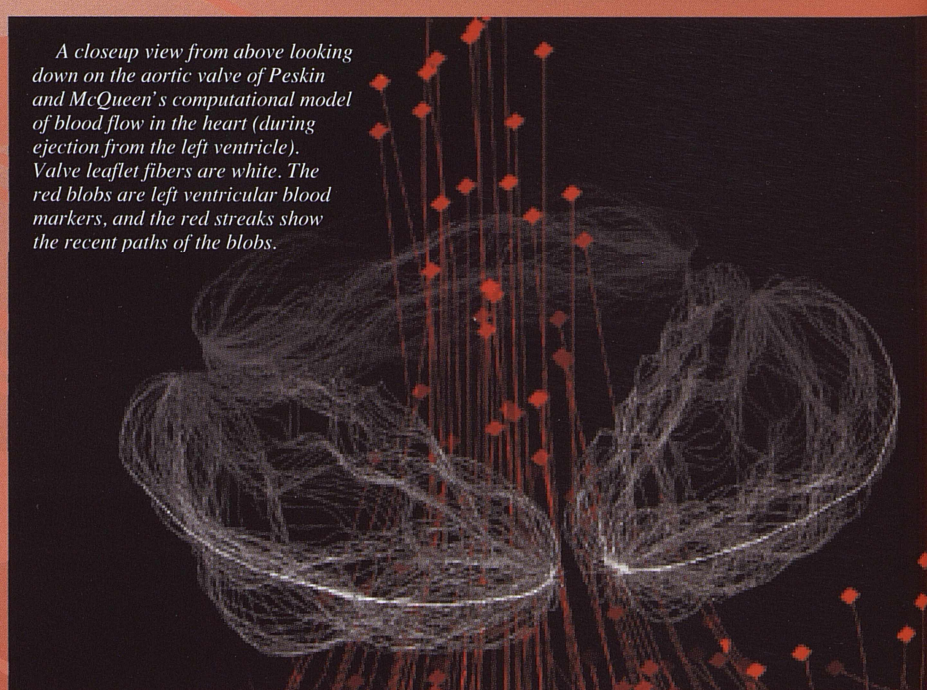
What they didn't realize was that improved resolution also would fix problems with the right side of their heart model, and the result was a breakthrough in their work. All the component parts of the model, each chamber and valve, did its job in cooperation with the other parts through a complete cycle — a full computational heartbeat. After 15 years of development, beginning with two dimensions and for the last five to six years working on the much more difficult three-dimensional version, Peskin and McQueen have achieved one of their goals: a fully functioning 3-D computational model of the heart, its valves and nearby major vessels.

Breakthrough Supercomputing

The breakthrough occurred as a direct result of improved computing technology in the form of the C90. Calculations on the predecessor supercomputer, the CRAY Y-MP (using a 64 x 64 x 64 computational grid), had progressed to the point where the researchers knew they needed improved resolution to realistically model the valves.

"In our computation," says Peskin, "there's a regular cubic lattice on which the flow variables are stored, and there's the more complicated geometry of the heart muscle that is stored as a series of fibers." Peskin and McQueen pioneered this technique, the "immersed boundary" method, in creating their heart model, and it is now used by other researchers on a multitude of problems involving fluids interacting with a movable, elastic boundary. The heart, in effect, is immersed within the computational grid, and if a valve ring — a hole that fluid should flow through — isn't big enough compared to the size of the grid, flow can be partially blocked.

Doubling the grid size in each dimension (128 x 128 x 128) meant eight times as many grid points and, consequently, required eight times as much computer memory. On the Y-MP, the researchers found that the increased data spilled over into disk storage, and the additional input/output processing slowed the calculations to such an extent that a single simulated heartbeat would have taken three years to complete. On the C90, the same calculation took a matter of days, and with work underway to "parallelize" their code — so it can maximize use of the C90's 16 processors running at the same time — the researchers look forward to even faster turnaround.



A closeup view from above looking down on the aortic valve of Peskin and McQueen's computational model of blood flow in the heart (during ejection from the left ventricle). Valve leaflet fibers are white. The red blobs are left ventricular blood markers, and the red streaks show the recent paths of the blobs.

"You gotta have heart,
All you really need is heart."

"With the new computer, we have a chance to succeed we didn't have before."

"It turns out there was no way we were going to get good results with the coarser grid," says Peskin, "and we were stuck there because of the computer that was available. With the new computer, we have a chance to succeed we didn't have before."

The Next Phase: Asking "What if?"

Problems remain to be worked out, notes Peskin, before the model heart is completely healthy. Contractions of the atria and ventricles are too abrupt, and the left ventricle, in particular, produces pressure that is too high and then falls off too quickly. The valves, though functional, don't work as they should — perhaps a result of the pressure abnormality.

Still, Peskin is confident a few more monster computations will shake out the bugs, and he has received a Grand Challenge allocation on the C90 to do it. (Grand Challenge allocations provide computing time for large-scale projects through a review process involving all four National Science Foundation supercomputing centers.) "It will be a somewhat slow process because each run is so big, but the only way this work proceeds is trial-and-error. You see a result and that tells you what you have to adjust."

The researchers look forward to the next stage — when the model can be used to pose "what if" questions that are difficult to address through animal and clinical studies. For instance, a heart attack often results in weakening a region of the heart, and this weakened part of the muscle may actually stretch while the rest of the heart contracts. What effect does this have on overall heart function? Does the location make a difference? When does the condition require surgery?

It will be possible to ask similar questions about diseased valves, says Peskin. "If a valve is stiffened, how much effect does it have on valve motion? Likewise, if there is a hole in the valve, what effect will that have?"

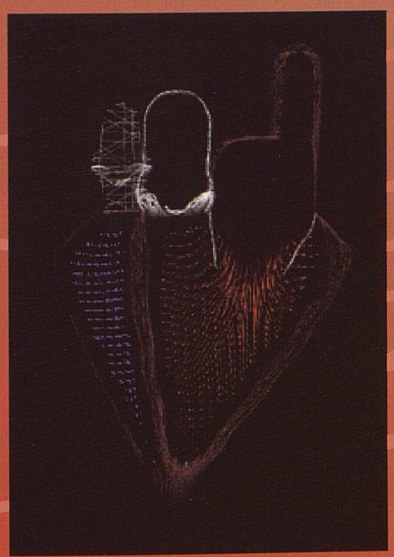
Peskin emphasizes that the model's ability to arrive at significant conclusions about normal and diseased heart functioning — and thereby to save lives — depends primarily on one factor. "The thing that sets the rate of progress in our project is the computer. If we had 10 times more computing time, or if the computer we have now were 10 to 100 times faster, we could use it effectively. We feel we're on the edge of starting to use this 3-D heart as a practical scientific tool. The C90 was a real breakthrough, but we're definitely in the market for similar breakthroughs in the future." (MS)

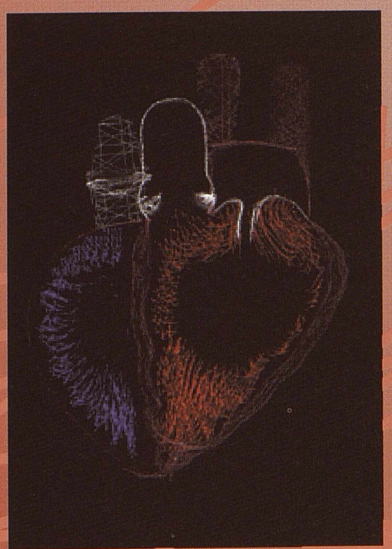
Reference:

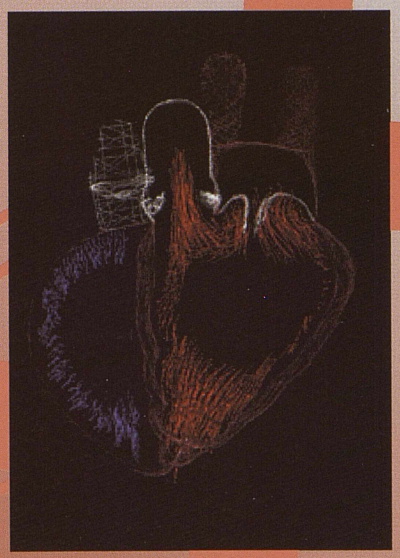
Charles S. Peskin and David M. McQueen, "Cardiac Fluid Dynamics," in *High-Performance Computing in Biomedical Research*, ed. T. C. Pilkington, et al. (Boca Raton, Fla.: CRC Press, 1993), pp. 51-59.

This research is supported by the National Science Foundation and has also been supported by the National Institutes of Health.

Computed flow patterns in a thin slice of Peskin and McQueen's three-dimensional heart model at different times during a single heartbeat. The slice is roughly through the middle of the left ventricle. The mitral valve (inflow) is on the right, and the aortic valve (outflow) is on the left. Red markers indicate oxygen-enriched blood (left-hand side of the heart) and blue indicates oxygen-depleted blood.





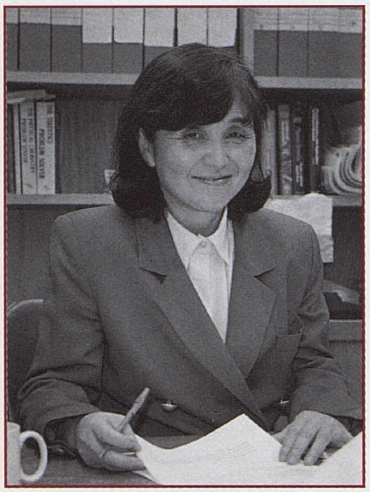


Your Cheatin' Heart

Cardiac Arrhythmias by Bifurcation Analysis
Teresa Chay, University of Pittsburgh

Personal Research

When we exercise, our heart rates increase, and when we sleep, they drop. But sometimes, when we're working during the day and minding our business, our heart skips a beat. During any 24-hour period, in fact, one-fifth of adults may experience this phenomenon and not know it. When extra beats happen in quick succession, the heart can enter a chaotic state, leading to a condition known as arrhythmia, which is treated with daily medication or, in more severe cases, electrical pulses that return the heart to normalcy.



Teresa Chay

Teresa Chay is a biophysicist who suffers from arrhythmia. She developed the condition as a child after rheumatic fever damaged two heart valves. Surgery seven years ago replaced the valves, but she still takes daily medication and occasionally needs to rush to the hospital for electrical pulse treatment.

Since her surgery, Chay has studied arrhythmia. Using supercomputers at the Pittsburgh Supercomputing Center, she has developed a mathematical model that explains why arrhythmia begins and why it stops when electric shocks are applied. Her goal, like other researchers in this field, is providing a foundation for developing better drugs to control arrhythmia. Currently, the drugs are not always effective, and in some cases, for reasons not well understood, rather than controlling arrhythmia they can trigger it and kill you.

Chaotic Hearts

The heart is a complex muscle with two chambers (left and right atrium) above two larger chambers (left and right ventricle) that are the prime pumpers. The beating action is regulated by the heart's natural pacemaker — the sinus node, a small piece of tissue at the top of the right atrium that sends out electrical pulses about once a second. Similar to a bucket brigade, the pulses stream into negatively charged muscle fibers lining the heart, creating a sudden electrical change beginning in the right atrium and extending into the ventricles. When the electrical pulses reach the ventricles, they contract, forcing the blood out. Thus, although the heartbeat seems like a single muscle contraction, it actually more resembles the wave phenomenon at a baseball game.

Research has shown that arrhythmia is triggered by an electrical impulse that arrives at the wrong time. The erratic beating starts when the mistimed signal travels in a continuous loop throughout the heart's muscle fibers — setting off what is called reentrant arrhythmia. If the reentrant arrhythmia occurs in the ventricles, the result is sudden cardiac death, which claims 250,000 U.S. lives each year. When it occurs in atrial tissue — known as atrial fibrillation, the condition can be treated with medication or electrical pulses.

Researchers have reproduced these results with dog hearts, but they still don't know why a small electric jolt throws a large muscle, like the heart, into a chaotic state. Instead, nothing should happen, Chay says, similar to when a person ignores a tap on the shoulder. And the key involves discovering why a mistimed pulse causes problems and why a properly timed pulse restores the normal heartbeat.

Chaos and Coexistence Theory

Because most functions in the human body, including the heartbeat, depend on other bodily events, such as hormone and chemical levels, the body is a dynamic system — very slight changes in startup conditions can result in drastically different outcomes, so that the system is virtually unpredictable. Recognizing this, Chay's research involves applying the insights of a branch of modern mathematical physics known as nonlinear dynamics or, more popularly, chaos theory. This approach involves examining the mathematical description of a physical phenomenon through the full range of its variables, to discern what patterns may exist within the apparent unpredictability. This field of study has come into existence only with the advent of powerful computers, because it requires a tremendous number of iterations of complex differential equations, and Chay's investigations can be accomplished only with supercomputing.

Chay first studied a simple system — just one cell and how it behaves as it gets sicker. In its healthy state, the cell is perfectly still and ready to receive an impulse from the sinus node. When that occurs, the cell sends the impulse to a neighboring cell and returns to a quiescent state. When the cell is sick, however, it holds too many positive charges, and under these conditions, Chay's calculations show that when an extra impulse arrives at the wrong moment — known as the vulnerable period — the cell begins sending its own signals across the heart, which interferes with pulses from the sinus node. As a result, arrhythmia kicks in.

Chay's studies showed this coexistence of states — dubbed the "coexistence theory" — for the first time.

Even for a sick cell, however, when a pulse arrives outside the vulnerable period, nothing happens. Arrived at using an approach from chaos theory called "bifurcation analysis," these results indicate that the beating and quiescent states of a single cell coexist; both possibilities are present, and which of the two occurs depends on the timing and magnitude of the triggering impulse. It is similar to the coexistence of water and ice when temperature is at the melting point. "At a certain level, you have two phases coexisting," Chay says, "so I was able to explain that complex dynamics exist even in a single cardiac cell." Her studies showed this coexistence of states — dubbed the "coexistence theory" — for the first time.

Multiple Cell Models

To make her model more realistic — for studying reentrant arrhythmia, Chay added a second cell. In this model, both cells begin at rest, then begin beating when an electrical pulse arrives from the sinus node. But a mistimed pulse causes signals to zing back and forth between the two cells, similar to reentrant arrhythmia — though it's not continuous and eventually stops. Since the two cells either beat independently or coupled for a short time, these results lend further support to the coexistence theory.

Calculations on an even more realistic, six-cell model, showed the coexistence of several states — reentrant arrhythmia, self beating and quiescent state — all of which can be triggered by electrical impulses given at the precise time and magnitude. "When a cell has several states coexisting, it becomes very susceptible to small perturbations," Chay says. "When cardiac tissues are in a healthy state, there's no coexistence. And even if you keep shocking it, nothing happens."

