Managed by UT-Battelle for the U.S. Department of Energy

Oak Ridge National Laboratory

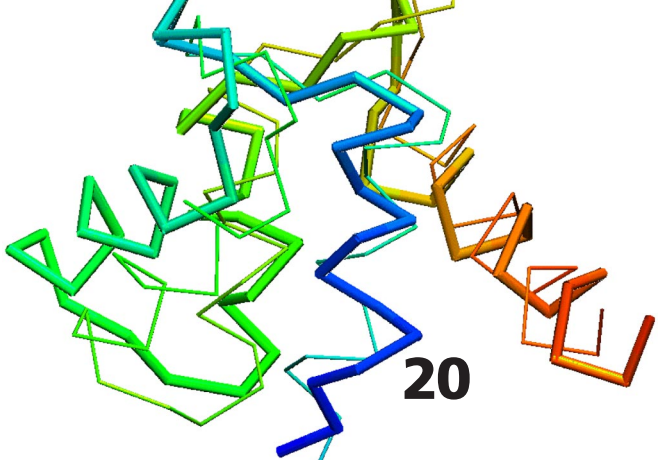
REVIEW

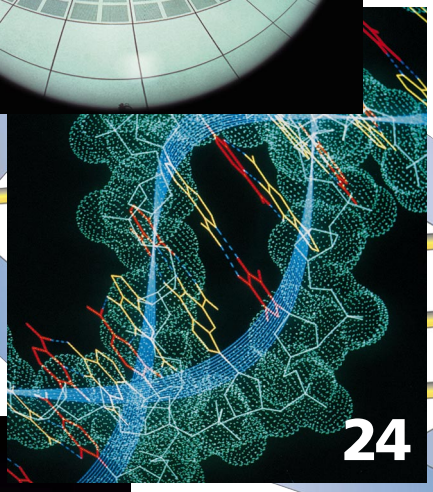
Vol. 34, No. 1, 2001

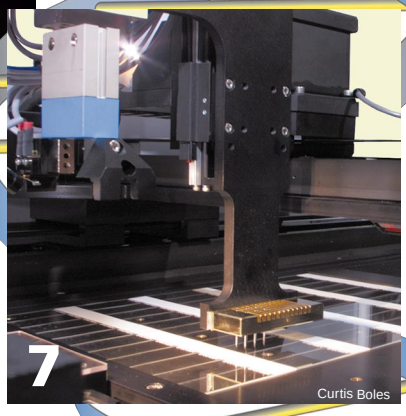
New Biology: Covering All The Bases

REVIEW Vol. 34, No. 1, 2001

Features







COVER: *Biologist Madhu Dhar uses the ORNL-developed MicroCAT to image fat deposits in a mouse. The MicroCAT and microarray (shown on the cover) are among the many technologies used at ORNL (and described in this issue) to understand what genes and proteins do to make organisms function well or poorly.* Dhar photo by Curtis Boles; cover collage by Jane Parrott.

REVIEW

Editor and writer—Carolyn Krause
Assistant editor—Deborah Barnes
Editorial Board—Lee Riedinger (chair), Fred Bertrand, Eli Greenbaum, Russ Knapp, Reinhold Mann, Stan Milora, Thomas Zacharia
Designer and illustrator—Jane Parrott
Photographer—Curtis Boles
Photographic enhancement—Gail Sweeden

The Oak Ridge National Laboratory *Review* is published three times a year and distributed to employees and others associated with or interested in ORNL.

Editorial office address:
Building 4500-North, M.S. 6266,
Oak Ridge, TN 37831-2008
Telephone: (865) 574-7183;
FAX: (865) 241-6776;
Electronic mail: krausech@ornl.gov
ORNL *Review* Web address:
www.ornl.gov/ORNLReview/

The *Review* is printed in the United States of America and is also available from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161

Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the Department of Energy under contract DE-AC05-00OR22725

ISSN 0048-1262

Oak Ridge National Laboratory is a multiprogram, multipurpose laboratory that conducts research in energy production and end-use technologies; biological and environmental science and technology; advanced materials synthesis, processing, and characterization; and the physical sciences, including neutron-based science and technology.