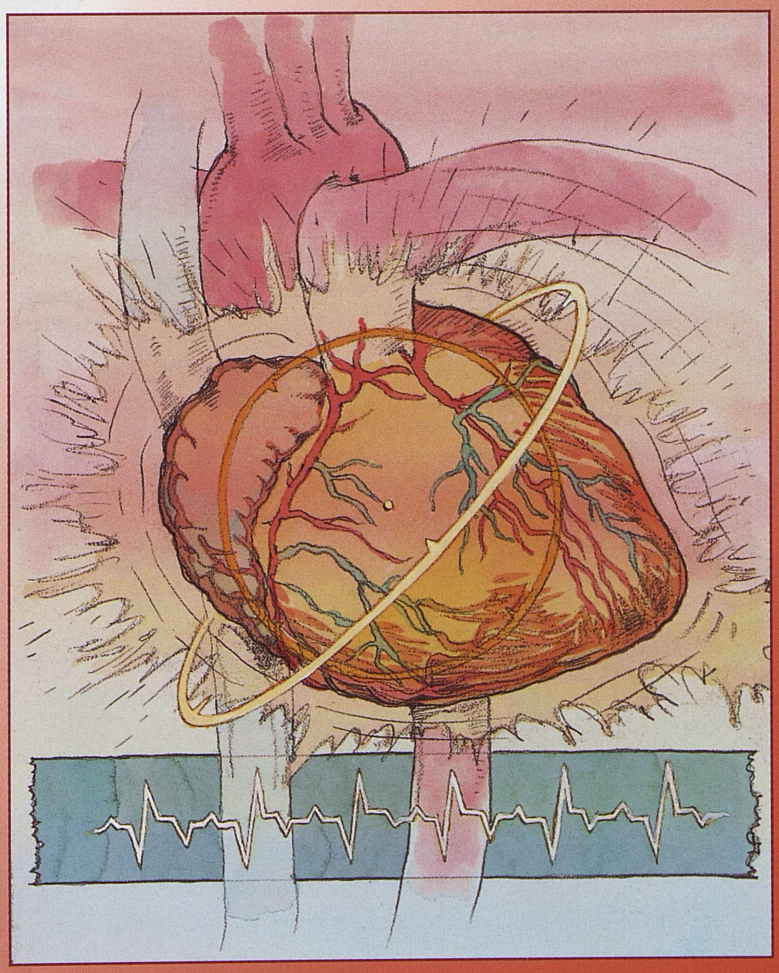
Now that Chay has developed a theoretical model to explain arrhythmia, she has extended the model to larger systems and plans to examine anti-arrhythmic drugs. Essentially, the drugs block channels that allow ions to flow into and out of cells. These moving ions, which are molecules with a positive or negative charge, create an electrical imbalance that in turn alters the passage of electrical pulses. "By blocking channels," Chay says, "I can see how to prevent arrhythmia, and I'd like to explain everything, instead of doing an experiment in the laboratory. I'd like to see precisely how the bifurcating phases change." (SE)

References:

Teresa Ree Chay, "Modelling for Non-Linear Dynamical Processes in Biology" in *Patterns, Information and Chaos in Neuronal Systems.*, ed. B.J. West (River Edge, N.J: World Scientific Publishing, 1993).

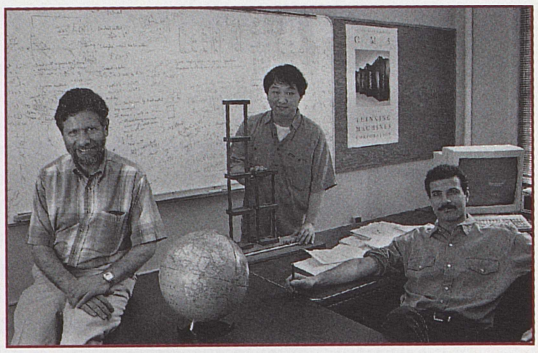
Teresa Ree Chay and Young Seek Lee, "Studies on re-entrant arrhythmias and ectopic beats in excitable tissues by bifurcation analyses," *Journal of Theoretical Biology* **155**,137-71 (1992).

This research is supported by the National Science Foundation and the American Heart Association.



When the Earth Moves

Earthquake Ground Motion Modeling in Large Basins
Jacobo Bielak and Omar Ghattas, Carnegie Mellon University



Jacobo Bielak (left), Xiaogang Li and Omar Ghattas.

Back to the Drawing Board

Minutes before the 1989 World Series game was scheduled to begin, the ground shook Candlestick Park, and millions of TV viewers watched in disbelief. The San Francisco Bay Area had been rattled again — this time, claiming 62 lives, injuring 3,757 people and causing $5 billion in damage. Four years earlier, another strong quake devastated parts of Mexico City. Some 300 buildings collapsed, and more than 10,000 people died.

These calamities have forced scientists to rethink what they know about making structures safe from earthquakes, and they're using supercomputers to more accurately gauge how earthquakes shake the Earth. The eventual goal, says Jacobo Bielak, professor of civil engineering at Carnegie Mellon University, is to develop three-dimensional models for earthquake-prone areas that pinpoint within a few city blocks — instead of several square miles — how the Earth moves, because ground motion can vary significantly over short distances. During the 1989 Loma Prieta, Calif., quake, for instance, scientists recorded significantly different levels of shaking only 30 meters apart.

"The 1985 Mexico earthquake was the strong one," says Bielak, "that really got people to think, 'Are we doing things right?'" Using massively parallel computing (the Connection Machine, CM-2) at the Pittsburgh Supercomputing Center, Bielak, assistant professor of civil engineering Omar Ghattas and graduate student Xiaogang Li have developed a three-dimensional model that helps explain the devastation of earthquakes such as the one in Mexico City.

Wreckage of an apartment in San Francisco's Marina district after the October 1989 quake.

Magnitude Can Be Misleading

The Mexico City quake registered 8.1 on the Richter scale; anything above 6 is considered serious, with each one-point rise meaning a 10-fold increase in strength. "But magnitude alone can be misleading," Bielak says. Seismologists measure the number 100 kilometers from the earthquake's epicenter, assuming the seismic waves from the quake travel only through rock and nothing else gets in their way. "So for design purposes, magnitude doesn't tell the full story," Bielak says. "Mexico City was an eye opener from this point of view."

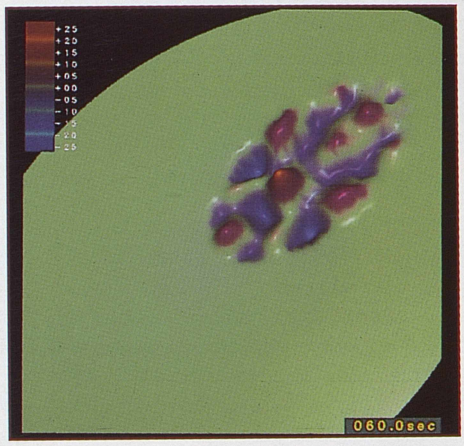
Since the 1950s, the building code in Mexico City recognized that the city had three distinct types of soil. Downtown is built primarily on a lakebed basin containing soft clays, surrounded by dense sands and rock, and engineers used this geological information to design buildings. Soil in each region vibrates at a different characteristic frequency when seismic waves disturb it, and the same is true with structures. So one important design criterion is making sure a building's vibrational frequency doesn't harmonize too closely with that of the soil it's built on — usually accomplished by stiffening the structure's steel or concrete frame. Otherwise, the building and ground may vibrate in unison — called "double resonance" — and eventually the structure sways violently. This phenomenon occurs on park swings every day when you pump your legs in time with the swing's oscillations to go higher and higher.

Although designers had addressed resonance concerns in Mexico City's three soil regions, they weren't precise enough, and several unexpected events occurred. First, the seismic waves resonated with the basin's soft clay, meaning the earthquake's strongest part lasted 40 seconds, instead of the normal five to 10 seconds, and the entire quake lasted about 200 seconds. As a result, there was ample time for the basin's clay to begin quivering like a bowl of jello being swished back and forth. This phenomenon, which wouldn't have been nearly as pronounced during a shorter quake, created an unequal distribution of ground motion; and unfortunately, some sections of the basin resonated with structures above, causing 99 percent of the quake's overall damage.

These calamities have forced scientists to rethink what they know about making structures safe from earthquakes, and they're using supercomputers to more accurately gauge how earthquakes shake the Earth.

This snapshot depicts horizontal ground motion in an hypothetical basin 60 seconds after the onset of an 80-second earthquake caused by southwesterly seismic waves lasting 10 seconds. The image — from a video made at the Pittsburgh Supercomputing Center — shows large variations in ground motion within the basin.

The color-coded graph indicates how far a particular area moves relative to the maximum distance that a rock formation moved outside the basin after seismic waves struck it. The red-orange region in the basin's center, for instance, moved a distance 25 times greater than the rock formation. Positive values indicate easterly movement; negative values westerly movement.



"So the questions came up, 'Do we know enough about ground motion? Do we need to know more? Should we subdivide the cities into smaller regions if a certain part is going to respond much more strongly than another one?'" Bielak says. "The work that we are performing here is geared to answer those questions."

Exploring 3-D Earthquake Simulations

Before Bielak and his colleagues began their study, other researchers had done little work on simulating three-dimensional, sedimentary basins with irregular qualities. And although most ground-motion simulations involved vector computers, such as the CRAY Y-MP, they chose instead to pursue their research on the massively parallel CM-2. "Massive parallelism has the promise of scaling to bigger and bigger machines," Ghattas says. "We know that ultimately this problem will require a speed of one trillion calculations per second and hundreds of millions bytes of memory, and it's not likely that traditional vector machines will achieve that."

For their prototype model, they developed a cylindrical region with 60,000 grid points. They placed the hypothetical basin in the center of this region and gave it a higher density of points than the outlying areas — essential for achieving accurate results. To simulate realistic seismic motion, during which a significant number of trapped waves eventually stop careening around the basin and escape into the surrounding rock, the researchers placed absorbing dampers on the boundary of the computational region.

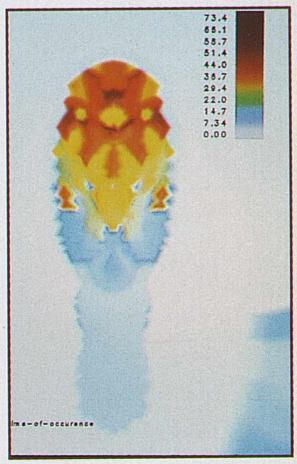
After developing an irregular grid to accommodate geological details, it was time to shake it up with a substantial seismic wave from the southwest. A series of preliminary runs served to refine the code, which uses the finite-element method, solving three equations at each grid-point. The final run took nine hours of CM-2 computing to simulate how three soil regions behave during an 80-second earthquake caused by seismic waves lasting only 10 seconds. A video made with the help of Pittsburgh Supercomputing Center scientific specialist Joel Welling clearly shows an uneven distribution of waves in the basin. The video also shows that areas with similar wave strength aren't active at the same time, thus different parts of elongated structures, such as bridges and tunnels, would move at different times, raising the odds of devastating destruction.

Bielak and his colleagues achieved an impressive sustained rate of 520 million arithmetic operations per second (using 32,000 CM-2 processors), yet they know they'll need more computing power for the project they're working toward: modeling the earthquake-prone Los Angeles basin, whose volume of 40,000 cubic kilometers will require two-billion grid points to get realistic results. Computing capability is the limiting factor, say the researchers, and they have set a five-year target to develop their code in anticipation of improved performance from upcoming generations of supercomputing. (SE)

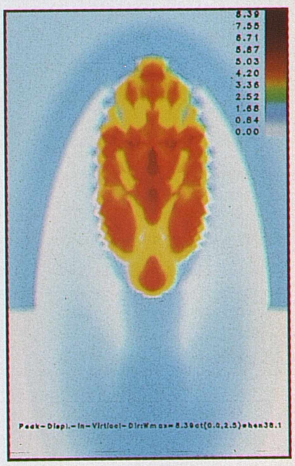
Reference:

X. Li, J. Bielak and O. Ghattas, "Three-dimensional earthquake response on a CM-2," *Proceedings of the Tenth World Conference on Earthquake Engineering: 19-24 July 1992, Madrid, Spain*, pp. 959-64.

This research was supported by the National Science Foundation.



These two images show peak vertical movement in the basin during the 80-second earthquake (top) and when that movement occurred in seconds (bottom). These results show that peak movements in adjoining segments occur at different times — a hazardous condition because different sections of elongated structures may undergo peak vibrations at different times.



Storm Warning

"Oklahoma, where the wind comes sweepin' down the plain."

Field Program to Validate a Thunderstorm Prediction Model

Kelvin K. Droegemeier, University of Oklahoma at Norman

Kelvin Droegemeier. "The service we received from Pittsburgh was remarkable. It's the way science ought to happen. You have an idea — and there's nothing as fragile as an idea — and you say OK, *we really want to do this, and it happened."*

Predicting Storms in Oklahoma

In Oklahoma, April through June means thunderstorms, huge ones that dump inches of rain in a few hours, swelling streams and flooding crops, businesses and homes. Storm season in Oklahoma also brews the ominous dark funnels that smash houses flat in seconds.

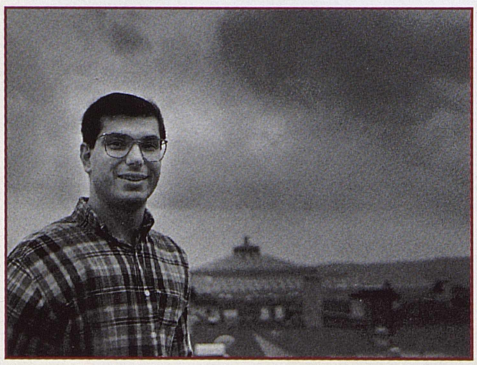
What would it be worth in reduced property damage and lives saved to have reliable warning, say four or five hours in advance rather than 30 minutes — as current forecasts give, of flooding or tornadoes headed in your direction? It's the sort of question that keeps Kelvin Droegemeier busy. An associate professor of meteorology and co-founder of the University of Oklahoma's Center for Analysis and Prediction of Storms (CAPS), Droegemeier leads a research team that has developed a state-of-the-art computer model for predicting storms, the Advanced Regional Prediction System (ARPS).

During May 1993, in a unique collaborative experiment, Droegemeier's team used the CRAY C90 to see how ARPS worked when run under the constraints of daily weather forecasting. Each stormy day for the entire month, the height of Oklahoma's storm season, numerical predictions from ARPS at Pittsburgh provided a real-time, direct feed of a variety of output products to forecasters at the National Weather Service's Experimental Forecast Facility in Norman, Okla. It is the first time a fully-automated system of this type has been used in an operational environment. "The model did a very good job," says Droegemeier, "of predicting storm type, motion and general characteristics."

By increasing the lead time for severe storm warnings, ARPS has the potential to reduce property damage and save lives.

Modeling at Storm Scale

Until now, storm-scale weather models have been used mainly for research, and an important goal of the experiment was to evaluate how ARPS would work in an operational setting. "We wanted to gain experience working with forecasters," says Droegemeier, "to understand what they need from a model like this as part of daily forecasting. It's like in football," he adds, "you can practice and you can scrimmage, but you don't really know how you're doing until you're in a game."



Paul Puglielli, user consultant at the Pittsburgh Supercomputing Center. Puglielli coordinated scheduling the ARPS model for daily runs during May using four of the CRAY C90's 16 processors. He set up a dedicated four-processor queue that met the project's need for reliable turnaround — from noon to 1 p.m. each day — under operational forecast conditions.

Although computer models are a standard part of weather forecasting, existing models operate at a much larger scale than ARPS. Models at the National Meteorological Center, for instance, predict atmospheric structure over the entire United States every 12 hours. These regional-scale models can't predict movement and severity of individual storms.

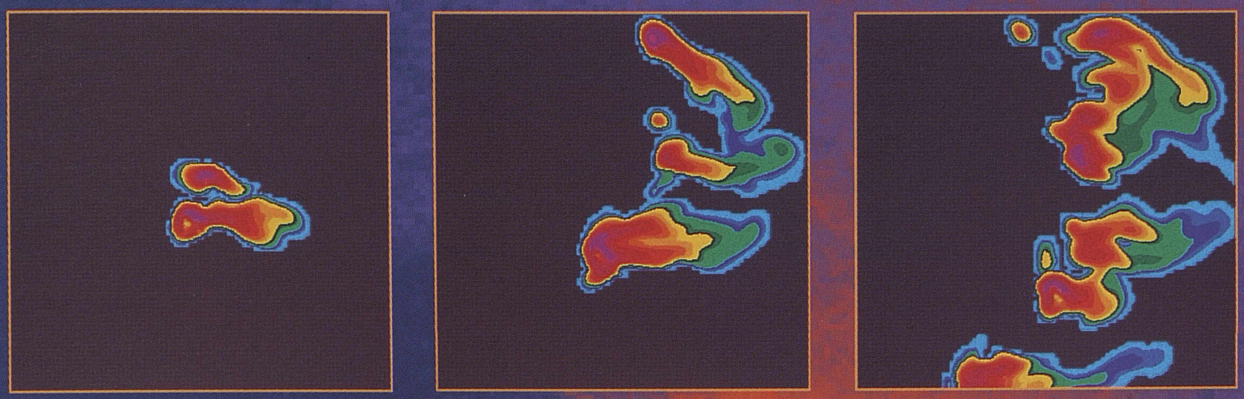
ARPS, on the other hand, is scaled to the size and duration of a storm. It predicts over a smaller area (up to 1,000,000 square kilometers) and gives detailed readings on key storm parameters — rainfall, wind direction and velocity, temperature and pressure, among others. It also predicts in time increments related to the way storms actually develop — every few minutes rather than every 12 hours.

Predicting in Real Time

During the month-long experiment, a dry run for a more formal multi-year evaluation of ARPS that will begin in 1994, forecasters in Norman determined the basic structure of the atmosphere each morning. If there was any potential for storms, these data became the starting conditions for an ARPS run. A CAPS researcher at Norman fed the data to the C90 by noon (central daylight time). By 1 p.m., ARPS output was available to forecasters in Norman, in time for the daily afternoon forecast.

To make possible the rapid, reliable turnaround required for the experiment, the center scheduled ARPS to run on four of the C90's 16 processors each day for the entire month. The model usually took about 45 minutes to complete its daily run, which simulated four hours of storm evolution. (ARPS runs at a sustained rate of over 400 million calculations a second on a single C90 processor.) Along with other output, the model sent forecasters in Norman a sequence of images they could display on workstations as a moving picture of how the storm would evolve during late afternoon and early evening.

Because the experiment was directed primarily at testing operational procedures, the model runs left out information — surface features of the land and Doppler radar data — that will be included in the more full-scale ARPS evaluation next year. Nevertheless, experienced forecasters found that even this simplified version of ARPS was useful. "It gave us an integrated estimate of rain water at various levels," says Paul Janish, a forecaster who participated in the experiment, "and we were able to evaluate other parameters to give us an idea what type of storm might develop." Forecasters gained enough confidence in the model that on May 21 they referred to it in an official National Weather Service release indicating the potential for intense thunderstorms over Oklahoma that evening.



Stormy Mother's Day in Oklahoma

These images show snapshots from the ARPS prediction for rainfall over central Oklahoma on Mother's Day, Sunday, May 9, 1993. Each image represents a 100 x 100 kilometer square horizontal slice through the storm four kilometers above ground level, showing what the storm would look like on a radar scope, where reflectivity corresponds to how hard it's raining. Color indicates rainfall intensity, increasing from light blue to pink.

Severe storms Saturday, May 8, and again on Sunday resulted in several tornadoes and heavy flooding in parts of Oklahoma, and ARPS successfully forecast this threat. It showed the development of "supercells" — storms with potential for rotating updrafts and tornadoes. The rainfall images show multiple supercells forming into a line. "The storm tracks were right on top of one another," says Droegemeier, "so a given point in space would get almost continuous rain. This agreed well with what was observed."

Saving Money and Lives

By increasing the lead time for severe storm warnings, ARPS has the potential to reduce property damage and save lives. It should also make it possible to give more precise information about impending storms. "It's one thing to say there are going to be thunderstorms," says Janish, "it would be another to say there's a likelihood of strong thunderstorms with potential tornadoes, or that tornadoes seem most likely in this part of the state and will form three hours from now." ARPS, says Janish, should give forecasters a better feel for the type of storm likely to form, where it will go and how long it will last.

A fully operational ARPS model could also reduce costs for the airline industry, which loses millions of dollars annually to weather-related flight re-routing and delays.

"Oddly enough," says Droegemeier, "there's no plan at the national level to run a model like this operationally. In some sense, we've been given a challenge to prove the usefulness of a forecast of this type. There are people who, for good reasons, question whether you can actually do storm-scale numerical predictions. The mission of CAPS is to demonstrate that capability." (MS, SE)

Reference:

K. K. Droegemeier, M. Xue, K. Johnson, K. Mills & M. O'Keefe, "Experiences with the scalable-parallel ARPS cloud/mesoscale prediction model on massively parallel and workstation cluster architectures," Proceedings, 5th Workshop on the Use of *Parallel Processors in Meteorology, European Center for Medium Range Weather Forecasts, Reading, England, 23-27 November, 1992* (in press).

This research is supported by the National Science Foundation and Cray Research, Inc.

Photograph of Kelvin Droegemeier: Thompson-McClellan Photography.

Cool Design

Concurrent Thermal Analysis and Design of Electronic Systems and Components
Cristina Amon, Carnegie Mellon University

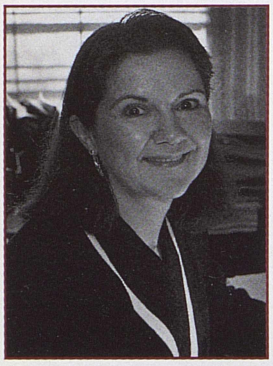
A Wearable Campus Guide

Next time you're in Pittsburgh and want a tour of the Carnegie Mellon campus — history and all — forget about tagging along with a tour guide. Just ask for VuMan, the first-generation wearable computer that scrolls campus information in front of your eyes as you manipulate three buttons on a belt pack.

Although VuMan has as much firepower as an IBM PC — all contained in its composite-metal headband — the $1,000 machine won't singe your hair. That's because its developers used a relatively new design approach, known as concurrent design, along with computer modeling performed at the Pittsburgh Supercomputing Center.

Instead of first embedding the electronic components into the headband and then worrying about how heat escapes, researchers at CMU's Engineering Design Research Center consulted with mechanical engineering colleague Cristina Amon from the outset. Amon and graduate student Jay Nigen modeled the thermal flow patterns of the preliminary circuit board designs on the CRAY Y-MP. "How do you locate the components within the enclosure so that it can be efficiently cooled?" asks Amon, who has seven years experience computing at the Pittsburgh Supercomputing Center. "How do you keep the temperature below what's required for reliability and safety?"

Previously, electrical engineers may have scratched an entire project and started from the beginning if thermal issues proved too difficult to solve. "With concurrent design, downstream concerns can be brought forward in the design cycle," Amon says, "with the objective of reducing failures and shortening the design cycle and time-to-market." For VuMan, it took only 12 weeks to design and build 30 units.

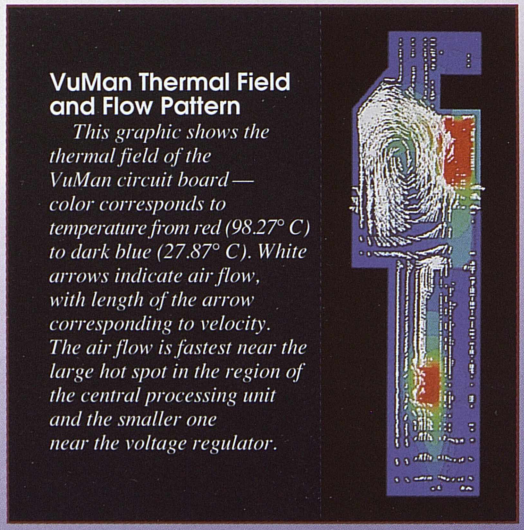


Cristina H. Amon

Electronic Cooling

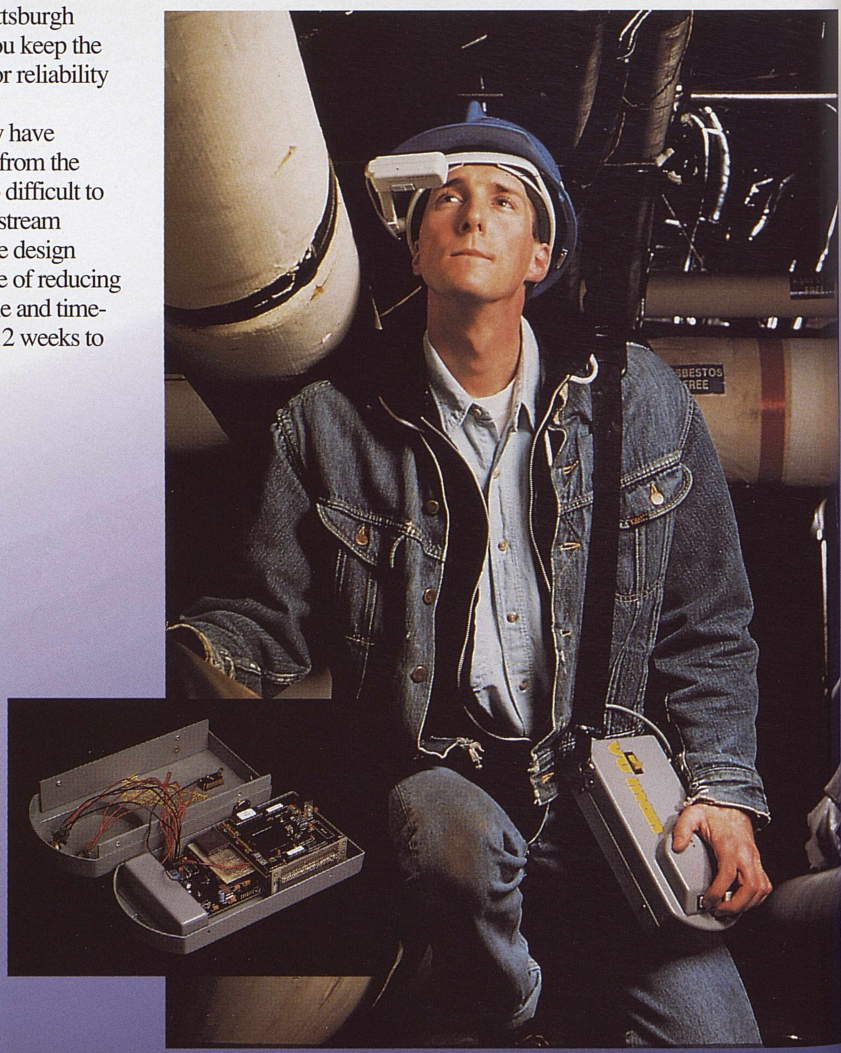
Advances in designing and manufacturing integrated circuits have brought about greatly improved performance through size reduction, leading to advanced electronic systems such as supercomputers. Today, over a million transistors fit on one semiconductor chip, with a billion foreseeable, and switching speed with these microcircuits is 10,000 times faster than it was 30 years ago. As chip circuitry becomes smaller and smaller, however, it produces more heat in a given area, and cooling requirements limit new product development. "Thermal considerations have become a controlling factor," Amon says, "and they rank among the major technical problems limiting the achievement of higher data processing speeds."

Amon's research group focuses on developing numerical methods to model thermal-flow phenomena. One important strategy, for instance, for cooling electronic components is to break up the flow, which tends to be fairly smooth. More complicated flow patterns tend to dissipate heat better, but they also require more sophisticated modeling techniques.



VuMan Thermal Field and Flow Pattern

This graphic shows the thermal field of the VuMan circuit board — color corresponds to temperature from red (98.27° C) to dark blue (27.87° C). White arrows indicate air flow, with length of the arrow corresponding to velocity. The air flow is fastest near the large hot spot in the region of the central processing unit and the smaller one near the voltage regulator.



Recently, Amon has brought her thermal modeling into conjunction with a novel manufacturing technique, and the combination creates a powerful design approach.

"To take full advantage of flow bifurcations and transitions," says Amon, "numerical techniques must be developed that accurately follow flow evolution." The numerical scheme must resolve sharp shifts in flow velocity and "hot spot" regions, which can degrade the performance of electronic components. To meet these challenges, Amon uses a technique known as the spectral element-Fourier method (SEFM). SEFM, she says, combines the geometric flexibility of finite-element methods (which use a mesh-like grid to calculate flow variables as a sum of many small, discrete regions) with the speed and accuracy of "spectral" methods (a mathematical approach that treats physical phenomena as a series of waves, each with a characteristic amplitude and frequency).

Because SEFM modeling requires massive amounts of computer firepower to solve realistic three-dimensional problems, Amon uses the CRAY to study thermal characteristics. Besides studying the electronic component itself, Amon looks at other segments of the integrated circuit, including solder and the circuit board. "The tangible results," says Amon, "are that the selection of the materials and where they're placed does significantly affect thermal performance."

One technique for dissipating heat, for instance, is placing metallic strands on a chip's edges. The strands, known as "heat spreaders," increase the area through which heat will dissipate. At the same time, they guide how heat is removed to the casing that encloses the circuit board. Amon has modeled the effect of this approach. "So here we have guidance," she says, "for what type of material, what type of shape and relative dimensions we need to achieve optimal heat removal."

Choosing the proper material for the heat spreaders, solder and circuit board is a tricky issue because engineers must consider cost, weight and performance. All these factors, as well as heat flow, must be considered as part of the design process.

Eliminating Casts

Recently, Amon has brought her thermal modeling into conjunction with a novel manufacturing technique, known as MD* thermal spraying, and the combination creates a powerful design approach. For the last two years, the Engineering Design Research Center has worked on developing a prototype headband for a more advanced VuMan, known as the Navigator. Previously, if the design team wanted to change the headband's shape and size, their only recourse was to make a new cast, which is both time-consuming and expensive. With MD* thermal spraying, they spray molten droplets to build complex structures comprised of layers of different materials on a surface. This process not only enables rapid prototyping of new designs, but also — in conjunction with Amon's modeling — it allows designers to optimize heat removal and to predict the operating thermal conditions.