Frank Harris

Unraveling Complex Biological Systems

A new age of biology is being ushered in with the sequencing of the human and other genomes. Armed with this information about the order of DNA bases in genes and the chromosomes on which they sit, scientists can now initiate the steps required to understand processes that occur in the human body, down to the level of the molecule. This new information promises a rich harvest: drugs targeted to our own personal DNA, early warnings of diseases to which we are genetically predisposed, and a more profound understanding of human evolution.

Now that we know where different genes are located on the 23 pairs of human chromosomes, scientists have begun to focus on what these genes do. Which genes are silent in each organ? Which genes are turned on, or expressed, directing cells to produce proteins in specific shapes that determine their functions? What are the exact structures of specific proteins that cause the body to become ill or work well? Can drugs be made to dock with these proteins, like two jigsaw-puzzle pieces fitted together, to block or enhance their action in the body, improving overall health?

In this postgenomic era of biology, scientists are learning about gene expression and the nature of the expressed proteins—understanding life's processes at the molecular level. Such knowledge could lead to therapeutic drugs targeted to each individual's genome, ensuring their effectiveness.

Researchers at ORNL—where the function of messenger RNA, the chromosomal basis for sex determination in mammals, DNA repair processes, and several important mouse genes have been discovered—are taking an interdisciplinary approach to unraveling complex biological systems. This approach also integrates studies of mouse mutations with studies of genes and proteins, using various technologies—automated DNA sequencing, biomedical imaging, microarrays (gene chips), mass spectrometry, neutron sources, and terascale supercomputers.

Barry Berven and Dabney Johnson inspect the hoods in the new biology laboratory at ORNL.

Curtis Boles

As described in this issue of the *Review*, we are determining which genes are expressed in microbes, fish, and mice during exposures to environmental toxins. We are searching for unique protein signatures of various microbes. We have already computationally analyzed the human and mouse genomes, to predict the structure of genes and proteins and make educated guesses about protein function. We have one of the world's top-ranking groups in the area of computational prediction of protein structure. We have identified mouse genes that play a strong role in genetic diseases similar to maladies that afflict humans, such as cancer, obesity, and epilepsy. We are collaborating with researchers in the Tennessee Mouse Genome Consortium to determine which of ORNL's 60,000 mutant mice are excellent models of human genetic diseases that can be used to test the effectiveness of various therapies.

Our experimental research is focused on a variety of organisms: microbes, zebrafish, hybrid poplar trees, and mice. Our computational research is focused on microbial, mouse, and human genes and proteins. We are now taking advantage of ORNL's leading mass spectrometry capabilities to study proteins expressed by these organisms under varying conditions. The success of much of this work has depended—and will continue to depend—on the strength of the collaborations involving scientists at ORNL and elsewhere who represent many different disciplines.

ORNL's goal is to be a center of excellence and a resource for our understanding of (1) complex biological systems, from the molecular to the cellular and to the organismal level, with emphasis on human susceptibility to becoming ill from exposure to low levels of radiation and other environmental agents, and (2) the interactions of organisms with the environment. Innovative ways to observe and understand the functioning of complex biological systems will be developed and applied through expanded partnerships, to meet the needs of the U.S. Department of Energy.

One proposed DOE program is the "Genomes to Life" initiative. Its goals include identifying "protein machines," the multiprotein complexes that carry out the functions of living systems; characterizing functions of microbes in their natural environments; and developing computer modeling technologies to determine how complex biological systems can serve DOE's environmental and health missions. ORNL hopes to shed light on these complicated processes through its systems biology expertise and resources, including DOE's new Center for Structural and Molecular Biology (which will take advantage of the combined expertise of resident and visiting researchers in mass spectrometry, computational biology, and neutron sciences) and DOE's Center for Computational Sciences.

This special issue of the *ORNL Review* showcases ORNL's achievements and capabilities not only in understanding genes and proteins but also in bioengineering developments that could improve health care, such as the lab on a chip, cantilever devices, and the multifunctional biochip. We also have pioneered a way to cure cancer in mice that may have applications to humans. These areas of endeavor are likely to be hallmarks of science and technology in this century.

W. F. Harris, *Associate Laboratory Director for Biological and Environmental Sciences*

Systems Biology: New Views of Life

ORNL scientists are conducting research in functional genomics—the study of genomes to determine the biological function of all the genes and their products—and proteomics—the study of the full set of proteins encoded by a genome.

The genomes of the human, mouse, fruit fly, a worm, a weed, and many microbes have been mapped and sequenced. We now have the parts lists for these organisms. We are learning that many of the parts—the genes that direct cell machinery to produce proteins—are related, from organism to organism. Researchers are now trying to figure out what these parts do in relationship to each other (systems biology) and how they vary among species and individuals within each species. Then researchers can write the operating manuals. The rewards will be great.

"One long-term goal of this research is to develop targeted drugs that are effective for a specific disease," says Michelle Buchanan, director of ORNL's Chemical and Analytical Sciences Division. "To design these therapeutic drugs, you need detailed knowledge about the many molecular-level processes that occur within a cell." Acquiring such knowledge is not an easy task.

"We frequently hear about new human genes that play a role in cancer and diseases of the heart, central nervous system, and other organs," says Reinhold Mann, director of ORNL's Life Sciences Division. "Some diseases can be traced to one altered DNA base pair in a particular gene. However, genome characteristics or changes that make some people more likely to get sick involve complex, intricately timed, and balanced interactions among a variety of genes and other signals encoded in the genome. Our current state of knowledge of how the genome is interpreted to provide the diversity of life is extremely limited."

"The next step," says Buchanan, "is to identify which genes turn on to make particular proteins. Then we must identify the protein complexes, or protein machines, in which proteins work together in the cell to carry out specific roles and help perform life's most essential functions. These protein complexes are involved in signaling pathways that tell cells what to do and allow them to communicate with each other."

Characterizing the roles of protein machines in cells is the objective of DOE's "Genomes to Life" initiative. This knowledge will help scientists predict how cells and their genes will respond to changes in the environment, such as exposure to a toxin.

Genes and Proteins: A Primer

Consider a living cell, the fundamental unit of life. Each human cell contains the entire human genome—some 35,000 genes. But only some genes are expressed within a specific cell, resulting in the production of specific proteins. The genes that turn on in a liver cell, for example, are different from the genes that are expressed in a brain cell.

Under the Human Genome Project funded by the Department of Energy and the National Institutes of Health, various sequencing tools have been used to determine the order of the building blocks of genes. Genes are made of millions of deoxyribonucleic acid (DNA) molecules. A DNA molecule is constructed like a spiral staircase, or a double helix. The rails of the staircase are made of a backbone structure of phosphates and sugars, and the steps are pairs of four nitrogen-containing bases—adenine (A), cytosine (C), guanine (G), and thymine (T). Through hydrogen bonds the two rails of the staircase are kept together, A and T pair together, and C and G are partners (see image below). For example, a sequence of TACAT would bond specifically with a sequence of ATGTA.

A gene is made of a unique sequence of DNA bases; it is like a message containing a unique combination of letters. This message is translated into information for protein production. A protein is a folded chain of amino acids in a specific order; up to 20 different amino acids exist. The message for each amino acid within a protein is dictated by a sequence of three DNA bases. If the key word in the message provided by the gene is misspelled—say, part of it is supposed to be CATTAG but instead is spelled CATGAG—then it will have a point mutation (G substitutes for the T that should be there). As a result of this altered DNA base, the gene may produce a protein that has an incorrect shape so it won't dock with another protein (e.g., a receptor), leading to a mistake in the resulting message. In other words, if the message in the gene is misspelled, the protein it encodes may be wrong and its function in the body may be changed, sometimes for the worse.

It was once believed that each gene codes for a single protein. However, experimental and computational evidence (partly obtained at ORNL) shows that many genes produce an average of three different proteins and as many as ten protein products. Genes have protein-coding regions (exons) interspersed with non-coding regions (introns). Through "alternative splicing," a gene's exons can be combined in different ways to make variants of the complete protein.