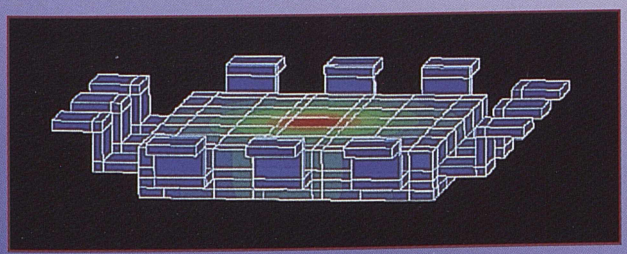
When the Navigator is operational, visitors to Carnegie Mellon no longer will need their hands to control their tour guide. The Navigator includes advanced voice-recognition circuitry that will allow it to respond to the human voice, and the more complex thermal-design problems posed by this circuitry present a new challenge for supercomputing. (SE)

References:

Jay S. Nigen & Cristina H. Amon, "Concurrent Thermal Designs of PCB's: Balancing Accuracy with Time Constraints," *IEEE Transactions on Components, Hybrids and Manufacturing Technology* **15**, 850 (1992).

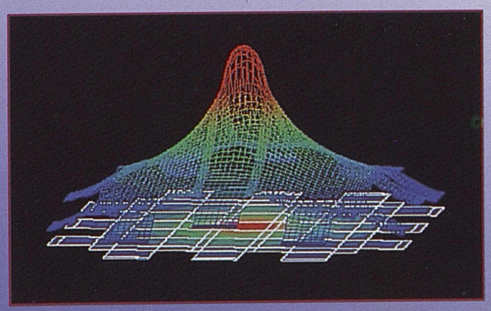
C. H. Amon, "Heat Transfer Enhancement by Flow Destabilization in Electronic Chip Configurations," *Journal of Electronic Packaging* **114**, 35 (1992).

This research is supported by the National Science Foundation, AT&T Foundation and Cray Research Inc.



Thermal Field with Heat Spreaders

These two graphics depict the thermal field of an electronic chip to which metallic strips called "heat spreaders" are attached to help remove heat. Color corresponds to temperature, with red being hottest. A hotspot occurs in the chip's center and the coolest spots are the heat spreaders themselves. For a cross section through the middle of the chip (below), the thermal field is also represented as a network structure with height corresponding to temperature.

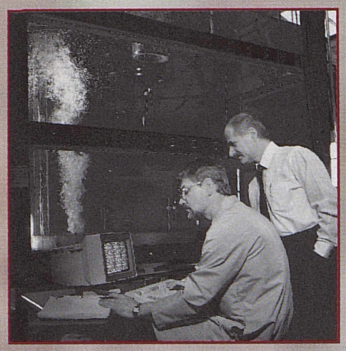


The Flow of Hot Steel

Modeling Steel Metallurgy in Continuous Casting Tundishes
Achilles Vassilicos & A. K. Sinha, U.S. Steel Technical Center

What Happens in a Tundish?

Computer modeling combined with physical modeling has paid off in steel manufacturing. That's the message from Achilles Vassilicos, research consultant at U.S. Steel Technical Center in Pittsburgh. Since 1989 when the Technical Center, the research arm of U.S. Steel, became a corporate affiliate of the Pittsburgh Supercomputing Center, U.S. Steel researchers have used the high-powered computing of the CRAY Y-MP, and now the C90, to help understand the flow patterns of molten steel.



Achilles Vassilicos (standing) with Ron Woodrow of the U.S. Steel Technical Center staff in the "flow lab," analyzing results from a physical model of gas stirring in a steel ladle.

The U.S. Steel research team wants to know what happens when a "heat" of steel — a huge ladle containing more than 200 tons of molten metal at close to 3,000° F. — empties its fiery brew into a continuous casting "tundish." In particular, they want to know as precisely as possible what happens inside the tundish as the molten steel churns and swirls around. The tundish holds the white-hot liquid and feeds it out the bottom into a continuous casting mold, where it forms a moving strand of steel that eventually cools from white to red hot and gets cut into slabs for further processing.

Continuous casting is the most up-to-date technology available for producing high-quality steel at low cost, and good understanding of what goes on in the tundish is critical because it affects the purity and chemistry of the output steel. Impurities, such as oxides of aluminum, calcium and iron, tend to float to the top of the tundish bath. The steel flow must be controlled to enhance this flotation and to prevent turbulence from drawing impurities back down into the bath. Furthermore, you need to know how the chemistry of the mix feeding out the bottom of the tundish varies as a new heat pours in the top.

"The objective is to have the caster running continuously," says Vassilicos, "and you usually aim for a string of several hundred heats. The chemistry often varies significantly from heat to heat. If you know exactly what is happening in the tundish in real time, you can precisely and intelligently disposition the output steel to meet the specifications of customer orders."

Modeling to the Rescue at Gary No. 2 Caster

U.S. Steel uses physical modeling — laboratory facilities include several large-scale models of continuous casting tundishes, ladle treatment stations and strand pouring systems — in conjunction with mathematical modeling on in-house workstations and on the CRAY at the Pittsburgh Supercomputing Center. "Using modeling to design improved processes has really paid off for us," says Vassilicos. "It has translated into hardware improvements running in our plants."

As one example, he cites improved grade intermix control at the No. 2 caster of U.S. Steel's Gary, Ind. plant: "One day our Gary works called. They wanted us to investigate if there was a way to predict the chemical composition of slabs produced during periods of steel grade change." (During a grade change, steel at the end of one heat mixes with steel of different chemistry from the next heat.) At Gary, because of the diversity of steel chemistries required by customer orders, many slabs produced during grade change had to be stored until they could be tested. By the time they were tested, even if the chemistry was good enough to meet customer specs, they lost their place in the processing line and ended up clogging inventory.

"Using modeling to design improved processes has really paid off for us. It has translated into hardware improvements running in our plants."

A U.S. Steel team devised a computational model based in a mixture of physical modeling, plant experiments and Y-MP computations. "We came up with a rather simplified model," says Vassilicos, "that can be used in a process computer, real time, to predict how the chemistry in the cast product changes as a function of time." Plant and research personnel implemented the model on the Gary No. 2 caster. It has proven highly reliable. Several months of trial run showed that the model predicted chemistry as accurately as laboratory testing of slab samples and it wasn't susceptible to errors inherent in the testing procedure.

With the new method, the Gary plant has significantly reduced the need to test for slab chemistry. "We have the confidence to supply our customers with quality steel grades on the basis of the model," says Vassilicos. "This is a cost savings because you save on the handling, you save on testing the chemistry, and you reduce the inventory."

Suppressing Tundish Turbulence

Another result of U.S. Steel modeling is a flow-control device called a "turbulence suppressor pad." As a result, U.S. Steel is the only company, probably worldwide, that can apply what is called "transition slabs" to the highest quality product manufactured — drawn and iron (D & I) steel used in thin-wall beverage cans.

Transition slabs occur because the initial flow into the tundish from a new heat — for about the first minute and a half — is much faster than the exit flow from the tundish to the strand. This initial high flow rate creates turbulence that draws impurities from the bath surface down into the liquid being cast, resulting in lower quality steel for one or two slabs.

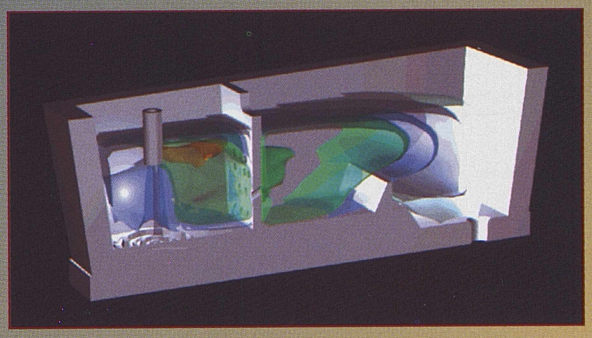
The turbulence suppressor pad, designed (and recently patented) by U.S. Steel through a combination of physical and computer modeling, suppresses this turbulence. "Modeling on the CRAY," notes Vassilicos, "helped to verify physical model observations of how the distribution of turbulence in the pouring region varies. Without the pad, the highest level turbulence is at the surface of the bath, where we don't want it. These computations verified that the design was doing what we expected." Plant testing has shown that using the pad in conjunction with special refractory baffles improves quality of the transition slabs to the point that they meet the criteria of D & I steel. The pad is now used for all steel produced at Gary, resulting in quality improvements for all products. (MS)

Computed and Measured Tracer Effects

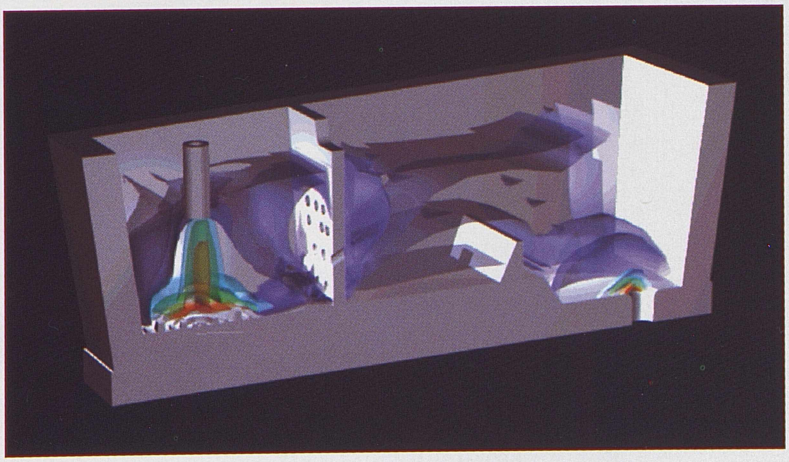
In recent calculations at the Pittsburgh Supercomputing Center, A. K. Sinha and Achilles Vassilicos compared physical measurements of tracer response in scale model and real tundishes to results from computer modeling. Tracers such as a pulse of copper are added to a tundish mix to give a reading of residence time — how long it takes for the tracer to exit the bath — and tracer density over time at the exit. This information gives a valuable index of the flow characteristics of a tundish.

The study shows that commonly used numerical techniques are not sufficiently accurate. The researchers developed FORTRAN code that adapts a more accurate algorithm, known as QUICK (quadratic upstream interpolation), for efficient use on the CRAY.

The fast response of the CRAY as compared to in-house workstations — a day turnaround versus as much as a week — is important to the U. S. Steel researchers. "These computations sometimes require a lot of tweaking and adjustments to parameters," notes Vassilicos, "sometimes with several restarts. With the CRAY, we can see right away what we're getting, and if something needs to be changed, we can do it."



Computed tracer concentration in Gary No. 2 tundish 120 seconds after an impulse of the tracer appeared in the input stream. Colors correspond to tracer concentration, with red representing the highest concentration and blue/gray the lowest.

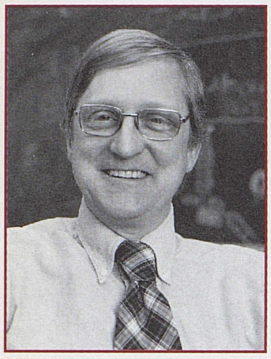


This graphic shows computed results from modeling turbulence in the Gary No. 2 tundish equipped with the turbulence suppressor pad (lower left of graphic). The colored surfaces represent turbulence kinetic energy, with red representing the highest values decreasing to blue/gray, which represents the lowest values. High values of turbulence occur near the pad, and the bath surface is relatively undisturbed.

Reference:
A. Vassilicos & A. K. Sinha, "Tracer Property Effects on Computed and Measured Response in Continuous Casting Tundishes," preprint, U.S. Steel Technical Center, 1992.

Stronger Steels by Design

Simulations of Stress-Corrosion Resistant Ultrahigh-Strength Steels
Gregory B. Olson, Northwestern University



Gregory B. Olson

Stronger than a Locomotive

Superman fans know that the man of steel had amazing strength because he was born on the planet Krypton. Greg Olson isn't superman, but he uses supercomputing, which in this case may be almost as good. With new theoretical insights about the microstructure of steel and powerful computing, Olson is figuring out how to use materials right here on Earth to make steel stronger, harder and tougher than ever before.

Olson directs the Steel Research Group (SRG), a large-scale, multi-institutional research program centered at Northwestern University's Materials Research Center. Over 30 investigators participate, including researchers at Harvard, Brown, MIT and Illinois Institute of Technology, several Department of Defense laboratories and a number of steel companies. "It was conceived in 1985," says Olson, "as a six-year program to develop the fundamental models and database by which you could actually design alloys on the computer."

The work progressed on schedule, and in 1991 Olson's group at Northwestern designed a new steel for bearings in the main engine turbo pumps of the space shuttle. This prototype ultrahigh-strength bearing steel withstands pressure, corrosion and high temperature beyond any previous steel. With this success under its belt, SRG kept going, and it is now applying the same design methods to develop advanced steel for other weight-critical applications such as helicopters, high-performance race cars and naval aircraft landing gear. "Steel is heavy," notes Olson, "but sometimes it's the only thing that can do the job. If you can push the strength up so you use less of it, you can save a lot of weight."

Quantum Blacksmiths

It is fundamental blacksmith's know-how that to harden steel you heat it and then quench it in water. That basic phenomenon underlies SRG's effort to design stronger steels. By developing sophisticated understanding of what happens at the level of atoms and electrons when heated steel is rapidly cooled, Olson and his colleagues are, in effect, taking quenching to its limits, and they have gained an unprecedented ability to control the properties of the resulting steel.

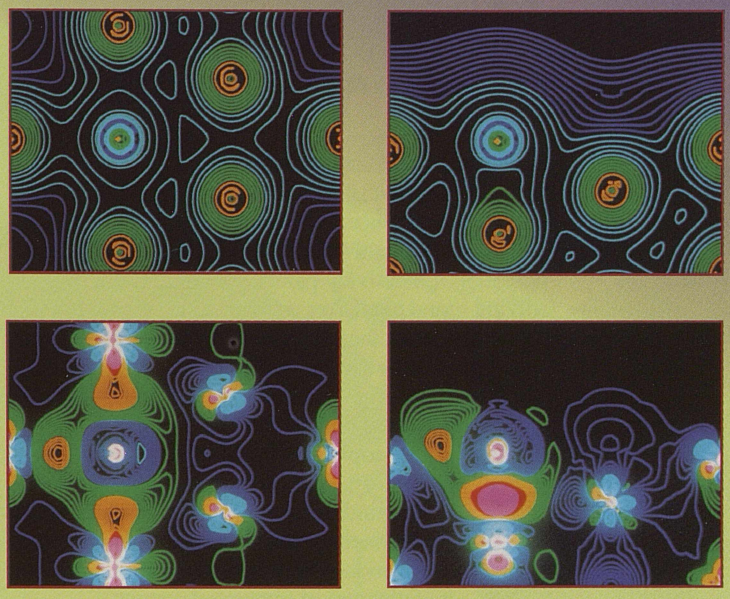
When steel is cooled rapidly, a particular structural transformation occurs, called a "martensitic transformation." "Instead of atoms moving around individually," explains Olson, "they all move in a coordinated mechanical way." The crystal structure spontaneously changes — technically speaking, it shifts from face-centered-cubic to body-centered-cubic. The result is a lower energy state with a fine grain structure that is more cohesive and harder.