Some of this knowledge will be obtained at DOE's Center for Structural and Molecular Biology at ORNL. Scientists at this user center directed by Buchanan will obtain information about protein interactions through mass spectrometry, computational biology, and small-angle neutron scattering (SANS). SANS will be conducted using the planned Bio-SANS instrument that is to be completed at ORNL's High Flux Isotope Reactor in 2003. Various studies of proteins and other biological materials are also planned. These studies will be conducted using biological instruments at the Spallation Neutron Source, to be completed in 2006 at ORNL.

Research by Oak Ridge biological scientists is aimed at learning which genes are expressed when certain microrganisms are exposed to environmental toxins or radiation. DOE is interested in funding microbial research partly because microbes can help remediate mixed waste sites by converting toxic metals from the soluble to the insoluble state to help keep them on-site.

In addition, there is great interest in learning about the proteins that are expressed by microbes under various conditions. ORNL's experts in mass spectrometry can identify proteins by determining their molecular weight and amino-acid sequence and comparing this information with that in a protein database. In this way, they can find protein signatures and then work backwards to determine the gene sequence coding for that protein, thus identifying the gene that was expressed, say, as a result of exposure to a pollutant.

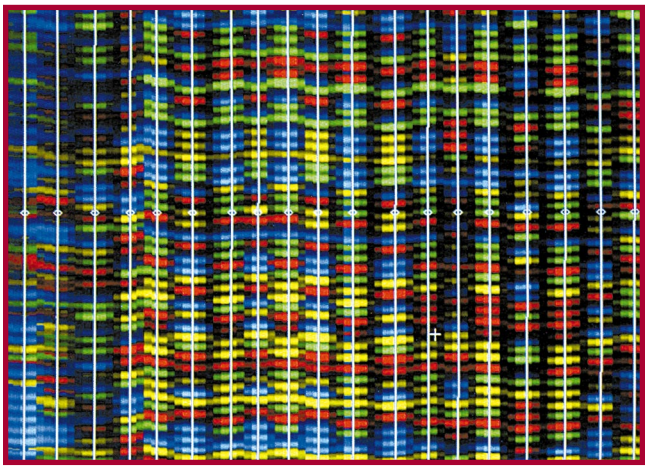
A protein can have as many as 200 modifications in response to the actions of other proteins or environmental influences. These post-translational modifications (PTM), such as the addition of a phosphate or carbohydrate to a protein, can change the protein's activity. For example, if a PTM is present on a regulatory protein A, it becomes a misshapen key that no longer fits into Protein B, preventing it from turning on a downstream gene, possibly causing miscommunication between cells. Mass spectrometry is an excellent tool for identifying proteins that have been modified.

ORNL researchers seek to understand complex biological systems at the organism as well as the molecular and cellular levels, says Mann. To understand how hormone-mimicking chemicals can affect development, ORNL researchers are studying gene and protein expression in see-through embryos of zebrafish. They are also trying to identify the genes that enable trees to produce better wood products and fuels and store more carbon from the air. To understand the functions of genes in mammals, ORNL researchers are determining which genes are expressed in the skin of mice and which mouse genes in their mutant form cause maladies also found in humans, such as polycystic kidney disease, obesity, chronic hereditary tyrosinemia, and epilepsy. Researchers use microarrays (gene chips) and computational tools for these expression studies. For example, our experimental researchers collaborate with our computational biology experts, who make sense out of gene expression data using supercomputers. They are practicing the discipline of bioinformatics, the study of genetic and other biological information using computational and statistical techniques.

"Knowing functions of all genes in the genome, by itself, will not lead to understanding the processes of a living organism," Mann says. "The reason is the biological system's complexity. Expression of genes can be regulated in a virtually unlimited number of ways, depending on location in the body, time in the development of the organism, and environmental conditions and exposures.

"Certain protein complexes can bind to specific locations in an organism's genome, thereby controlling the expression of a gene sometimes far away from these binding sites. The number of these regulatory protein complexes is finite, perhaps some 10,000, but taken together with the number of genes and regulatory binding sites in the genome, there is a combinatorial explosion that works against any brute force approach solely based on experimental research. That is why collaborations between experimenters and computational biologists are so important. They are a hallmark of biological investigations at ORNL."

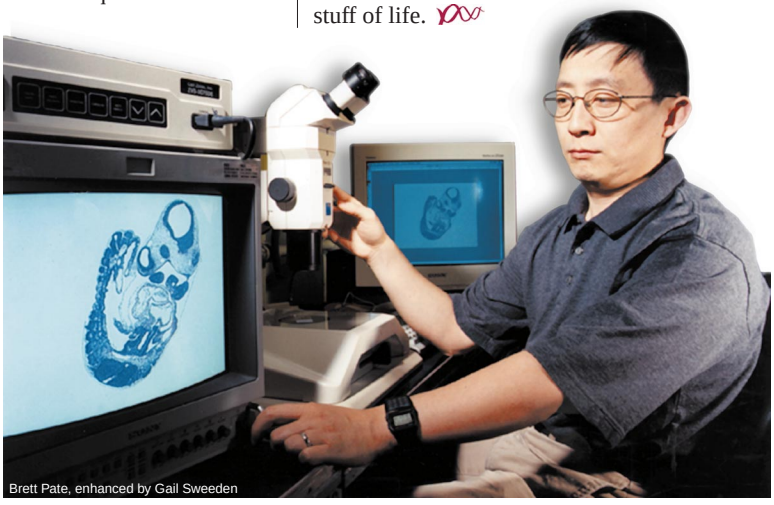
Computational biology researchers in ORNL's Life Sciences Division have identified many genes in bacterial, mouse, and human genomes and have computationally analyzed the human genome using an ORNL-developed gene-finding computer program. They have written and used assembly programs and analysis tools to produce draft sequences of the 300 million DNA base pairs in chromosomes 19, 16, and 5 for DOE's Joint Genome Institute (JGI). They have also analyzed 25 complete microbial genomes (52,000 genes) and many JGI draft microbial genomes (1000 genes/day). They have predicted the structures of proteins (100 proteins/day) from amino-acid sequences using an ORNL-developed, protein-threading computer program.



ORNL researchers use various technologies to characterize DNA and proteins. This image shows the order of chemical bases in a strand of DNA. The bases are labeled with dyes that fluoresce in different colors when exposed to laser light. The sequence was obtained by gel electrophoresis in a PE Biosystems DNA-sequencing machine at ORNL.

The section's programmers have written algorithms and developed other tools to make it easier for biologists to use computers to find genes and make sense out of the rising flood of biological data. These data are produced in studies of biochemical pathways and processes, cellular and developmental processes, tissue and organism physiology, and ecological processes and populations. Through ORNL's user-friendly Genome Channel Web site, its Genomic Integrated Supercomputing Toolkit, and the IBM supercomputer at DOE's Center for Computational Sciences, the international biology community, including pharmaceutical industry researchers and academics, have easily obtained genetically meaningful interpretations of their DNA sequences and other data. ORNL's Web site is popular in the biological community (150,000 sessions per month).

Thanks to ORNL's interdisciplinary approach to complex biology using state-of-the-art technologies, we believe we have the right stuff to better understand the stuff of life.



Yun You examines the image of a mouse embryo magnified in an optical microscope. The image is captured by a digital camera and then transferred to a computer.

Complex Biological Systems in Mice

Using genetic engineering, gene microarrays, and computational technologies, ORNL researchers are deciphering genetic variations in the skin that lead to increased risk of disease from environmental factors.

In a car, when the radiator springs a leak, the car engine heats up dangerously. You stop the car, fill the radiator with coolant, and drive the "sick" car to the repair shop to plug the leak or replace the radiator. The leaking radiator is like an altered, or mutated, gene. By learning that a defective radiator can make the car dangerously hot, you find out that a properly functioning radiator keeps the car engine cool enough to ensure normal operation.

Similarly, ORNL biologists "break" genes in mice to find out their normal roles in a healthy mouse. "We induce mutations in mouse genes and study the resulting disease state in mice so we can determine what the normal versions of the genes do in the body," says Ed Michaud, a senior research biologist in ORNL's Life Sciences Division (LSD). One way that mutations are created is to expose male mice to ethylnitrosourea (ENU), a powerful chemical mutagen discovered by ORNL's Bill Russell in 1979 that alters a single base pair in a gene. Another way is to use recombinant DNA technology or ENU to alter a gene in embryonic stem cells later inserted into mouse embryos. Or an altered gene can be inserted into a fertilized egg, which is implanted in a female mouse and brought to birth as a "transgenic" mouse.