By learning to control this structural change — in particular, by using it to create finer spacing between the carbide particles in steel, SRG has made steels 50 percent stronger (meaning 50 percent more resistant to permanent deformation) than conventional steels of the same carbon content. SRG's biggest challenge, however, has been to improve the ability of high-strength steel to resist cracking under the stresses of a hostile environment, like the salt water that affects naval aircraft. To meet this challenge, Olson turned to quantum theory and supercomputing.

Grain Boundary Cracks

Much like the grain in wood, steel has a grain structure, and when steel cracks it tends to do it along the boundaries between grains. At high-strength levels, steel is very sensitive to impurities, and even hydrogen absorbed from the air can cause the grain boundaries to crack. Research has shown that when phosphorus, a common steel impurity, is located at a grain boundary, the steel is more susceptible to this hydrogen-induced cracking. "We know what the bad actors are," says Olson, "but with the supercomputer, we're getting at the electronic mechanism of how they do it."

"Engineers no longer need to be limited by materials that exist, but can start designing materials they would like to have."



Phosphorus and Grain Boundary Cracking

These contour plots from Greg Olson's quantum calculations at the Pittsburgh Supercomputing Center (with Art Freeman and Ruqian Wu) show how phosphorus at a grain boundary affects electron distribution in steel. In the top panels, color corresponds to electron density with yellow representing the highest density decreasing through green, blue and purple. The left panel shows iron atoms and a phosphorus atom at the grain boundary (the boundary runs horizontally through the center of the panel). The right panel shows the same steel fractured along the boundary with the top segment removed, so that phosphorus is at the "free surface" of the crack.

The lower panels are difference plots — showing how the electron distribution changes when phosphorus is in the steel compared to only iron. Electron gain is shown in pink (greatest), red, yellow and green, and electron loss is purple and blue (more negative). These plots show that phosphorus bonds much more strongly with the iron atom below it at the free surface (large pink area) than it does at the grain boundary. "Phosphorous at the free surface has lowered energy," explains Olson, "which means that phosphorus at the grain boundary reduces the work of fracture — promoting embrittlement of the steel."

In collaboration with solid-state physicist Art Freeman of Northwestern, Olson has done quantum theory calculations at Pittsburgh that delve into the roots of the problem — how the electrons associated with iron and phosphorus atoms interact at the grain boundary. Using a highly precise and computationally demanding approach (known as the full-potential linearized augmented plane wave method, FLAPW for short), they have addressed this question of why phosphorus atoms at the grain boundary reduce resistance to cracking. Their results, says Olson, (shown as electronic contour plots) provide "the first definitive electronic-level answer."

As part of this work, Olson and SRG colleagues developed a new method to remove phosphorus from the grain boundaries, and steel they produced this way has significantly higher resistance to hydrogen-induced cracking. With future calculations on the CRAY C90 at Pittsburgh, they will explore how other alloy-elements like boron and molybdenum affect grain-boundary cracking. The objective, says Olson, is to design steel the way semiconductors are designed, with elements deliberately added to get the electronic structure that will provide precisely the properties needed to satisfy the demands of a particular application.

Boundaries of a New Science

"Engineers no longer need to be limited by materials that exist," says Olson, "but can start designing materials they would like to have." SRG exemplifies an evolving discipline Olson calls engineering science, because it brings together — in a total systems approach to design — the opposed philosophies of science and engineering. "Science seeks to distill simple laws from the complexity of nature, and engineering combines those simple laws to create new complexity."

The prototype bearing steel for the space shuttle proved the feasibility of SRG's approach, and industry has become more interested as their work moves closer to other applications. "Probably the most exciting area right now," says Olson, "is gear steels. It looks like we may be able to cut the weight of gears in half." For Michael and Mario Andretti and the Newman-Haas Indy 500 racing team they drive for, this is nothing to sneeze at. Olson began working with Newman-Haas in 1990.

Helicopters and airplanes also require high-strength structures. Navy fighter planes landing on a carrier flight deck stall and essentially drop onto the deck from about 18 feet high, placing a tremendous load on the landing gear and arrestor hook, both prime applications for ultrahigh-strength steel. These planes are about 30% steel by weight — a lot of steel, says Olson, and nothing else will do. (MS)

Reference:

Ruqian Wu, A. J. Freeman & G. B. Olson, "On the electronic basis of the phosphorus intergranular embrittlement of iron," *Journal of Materials Research* **7**, 2403 (1992).

This research is supported by the Office of Naval Research and the National Science Foundation.

The Nitty Gritty of Silicon

Ab Initio Theory of the Si (111)-(7x7) Surface Reconstruction
J. D. Joannopoulos, Massachusetts Institute of Technology

Breakthrough Supercomputing

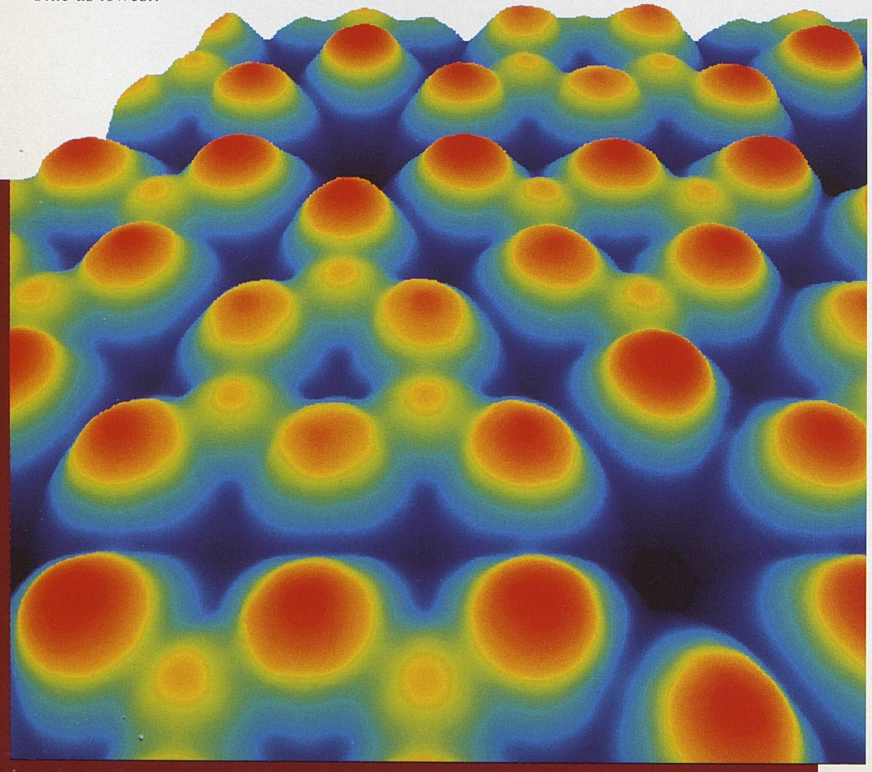
An MIT physicist has done computer simulations that reproduce the full complexity of the surface structure of silicon, showing the great potential of massively parallel computing for solving grand challenge problems in science. John Joannopoulos achieved the milestone with the Connection Machine, CM-2, at the Pittsburgh Supercomputing Center. His 700-atom model closely resembles microscopic images produced in the laboratory.

John Joannopoulos

The finding is significant because silicon is the key component in the semiconductor industry and because the manufacture of integrated circuits depends largely on processes occurring at surfaces. Joannopoulos' work also has broader implications; researchers now can better understand how defects in materials affect their atomic structure and properties.

"Before the Connection Machine," says Joannopoulos, "we were able to calculate the properties of only about 100 atoms in a model for the silicon surface." But in the near future, he says, it will be possible — with more advanced computers such as the CM-5 — to study fractures, which require more than 1,000 atoms to describe their detail. "With the new method, you can make the material behave the way you want. Theorists will be able to perform accurate calculations of materials with true predictive power."

Computer image of the silicon surface reconstruction as calculated by John Joannopoulos. Colors indicate electron density, with red as highest going toward blue as lowest.



A Massively Parallel Shakeup

For the MIT simulation, Joannopoulos and his colleagues — graduate student Karl Brommer, Brond Larson of Thinking Machines, and former graduate student Marc Needels — modeled one repetitive section of the silicon surface, known as the 7 x 7 Takayanagi reconstruction (see sidebar). They developed a five-layer cell and, to create the illusion of a micro-thin silicon wafer, they produced the cell's mirror image, with the two innermost layers considered the wafer's innards. "All atoms were allowed to move freely except those in the middle two layers, each of which was frozen in its bulk position," Joannopoulos says. The newly created super cell contained a slab of 400 silicon atoms — with one half-slab of vacuum on either surface side — for a total volume equivalent to more than 700 atoms.

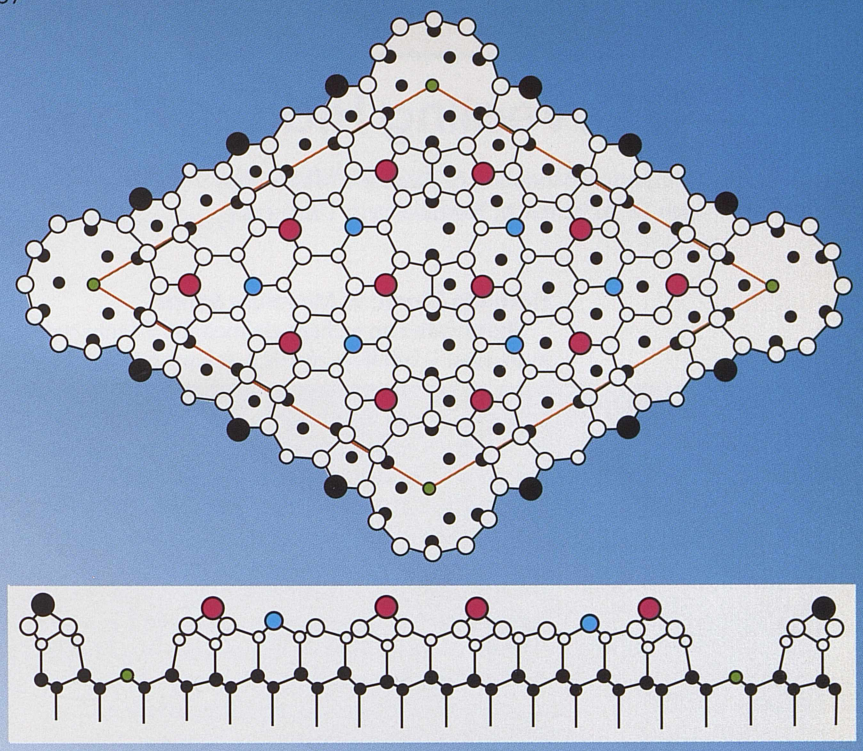
Then, it was a matter of developing the correct program so that the computer would "shake" the atoms from their experimentally estimated locations. The goal was finding the lowest energy state for the entire structure. During the shakeup, the CM-2 computed the forces on each atom and shifted them accordingly. Several hundred hours of computing later — at an impressive speed of 600 million calculations per second — Pittsburgh's 16,000-processor CM-2 recomputed the structure's total energy numerous times, and the process was complete. Part of the computer work was performed at Los Alamos National Laboratory. "Previously, a lack of memory and speed limited us," says Joannopoulos, "but massive parallelism gave us the resources to do the job."

Transatlantic Hookup

Joannopoulos had an unusual way in academic circles to check his results. His former postdoctoral fellow, Michael Payne — now at the University of Cambridge in England — was working on the same problem, using a different program and a different computer. The two agreed to contact each other when they finished their work. "We decided to give each other one day to finish after the phone call and if we couldn't, then the other person would go ahead and submit their paper to the journal," Joannopoulos remembers. "Well, one day he called me, and I said we were finished with our calculations, but we hadn't finished writing our paper yet."

So the two researchers compared their results. "We both laughed in relief because the numbers were within 3 percent. It's unusual for two research groups working on the same problem to come up with virtually the same answer in the

"Within 10 years, we might be able to do all of this routinely."



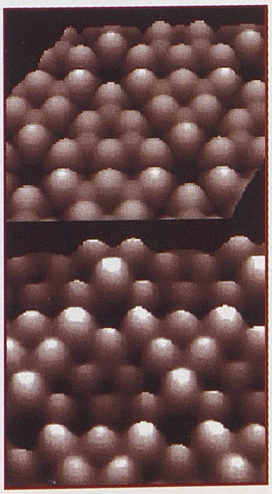
The 7 x 7 Takayanagi Reconstruction

Creating a computer simulation of silicon's surface structure took substantially fewer years than deciphering the surface experimentally. In the mid-1980s, after a quarter-century search, a research team led by Japanese scientist Kunio Takayanagi discovered the correct structure, shown in the schematic in top and side views. It's known as the silicon (111) 7 x 7 reconstructed surface and is perhaps the most complex and widely studied surface of a solid. The surface itself is a repetitive pattern built from a basic unit — the "unit cell" — which Takayanagi's team discovered. The cell includes an array of silicon atoms resembling two equilateral triangles butting heads to form a rhombus. The unit-cell boundary is outlined by an orange line.

After a silicon crystal is cleaved, bonds between the atoms break, and one-time coupled atoms become estranged. Other atoms come to their rescue during the reconstruction and jockey for position, forming many new bonds. These newly bonded atoms adhere to the surface — 12 of them become bumps on the rhomboid surface of the unit cell, represented in the schematic as red circles. Also visible are ditches (green) at each corner of the rhombus, below which the atoms are arranged in layers. The blue circles represent one layer below the 12 outermost atoms, and other atoms at increasing distances from the surface are indicated by circles of decreasing size.

Although the bonds in the Takayanagi surface structure are considerably longer and, thus, less stable than bonds in the crystal's guts, the structure nevertheless represents the lowest energy state of any proposed model for the silicon surface. Because of these longer bonds, the unit cell is 49 times larger than that found in the crystal's interior. More important, the cell contains the smallest number of broken bonds of all proposed surface models — 19 down from a possible 49 — resulting in the most stable arrangement for the silicon surface.

beginning. Either we were both lucky, or we both screwed up," he says. Of course, comparison with the microscopic images validated their findings.The two researchers sent their papers to *Physical Review Letters*, and the journal considered them as simultaneous publications, publishing them in the Mar. 2, 1992 issue.



Computer Simulation versus the Microscope

A computer-generated image (top) of Joannopoulos' silicon-surface model compares well with a silicon-surface image from the scanning tunnelling microscope (below). The position difference is less than 0.01 angstrom.