"Because genes operate in complex interacting networks, the mutation of one gene often results in an alteration of other genes in the same network," Michaud says. "Therefore, we are interested not only in determining the functions of individual genes but also how these genes interact with each other and with the environment."

Michaud and other ORNL biologists are studying complex biological systems in mice in collaboration with ORNL researchers using microarrays, mass spectrometers, and bioinformatics—the discipline in which large amounts of data are sorted into intuitive databases, analyzed, and presented in an understandable form. The first complex biological system he has focused on is the network of genes that affect the development and functioning of the skin. In 1992 Michaud and fellow ORNL biologists identified and cloned the mouse agouti gene, which plays a role in the development of skin and hair pigment. He subsequently identified and cloned mutant forms of the agouti gene that cause obesity, diabetes, and skin cancer in the mouse. The mouse agouti gene has a counterpart in the human genome.

Mice and humans each have some 35,000 genes. These genes are distributed among chromosomes, which are long strands of DNA packed in the nucleus of each cell whose job is to determine and transmit hereditary characteristics. The human has 23 pairs of chromosomes and the mouse has 20 pairs of chromosomes.

"We are now interested in genes scattered among all chromosomes of the mouse genome that affect skin," Michaud says. "The skin is a highly metabolic organ with the largest surface area in the body. The skin has many important protective and defensive functions, such as regulating water loss and body temperature and defending the body against chemical and biological agents in the environment. Because the skin comes into direct contact with the external environment, it is ideally suited for studying genetic and environmental interactions. Mutations affecting the skin are also easy to observe and to study throughout the life of the mouse.

"We look at how disease affects animals under different conditions. We are interested in determining how certain genetic mutations make animals more sensitive to environmental toxins. The Department of Energy is interested in the effects of environmental exposures on mice because mice and humans have a similar genetic makeup, and the information from mice can be used to better understand human health risks."

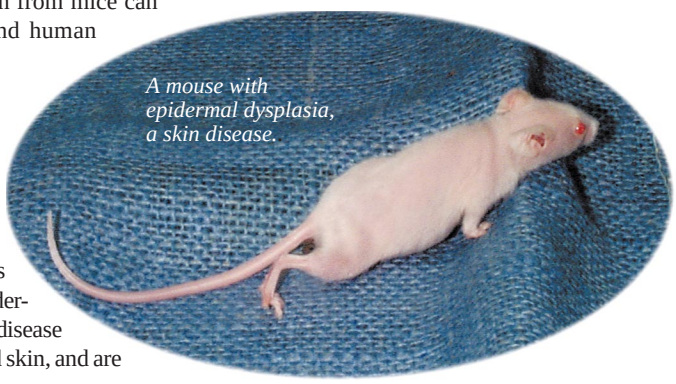
MOUSE MUTANT WITH SKIN DISEASE

Michaud and his colleagues are studying an ORNL mouse mutant that was born with a disease called epidermal dysplasia. Mice with this disease lose their hair, have thickened skin, and are more susceptible to getting skin cancer, observed as tumors on the skin.

"We are interested in seeing which genes are altered in mice with epidermal dysplasia because that information will then point to the normal role that these genes play in the development and functioning of the skin," Michaud says. "Like all organs in the body, the health of the skin is dependent upon the well-orchestrated interactions between complex networks of genes and exposure to numerous environmental variables."

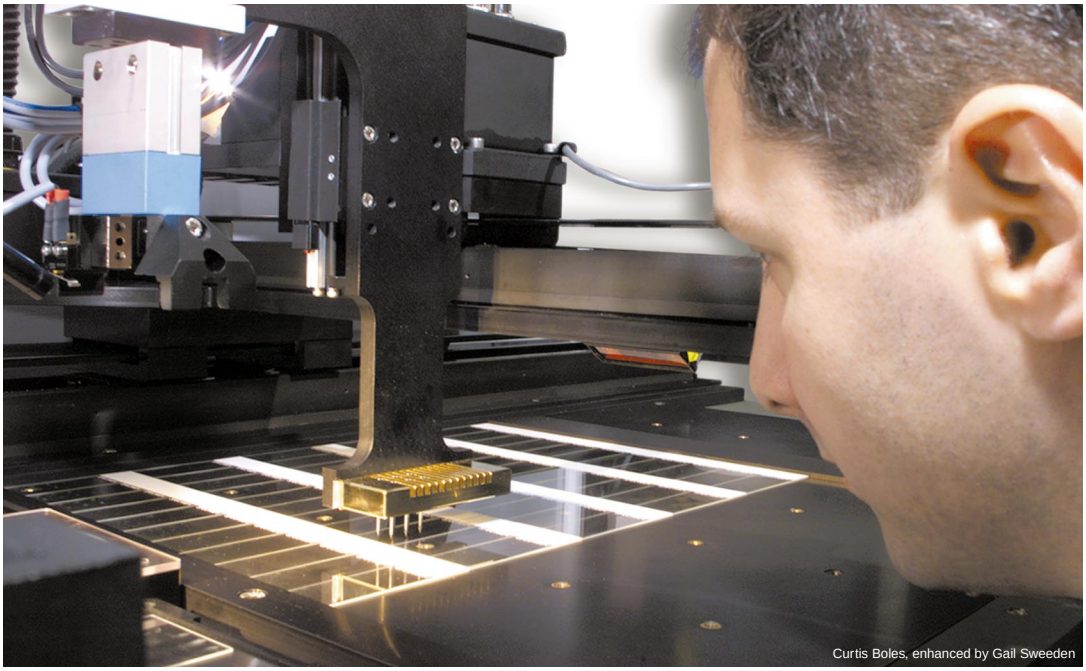
Michaud gives an example of a genetic network. "Suppose that gene A regulates genes B, C, and D," he says. "What happens to these three genes if A is knocked out or mutated?" It's a little like making a hole in the oil pan of a car and then driving it. Eventually, the lack of lubrication of all the mechanical parts will cause the car to grind to a halt. The car breaks down because the motor seizes up, but the hole in the oil pan is the main problem. Determining these types of interactions between genes and the environment gives Michaud and colleagues a better understanding of how the skin protects the body.

Michaud suspects that epidermal dysplasia in the ORNL mice is due partly to a mutated transcription factor gene, whose job is to regulate the function of many other genes. To find out what genes this transcription factor is responsible for, he and his colleagues had to determine which genes are turned on normally when the transcription factor gene is working. Many of these genes would likely have a known function in the skin, such as DNA repair, cell growth, cell



A mouse with epidermal dysplasia, a skin disease.

differentiation, and programmed cell death. If the transcription factor is knocked out, the expression of some of those genes will be altered.



Mitch Doktyz observes the Virtek microarray in a test run in preparation for gene chip production.

MICROARRAYS AND MICE: HOW A GENE CHIP WORKS

To obtain this information, mouse geneticists Brynn Jones, Bem Culiat, and Ed Michaud turned to Mitch Doktycz, Peter Hoyt, and their colleagues in LSD who design microarrays, or gene chips, and use them to perform analyses of gene expression patterns (See following article.) A microarray allows a comparison between genes expressed by a normal organism and genes expressed by a mutant organism or an organism exposed to an environmental toxin. When a gene is turned on, a specific DNA segment corresponding to this gene is copied into a messenger RNA (mRNA) molecule, which is chemically very similar to DNA. This shorter-lived RNA copy moves from the cell's nucleus to the cytoplasm where its code—its sequence of DNA bases, or nucleotides—is translated, causing specific amino acids to be strung together in a specific order to form proteins.