Joannopoulos says the eventual goal of the research is discovering new materials. "Suppose I'm looking for a material with a particular property and I know other materials behave similarly. For instance, I want a material with AB properties, and I have a material with A property and another with B property. I combine the two materials on the computer and see if the mixture is stable, whether it has the desired properties and also whether I could actually make it. There are lots of ifs, but it's exciting. Within 10 years, we might be able to do all of this routinely." (SE)

References:

Karl D. Brommer, M. Needels, B.E. Larson and J.D. Joannopoulos, "*Ab Initio* Theory of the Si(111)-(7x7) Surface Reconstruction: A Challenge for Massively Parallel Computation," *Phys Rev Lett* **68**, 1355 (1992).

This research is supported by the Office of Naval Research and the Joint Services Electronics Program.

Revealing Hidden Flaws

Numerical Modeling of Acoustic Microscopy
Jan D. Achenbach, Northwestern University

From the Titanic to Miniature Sonar

Shortly after an iceberg claimed the Titanic and more than 1,500 lives in 1912, scientists looked for ways to pinpoint underwater objects. Soon, however, World War I began, and the focus shifted from icebergs to submarines. Gargantuan lights were considered, but visible light didn't penetrate sea water very well. Instead, researchers suggested sending sound signals that bounce off targets — a process known as sonar, which provides the range and bearing of the target because the speed of sound in sea water is known.

Since sonar's discovery, researchers have made great advances in acoustic technology. For medical applications, they have developed ultrasound, which goes one step beyond sonar, converting reflections of sound waves inaudible to human ears into visual images, such as a fetus inside a womb. The procedure is a better option than x-rays for pre-natal care because it distinguishes between soft tissues and does not damage the fetus.

Now, researchers have turned up the frequency still further; this time, to produce sub-surface microscopic images of objects that visible light can't penetrate. The process is known as acoustic microscopy, and instead of ultrasound's frequency of 3.5 million cycles per second, the new breed of microscope cranks out one billion cycles per second. The increased frequency means a shorter wavelength, and as a result, the miniscule pulses of sound can weave their way through all sorts of nooks and crannies, providing the smallest sub-surface details.

"The ideas behind acoustic microscopy are identical to those found in sonar and ultrasound, but the applications and technology are more refined," says Jan Achenbach, a professor of civil and mechanical engineering at Northwestern University, where he founded the Center for Quality Engineering and Failure Prevention eight years ago.

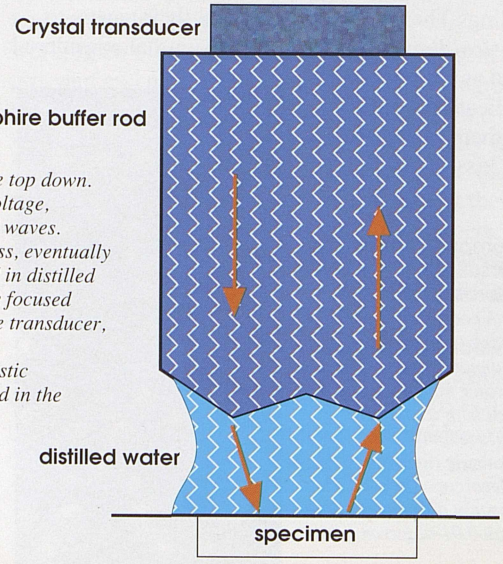
Jan D. Achenbach

Seeing with Sound

Unlike optical microscopy, where we can see the object under examination, acoustic microscopy creates an electrical output from the reflected ultrasound pulse. In order to "see" the specimen, some researchers convert these readings into images by assigning colors and intensities to different numbers. This allows them to produce images unavailable from optical microscopes, such as fibers in composite materials and the random distribution of individual grains in metals.

Other researchers use computer software to develop property-revealing curves from the numbers. But because the electrical signals generated by specimens are usually garbled, it is difficult to decipher their meaning. For that reason, Achenbach, who was elected to the National Academy of Sciences last year, relied on the CRAY Y-MP at the Pittsburgh Supercomputing Center to create numerical simulations of an acoustic microscope examining specimens. The models not only determine material properties, but also detect and characterize defects in the specimens. The goal, Achenbach says, is creating a catalog of these models so he can match their ultrasonic "signals" with the signals obtained from laboratory or industrial specimens.

In industry, thin layers of materials are sometimes deposited on other materials to harden them. For instance, in the tool industry, diamond is placed on steel to harden it. And in the engine industry, thin ceramic layers protect engine components against heat. "When you deposit these layers," Achenbach says, "you may know their thickness, but not their elastic properties."

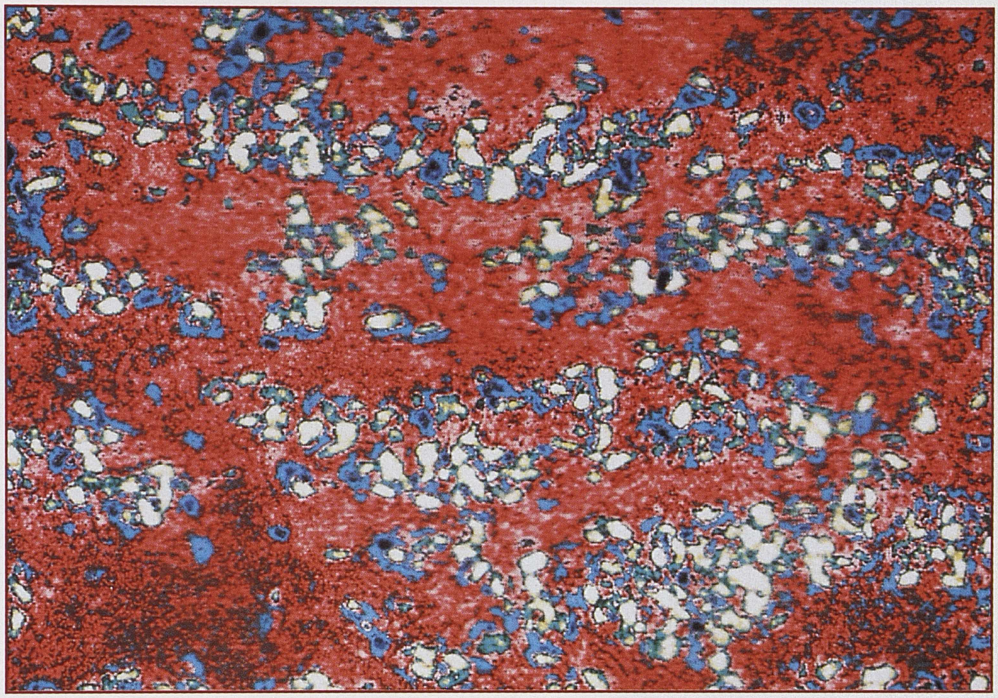


Schematic of acoustic microscope

The acoustic microscope's apparatus is fairly straightforward from the top down. Replacing the observer's eyes is a crystal transducer, which receives a voltage, expands, and pushes against a sapphire buffer rod, producing ultrasound waves. The rod itself resembles a pencil, through which the ultrasound waves pass, eventually focusing themselves in the pencil's point. The point, meanwhile, is bathed in distilled water and is resting just above the specimen. The sound waves, which are focused additionally by the water, strike the specimen and then bounce back to the transducer, registering a particular voltage.

For his numerically based research, Achenbach uses a line-focus acoustic microscope, which is manufactured in Japan and is the only one of its kind in the United States. This variety has several advantages over the usual point-focus configuration because it can make both horizontal and vertical measurements of material properties near the surface, which is essential for analyzing substances that may be layered on each other.

The goal is to create a catalog of models that can be matched with signals from laboratory specimens.



This acoustic microscope image shows aluminum (red) reinforced with ceramic particles (white). This composite material is being developed as a potential replacement for steel in automobile engine components. Aluminum is lighter and dissipates heat better than steel, but requires reinforcement to increase its toughness. The acoustic microscope reveals pores (black) and cracks (blue-green) in the material. An Olympus UH-3 Scanning Acoustic Microscope produced this image, provided courtesy of B. R. Tittmann, Pennsylvania State University.

Currently, industry is using an indentation test to determine that property and others, but it leaves a mark. Also in industry, microscopic surface scratches on materials sometimes develop, and technicians want to measure their depth. "In both cases," Achenbach says, "you can send light to the surface with a light microscope, and it reflects back, never penetrating the surface." Other times, friction and wear generate subsurface cracks that may lead to material fatigue, and acoustic microscopy provides the best potential analysis method.

Supercomputer Modeling

Using a supercomputer to model this complex process requires a series of mathematical equations, none of which can be solved completely. "The system is too difficult to solve analytically, and you cannot write out the solutions," Achenbach says. "So you must break up the system into components because smaller parts have simpler equations and geometry, and then you solve equations for each part, including their interfaces." The processes are known as boundary element analysis and finite element analysis, and they help determine the system's overall properties. A supercomputer is needed here because the researchers transform a one-millimeter specimen into a two-dimensional 2,000-by-2,000 computer grid, where a series of equations is computed at each point.

Achenbach and his colleagues — John Harris, V.S. Ahn and Jin O. Kim — have simulated measurements when either titanium nitride or niobium nitride are deposited on magnesium oxide. And so far, their results have been favorable because when they compare defective or defect-free specimens simulated on the computer with those tested in the laboratory, the findings are identical. The researchers have obtained reliable results from layers as thin as a few *microns* — each micron is one-millionth of a meter. "As industry improves their processing techniques," Achenbach says, "they will use our methods to test and develop new products." (SE)

References:

G.R. Liu, J.D. Achenbach, J.O. Kim and Z.L. Li, "A Combined Finite Element Method/Boundary Element Technique for V(z) Curves of Anisotropic-Layer/Substrate Configurations," *Journal of the Acoustical Society of America* **92**, 2734 (1992).

V.S. Ahn and J.D. Achenbach, "Response of Line Focus Acoustic Microscope to Specimen Containing a Subsurface Crack," *Ultrasonics* **29**, 482 (1991).

This research is funded by the Office of Naval Research and Cray Research Inc.

The Structure of Steps

Quantum Molecular Dynamics Simulations of Semiconductor Growth Mechanisms

Jerzy Bernholc, North Carolina State University

"Ah Sun-flower!
weary of time,
Who countest the
steps of the Sun."

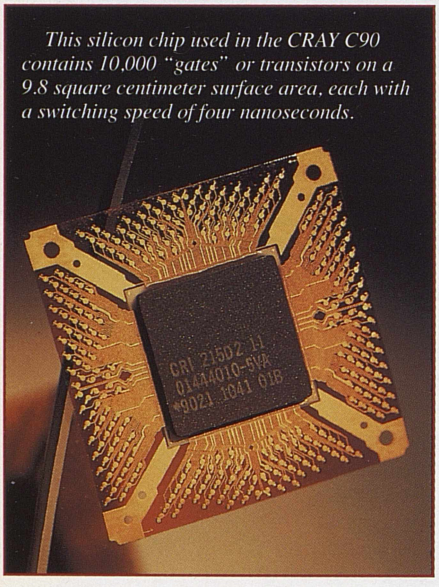
—Blake

Symbiotic Supercomputing

Like honeybees and flowers, like a hand fits a glove — supercomputing and semiconductor research go together well, and the relationship is thriving in solid-state physicist Jerzy Bernholc's work at the Pittsburgh Supercomputing Center.

One of the main forces driving development of more powerful supercomputers like the CRAY C90 has been the ability to make transistors, the basic microelectronic device of computer circuitry, smaller and faster. A transistor is essentially a tiny electronic switch, and the first one — invented at Bell Labs in 1948 — took a few microseconds (millionths of a second) to switch between on and off states of current flow. In today's supercomputers, switching times are a thousand times faster and measured in nanoseconds (billionths of a second), and new "chip" materials, such as gallium arsenide and diamond, hold promise of picosecond (trillionth of a second) switching.

Jerzy Bernholc



This silicon chip used in the CRAY C90 contains 10,000 "gates" or transistors on a 9.8 square centimeter surface area, each with a switching speed of four nanoseconds.

As faster electronics increases supercomputing power, it becomes possible for researchers like Bernholc to ask more detailed and sophisticated questions about the atomic structure and electronic properties of the semiconducting materials used to create these minuscule instantaneous switches — research that leads the way in a kind of mutual positive feedback loop toward new and even more powerful computers.

Growing Crystals by Molecular Beam Epitaxy

Semiconductor material used to manufacture chips must have nearly perfect orderly arrangement of the atoms in the crystal structure, and naturally occurring materials have too many impurities and defects to be used directly. Producing crystalline material under controlled conditions is known in the trade as "growing" the crystals, because the process usually involves starting with a small "seed" crystal of the desired structure and adding to it atom-by-atom. Several different techniques are used to grow relatively pure, defect-free crystals. In the case of silicon, the most important and best understood semiconductor, chip-quality crystals are usually grown by carefully pulling a seed crystal from molten silicon.

Bernholc has used the newest and most advanced system from Cray Research, the C90, to do detailed quantum calculations on the "step-flow mechanism," believed to be the dominant growth process involved in molecular beam epitaxy (MBE), a relatively new method of growing chip-quality crystals. As the trend continues toward smaller (and thus faster) devices, it is increasingly important to gain microscopically precise control of the crystal structure, and in this respect, MBE is superior to melt-based techniques. Epitaxy means that one layer is deposited on another, so that the atomic surface of the bottom layer, called the substrate, serves as a template controlling the arrangement of atoms in the top layer. Somewhat like eggs being placed in an egg carton, deposited atoms in the top layer register in the inter-atomic spaces of the substrate. In MBE, a gaseous beam of semiconductor material deposits a layer only one or two atoms in thickness.

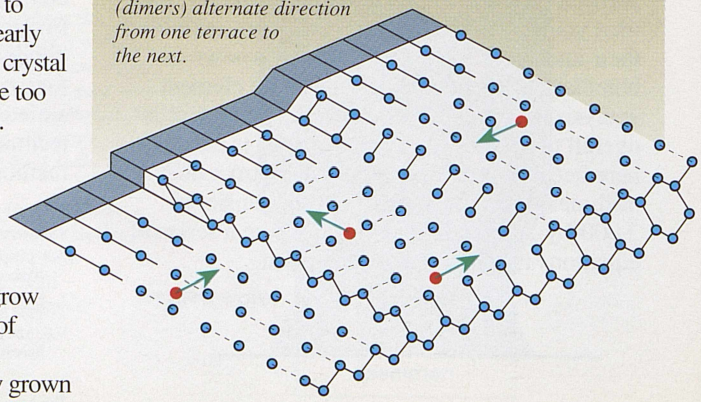
Step-by-Step

Bernholc's research focuses on silicon homoepitaxy, the process of depositing silicon on a silicon substrate. A good deal of research has focused on this process, because of its importance, and it is known that even the highest quality surfaces grown this way contain "steps" — one or two atom layers of stairstep-like formations on the edge of a growing crystal.

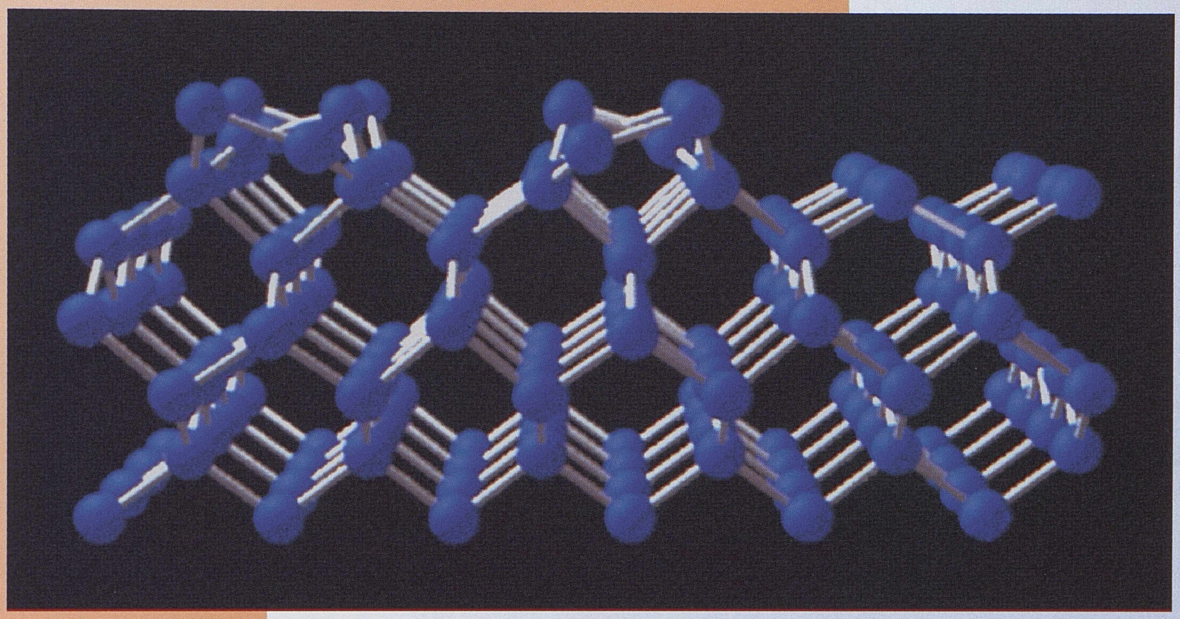
"In recent years," says Bernholc, "it has become apparent that steps play a crucial role in controlling the growth process on the silicon surface. At standard temperatures, atoms that have condensed on the flat regions of the surface (terraces) diffuse to the step edges, where they are more readily incorporated into the crystal. This step-flow mechanism results in the desired layer-by-layer epitaxial growth."

Steps and the Step-Flow Mechanism

This schematic represents three terraces with single-atomic-layer steps between them. Deposited atoms (red) arrive at a surface and travel to the step edges, where they fill out the crystal by extending the terrace. This is believed to be the prevailing growth mechanism for semiconductors. The bonds (dotted lines) between paired surface atoms (dimers) alternate direction from one terrace to the next.



As faster electronics increases supercomputing power, it becomes possible to ask more sophisticated questions about the semiconducting materials used to create these minuscule instantaneous switches.



The atomic structure of steps, as determined in calculations by Jerry Bernholc on the CRAY C90. The silicon crystal is constructed of hexagonal rings, but geometry at the step edges (top of graphic) is more complicated. Joel Welling of the Pittsburgh Supercomputing Center scientific visualization group produced this graphic from Bernholc' data.

To gain precise understanding of this process, explains Bernholc, it is necessary to track the pathway by which deposited atoms diffuse to the step edges and to know the energy required for them to form bonds with the crystal. Scanning tunneling microscope images show the approximate structure of steps, but the precise positions of the atoms is still unknown.

First Principles Structure of Steps

Recent improvements in numerical techniques, in particular a method called quantum molecular dynamics (QMD), coupled with a more powerful supercomputer made it feasible for Bernholc to undertake a series of calculations aimed at providing this information. "Memory constraints prevented us from running these calculations on the Y-MP," notes Bernholc, "and with our re-optimized code the C90 is 2.9 times faster than the Y-MP [Cray's predecessor system]."

Taking advantage of C90 "friendly user" testing in the early months of 1993, Bernholc let his QMD code go to work — at the rate of 693 million calculations a second on a single processor, with most calculations running on four processors (2.3 billion calculations a second). This is 70 percent of the theoretical peak maximum performance available on the C90, very fast indeed and currently the fastest single processor performance of any code running at the Pittsburgh Supercomputing Center. The results provide for the first time *ab initio* calculations — i.e., atomic interactions calculated directly from the electron interactions — of the atomic and electronic structure of steps.

In future work, Bernholc will do calculations to precisely identify the diffusion pathway and binding energies of deposited atoms. With results from these QMD studies, he will then be in a position to do simulations using a simpler technique (Monte Carlo) that map the entire growth sequence of steps. (MS)

Acknowledgement:

P. Boguslawski, Z. Zhang, Q.-M. Zhang & J. Bernholc, manuscript in preparation.

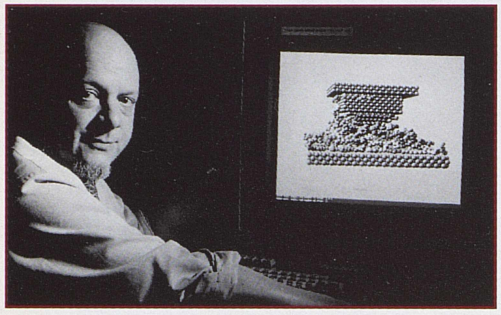
Bernholc's research is supported by the Office of Naval Research and the National Science Foundation.

The Long and Short of it

Fundamental Studies in Microtribology, Microrheology, Finite Aggregates and Reaction Dynamics
Uzi Landman, Georgia Institute of Technology

Lights, Camera, Action!

Uzi Landman is not your typical movie producer. Take, for instance, the next video project of Landman and senior research scientist Charles Cleveland. The opening scene will show short- and long-chained carbon molecules in a mixture, their green and blue bodies intertwined above a metallic surface. Then, as the video runs, the longer molecules will jockey for position, attempting to hook up with the surface, one atom at a time, knocking the shorter actors out of the way. A short time later, the longer molecules have won the battle, two to three layers deep, all lying in parallel above the solid surface.



Uzi Landman

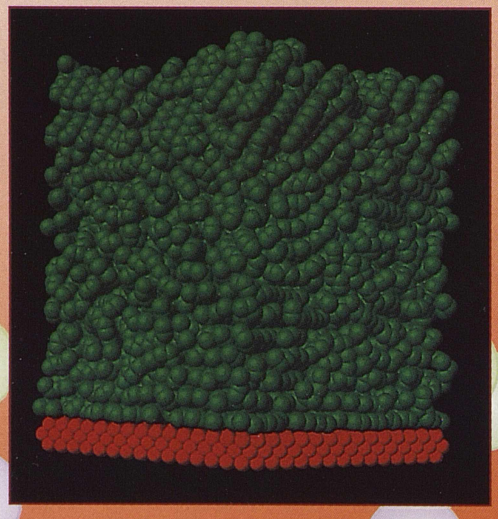
The set of Landman's latest video was the CRAY C90 at the Pittsburgh Supercomputing Center, and the entire simulation for this production took 500 hours. Landman was among a handful of Pittsburgh users granted access to the new machine after its installation last October. Their goals: to carry out leading-edge scientific research while working the bugs out of the new machine and bringing it up to peak performance.

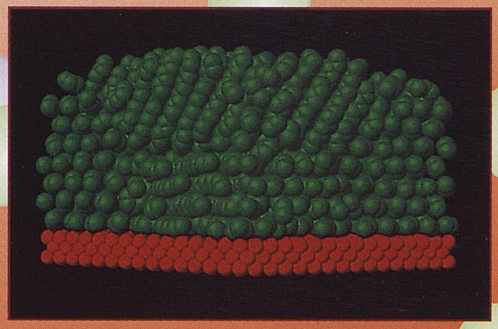
"In addition to providing an excellent medium for presenting simulation results, animation is an analytical tool," says Landman, a professor of physics at the Georgia Institute of Technology in Atlanta, "because you can animate not only how atoms move in space from A to B, but you can also animate the dynamical evolution of their properties. For instance, you can color them according to their instantaneous velocities. If an atom gets hit and starts moving fast, it turns from blue to red."

Supercomputing with Complex Liquids

A portion of the research at Landman's Center for Computational Materials Science focuses on systems of long-chain molecules, known as complex liquids, that have basic science implications, as well as biomedical and engineering applications. Of particular interest is how these molecular chains behave near surfaces. When a surgeon inserts a metallic plate into the body, for instance, the plate's surface interacts with tissue and may disturb the natural surroundings. Another example involves lubricants and how they act when sandwiched between two materials. Knowing how the long-chain molecules stretch, coil around each other and move is the goal. "We are trying to provide sometime in the future the recipe book that will give you details on the molecular level — what it is you have to change about the molecular structure," Landman says, "to make it desirable for the technological or biological applications you wish to achieve."

Simulations on supercomputers, Landman says, are playing a key role in devising these recipes for several reasons. First of all, the complexity of these systems requires that thousands of atoms be included in the simulation. To obtain meaningful results, furthermore, researchers need millions of time intervals to simulate the system's behavior, even for a short time span stretching only billionths of a second. The simulation described above of segregation between short- and long-chained carbons, for instance, included 2.4 million intervals during eight billionths of a second.

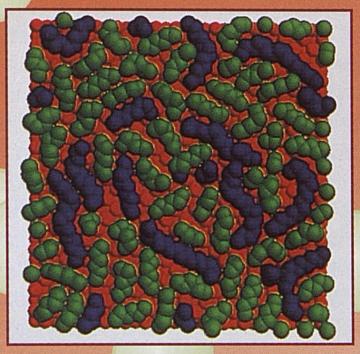


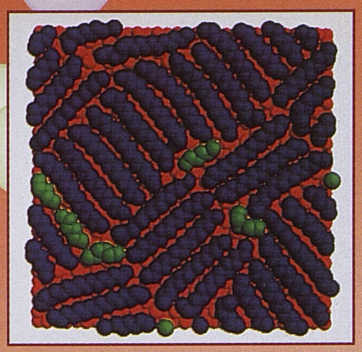


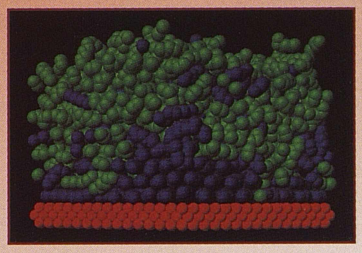
Crystallization

At left, a representation of n-heptadecane film (n-$C_{17}H_{36}$) at 297.15° K, two degrees above its freezing point, shows molecular ordering only in the crystallized top monolayer (n-heptadecane molecules are green; gold atoms are red). At right, a film of slightly smaller long-chain molecules, n-hexadecane (n-$C_{16}H_{34}$), is shown at the freezing point, when total crystallization has occurred.

"This is the type of thing I suspect very much would not have been predicted without computer simulation."

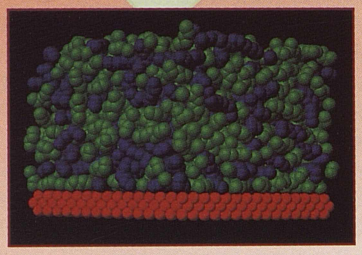






Segregation

Top and side views of a mixture of short- (green) and long-chain (blue) carbon molecules (n-C_6H_{14} & n-$C_{16}H_{34}$) before segregation (left) and toward the end of segregation (right). The top view shows the layer of carbon molecules in contact with the gold (red) surface.



Like a Zipper

For the segregation study soon to be a video, Landman and research scientist T. K. Xia prepared a 50/50 mixture of six- and 16-carbon molecular chains. The objective was to understand how long-chain molecules migrate to the surface and stay there. Experiments had suggested that long-chains seek out the surface. The simulations revealed that in doing so, the long molecules lie parallel to the surface in order to achieve the most stable configuration.

The simulation shows that as the long chains reach the surface, each molecule's head seeks out a place to anchor, and in the process, they push away and displace short chains. Then, in rapid succession, each atom finds an anchor position. "It's like a zipper," Landman says. "The first one goes in, and in order to have space for the next one, it must push something away and who does it push? The short chain."

After the first layer forms, other layers form on top, similar to the growth of skin. In the process, the short-chain molecules are continuously displaced into the liquid, away from the surface.

This finding may have significance for biomedical applications, as well as lubrication. For instance, if the surface in this case had been a steel plate in a human body, the tissue around it might fail to function properly. "Now that we understand this anchoring mechanism," says Landman, "maybe we can diminish the surface effect by modifying the adhesive properties of the surface."

Teeth of a Comb

In another recent project, Landman and Xia looked at crystallization in complex liquids, specifically how the surface of a film made of long-chain carbons freezes. Normally, a liquid freezes in a concerted reaction at a specific temperature — the freezing point — and that's what Landman and others had expected to see. The simulations showed, however, that the top layers of films made of complex molecules containing 16 and 17 carbon atoms crystallize two to four degrees above the freezing point. "The surface molecules stood up like the teeth of a comb," Landman says, "almost perpendicular to the surface, with a small tilt angle." Then, at the freezing point, all the layers in the film crystallized.

Landman and his colleagues, Xia and David Luedtke, wanted to know whether shorter chains would behave similarly, so they simulated freezing in a film made of six-carbon chain molecules. Crystallization of the top layers did not occur. Landman believes that the break-off point probably is chains longer than 10 carbons. Freezing of the film's top layer, with long chains lining up like pencils, helps to stabilize the more complex system by lowering its free energy.

Shortly after completing these simulations, Landman learned that coincidentally and independently researchers had produced similar results in the laboratory. With theoretical predictions and experimental findings corroborating and complementing each other, these investigations show a new phase between liquid and solid. "This is the type of thing," Landman says, "which I suspect very much would not have been predicted without computer simulation." (SE)

References:

T.K. Xia and Uzi Landman, "Molecular Dynamics of Adsorption and Segregation from an Alkane Mixture," *Science* 1993 (in press).

T.K. Xia and Uzi Landman, "Structure and Dynamics of Surface Crystallization of Liquid n-Alkanes," *Physical Review Letters* (in press).

This research is supported by the U.S. Department of Energy, the Air Force Office of Scientific Research and the National Science Foundation.

Faster Electronic Structure

New Methods for Electronic Structure Calculations on Large Molecules

Richard A. Friesner, Columbia University

Cracking Chlorophyll's Code

Photosynthesis is the energy source for life as we know it. It occurs in all green plants and produces the raw materials that fuel metabolism in virtually all organisms. Nevertheless, scientists don't completely understand this fundamental process, in which a plant's energy workhorse — known as a chloroplast — uses sunlight to combine water and carbon dioxide to form sugar and similar foods like starch. As a byproduct, oxygen is released into the atmosphere.

For the last 15 years, chemist Richard Friesner of Columbia University has been interested in the region where the energy conversion mostly occurs — the photosynthetic reaction center, a complex of four large proteins containing six chlorophyll (green pigment) molecules. The center is in essence a large photocell. As light strikes it, electrons in the chlorophyll become excited, and the chemical process begins. To better understand the reactions, it's imperative to first determine how electrons are distributed in the chlorophyll. This poses a challenge because chlorophyll is a large molecule, containing about 70 atoms, and researchers typically have had success calculating electronic structure only with molecules up to 20 atoms — because computing demands rapidly escalate with increasing numbers of atoms.

Richard Friesner, explaining his computational methods at a Pittsburgh Supercomputing Center seminar in February 1993.

To address this problem, Friesner has refined traditional electronic structure techniques by taking some clues from how physicists model turbulence. With the help of the CRAY Y-MP (and now C90) at the Pittsburgh Supercomputing Center, he has developed a new approach to these calculations, called the pseudospectral method.

Using this method, he has modeled a 38-atom molecule known as porphine, which resembles chlorophyll, and the significantly improved efficiency of his pseudospectral approach compared to other electronic structure techniques suggests that it could allow researchers to address other large molecules. "Enhanced computer capabilities," says Friesner, "now have made molecular modeling a serious possibility for virtually any system one cares to study."

Computing with Quantum Theory

To determine the distribution of 10 electrons around an atom in three-dimensional space, a fully exact approach to quantum theory requires calculating each electron's instantaneous interactions with every other electron in the system — in other words, 10 calculations for each electron, 100 in all for the molecule. These calculations are extremely demanding and stress the capabilities of even the most advanced supercomputers, and scientists usually opt for an approach known as the Hartree-Fock approximation, which computes each electron's distribution as a result of interaction with an averaged field representing all the other electrons in the molecule.

About 10 to 15 years ago, scientists began improving on the Hartree-Fock approximation. Instead of having electrons see an averaged field, researchers developed correlation methods, which more accurately describe the interaction of electrons and the molecule's energy. But the demands of these calculations are sometimes prohibitive, even with supercomputers.

At the Pittsburgh Supercomputing Center, Friesner has developed his pseudospectral approach to speed computations on both the Hartree-Fock approximation and one electron correlation method, known as the Generalized Valence Bond Theory — developed by his collaborator Bill Goddard at the California Institute of Technology

The efficiency in Friesner's pseudospectral approach comes from reducing the escalation in computing time that usually goes along with increasing the number of "basis functions" involved in the calculation. Quantum theory represents the electron distribution as a wave function, which is written as a sum of convenient mathematical terms known as basis functions. There are many varieties of basis functions, and researchers decide which ones are appropriate to solve their particular problem.

To solve Hartree-Fock equations, the difficulty rises substantially as the number of basis functions increases. The average number typically ranges between 100 and 1,000. Doubling the number of basis functions means a 16-fold growth in the amount of computing required. The pseudospectral approach halves this exponential escalating effect, and as a result, it's possible to tackle much larger molecules.

"Enhanced computer capabilities now have made molecular modeling a serious possibility for virtually any system one cares to study."

The Pseudospectral Speedup

Friesner examined the basis functions in the traditional wave functions and decided that while it works relatively well to perform calculations involving the same atom, there was a better way to study molecules formed from different atoms. He throws the functions on a grid, similar to what physicists do with calculating turbulence.

To make the grid approach work, Friesner represents the atoms as spherical shells of grid points. Along with these atomic grids, there are specially designed sets of fitting functions that make the connection between the grid representation and the basis function representation. By optimizing the grid point positions and the fitting functions, Friesner is able to use about 100 grid points to represent each atom, while retaining a high level of accuracy.

Friesner has compared his pseudospectral calculations to GAUSSIAN 92, one of the most frequently used programs for electronic structure, and has achieved speeds four to 10 times faster, depending on the desired accuracy, which is influenced by the complexity of the basis functions.

"Our first step was to make sure the program worked properly," Friesner says, and Pittsburgh's CRAY C90 helped. Over a three-month period since the C90 was installed, Friesner put his code through its paces on a large number of test cases involving hundreds of runs and 200 to 300 hours of C90 computing. The new system eliminated concern, says Friesner, about memory and disk restrictions, and it greatly improved turnaround. "The C90 saved us a substantial amount of time in this development process," says Friesner, "probably a factor of two or three in real time."

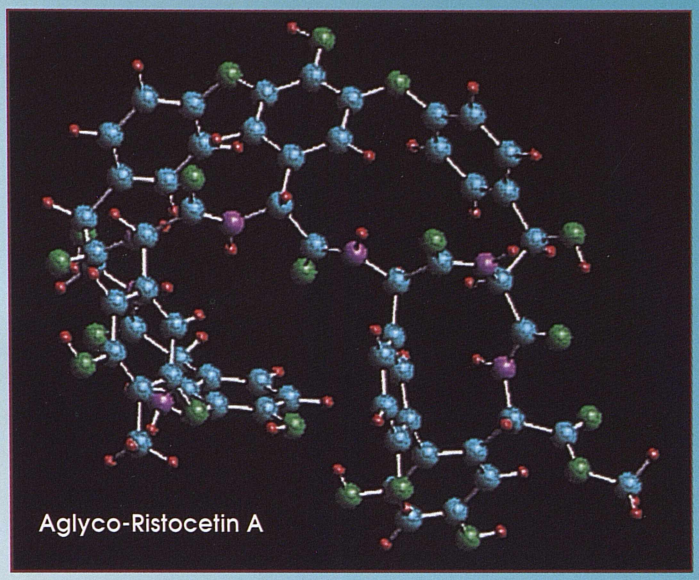
Now that the pseudospectral approach has proved effective, Friesner will begin applying it to new problems and expects to get answers to questions that haven't been addressed before. He foresees widescale study of large molecular structures. "Over the next five to 10 years, reasonably accurate methods will become applicable to these structures on a routine basis. Furthermore, the development of user-friendly, commercialized software insures that this technology will be accessible to chemists in all fields and institutions." (SE)

References:

Richard A. Friesner, "New Methods for Electronic Structure Calculations on Large Molecules," *Annu Rev Phys Chem* **42**, 341 (1991).

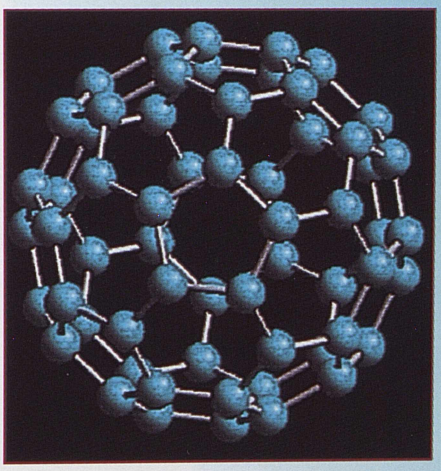
This research is funded by the National Science Foundation, National Institutes of Health and the Department of Energy.

The graphics were produced by Robert Donnelly of Schrödinger, Inc., Pasadena, Calif., from calculations by Friesner, Bill Goddard of Caltech and Schrödinger, Inc.



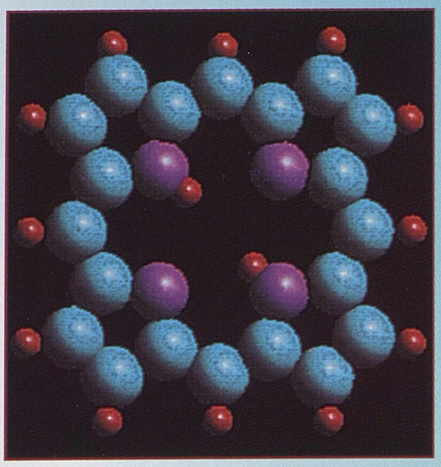
To test his pseudospectral technique, Richard Friesner did calculations on three large molecules for which the electronic structure was already known. These representations of molecular structure show carbon atoms (blue), hydrogen (red), nitrogen (purple) and oxygen (green).

The largest molecule, Aglyco-Ristocetin A, is an antibiotic with 138 atoms. For this molecule, the pseudospectral approach was four times quicker than GAUSSIAN 92.



Carbon 60 "Buckyball"

The spherical "buckyball" molecule (C_{60}) tested whether the pseudospectral approach could determine electronic structure when a molecule is symmetric. In this case, symmetry reduced computation time by 20 to 30 times.

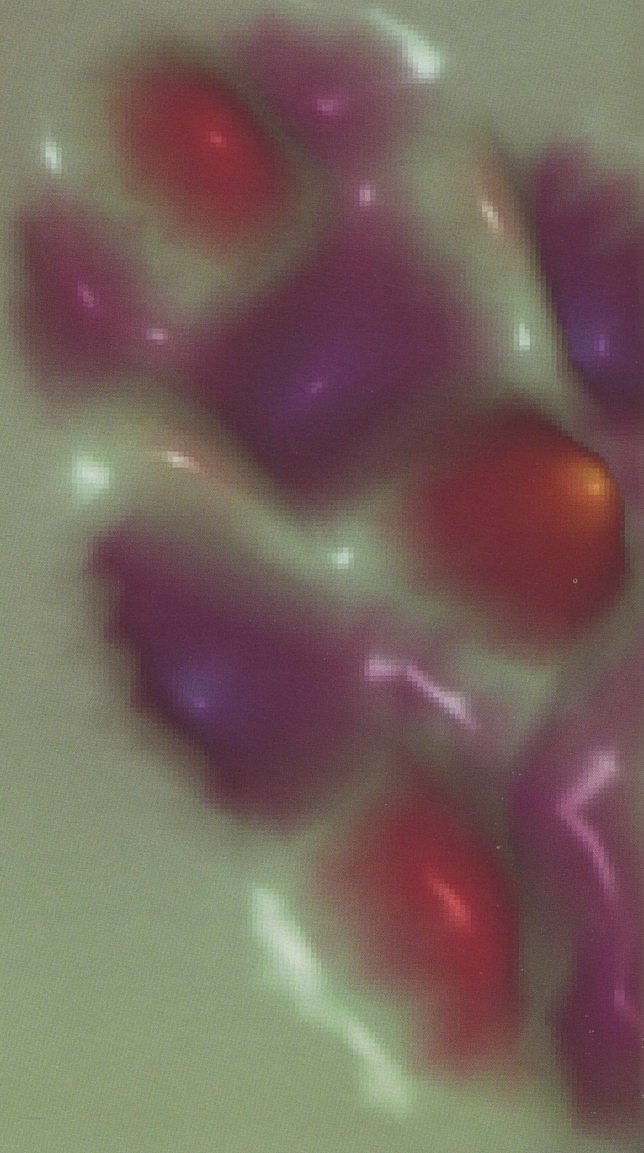


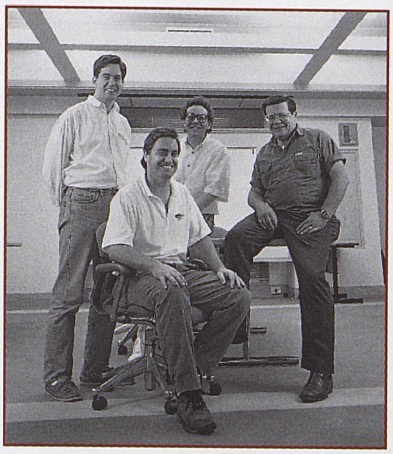
Porphine

Porphine is a 38-atom molecule that resembles chlorophyll, a key component of the photosynthetic reaction center in plants. In the future, Friesner plans to determine electronic structure of the 70-atom chlorophyll molecule because it would help researchers answer questions about photosynthesis.

The Academic Affiliates of the Pittsburgh Supercomputing Center

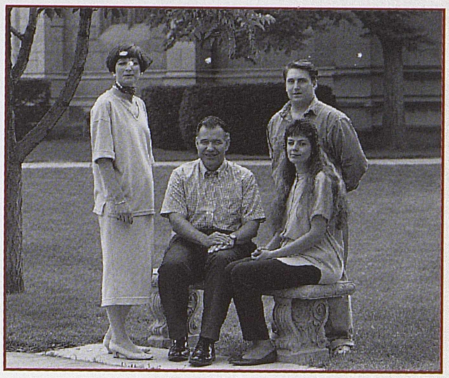
Carnegie Mellon University
Case Western Reserve University
Drexel University
Duke University
University of Florida
Georgia Institute of Technology
Johns Hopkins University
Lehigh University
University of Maryland
Massachusetts Institute of Technology
University of Michigan
Michigan State University
North Carolina State University
University of North Carolina
Northwestern University
Ohio State University
University of Pennsylvania
Pennsylvania State University
University of Pittsburgh
Rensselaer Polytechnic Institute
University of Rochester
Temple University
University of Tennessee
University of Virginia
Washington University
West Virginia University
University of Wisconsin
Yale University





Rich Raymond (right), user consultant coordinator, with user consultants Andy Adams (seated), Bryan Webb (left) and Rick Costa.

Elvira Prologo (left), D. Cook and Tracy Mills of the administrative staff.



Debbie Azzinaro (left), Ed Wozniak, Shirl Grant and Mark Porterfield of the support systems group, which also includes (not pictured) Chuck Maiden.

The Pittsburgh Supercomputing Center is a joint project of Carnegie Mellon University and the University of Pittsburgh together with Westinghouse Electric Corp. It is pleased to acknowledge significant support from:

The National Science Foundation
The Commonwealth of Pennsylvania
The National Institutes of Health
The Advanced Research Projects Agency
The Ben Franklin Technology Center of Western Pennsylvania
Buhl Foundation
Cray Research, Inc.
Digital Equipment Corp.
International Business Machines Corp.
The W. M. Keck Foundation
Network Systems Corp.
Storage Technology Corp.
Sun Microsystems, Inc.
Thinking Machines Corp.
Bell of Pennsylvania

Credits

Editor: *Michael Schneider, Pittsburgh Supercomputing Center.* **Design & Illustrations:** *Steve Palmer, Pittsburgh.* **Additional Illustrations:** *Fred Carlson, Pittsburgh (16, 18, 20, 24, 34).* **Writing:** *Steve Eisenberg & Michael Schneider, Pittsburgh Supercomputing Center.* **Photography:** *Jim Schafer (1, 5), Jeff Macklin, Photography and Graphic Services at Mellon Institute (5-9, 13, inside back cover), Michael Haritan (12).* **Printing:** *Westinghouse Printing.* **Cover Masthead Design:** *Jamie Christner, Pittsburgh Supercomputing Center.*

EdPress Award Winner 1993

Page layout & design were done by Steve Palmer on a Macintosh Quadra 700 at the Pittsburgh Supercomputing Center using PageMaker 4.2 desktop publishing software and Adobe Photoshop. These digital layouts included spot color, encapsulated postscript graphics, and tag image file format (TIFF) scientific graphics created from calculations by researchers at the Pittsburgh Supercomputing Center. Special thanks to Joel Welling, Anjana Kar and other members of the Pittsburgh Supercomputing Center graphics group for their assistance.

Printed on recycled paper with vegetable-based inks. Please pass to a friend.

Pittsburgh Supercomputing Center
Mellon Institute Building
4400 Fifth Avenue
Pittsburgh, PA 15213

Nonprofit Org.
U.S. Postage
PAID
Pittsburgh, PA
Permit No. 251