To determine which genes are being expressed at any given time for an important cellular activity, scientists collect the mRNA molecules transcribed in a cell or tissue at that moment. In the laboratory, those RNA messages are reverse transcribed to form more stable complementary DNA (cDNA) molecules. These cDNA samples are prepared from tissues in which biologists want to examine differences in gene expression, such as in skin from mutant mice and normal mice. To detect changes in gene expression, the two cDNA samples are labeled with fluorescent dyes (one for each sample) and then allowed to bind with their complementary DNA templates on a gene chip. The gene chips are glass microscope slides that are spotted with DNA sequences from many hundreds or thousands of different genes in an orderly array. The slides are then exposed to laser light of two different wavelengths and the ratio of fluorescent intensities is measured for each gene. If a gene on the chip is expressed at a high level in the skin of the mutant mouse compared with the normal mouse, the dye used for the mutant sample will shine brighter than the dye used for the normal sample, and this difference can be quantified.

Computers are used to keep track of information for each gene on the microarray. Scientists need to know many variables related to the microarray experiments, including the location of each gene on the array, the DNA sequence and identity of each gene, the mouse tissue that the labeled samples came from, the hybridization conditions, the fluorescent ratios for each gene, and the analysis of the data. At ORNL researchers in LSD's Computational Biology Section—Jay Snoddy, Denise Schmoyer, and Sergey Petrov—are writing the computer programs to handle the data as part of the development of ORNL's Genosensor Information Management System (GIMS).

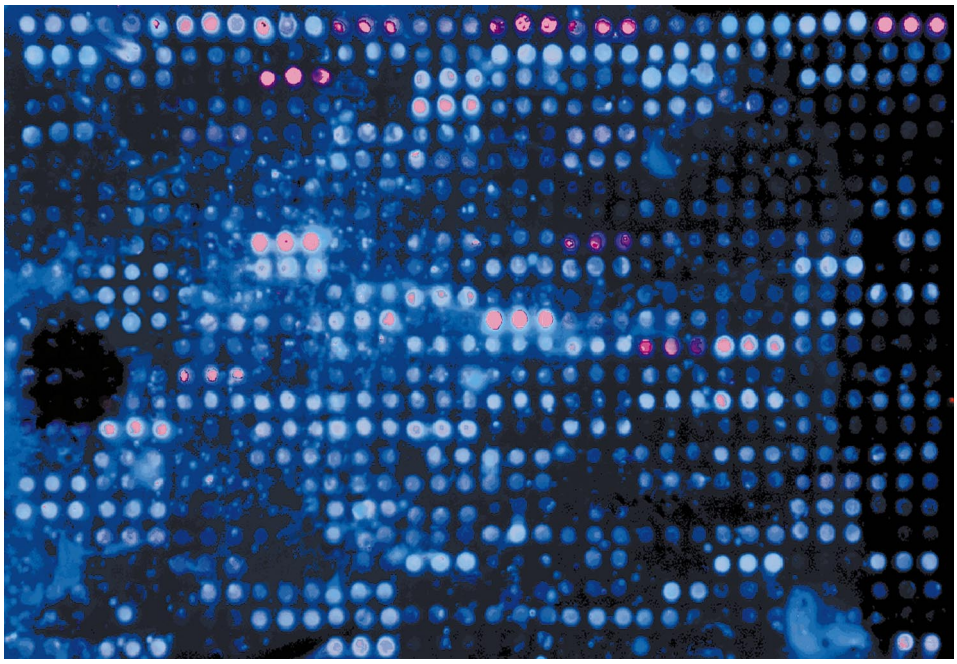
Christine Schar, ORNL-UT Genome Science and Technology Fellow, and Brent Harker check the alignment of the Virtek spotter. This state-of-the-art instrument can produce more than 100 gene chips a day. Additionally, it can place more than 40,000 spots onto a standard microscope slide.

For the analysis of genes from normal mice and mouse models of skin and hair disease, Jones, Culiat, and Michaud designed microarrays containing about 500 mouse genes that were arrayed in triplicate on glass slides. These gene chips were used to determine which genes in the skin of the mice with epidermal dysplasia have altered expression.

As a result of this work, these ORNL investigators made some important discoveries. "We found 30 different genes with altered expression—that is, the levels of mRNAs produced by these genes were different in the mutant mice from those in the normal mice," Michaud says. "We found six altered keratin genes, which are responsible for scaffolding in the skin. We also found four programmed cell death genes and two genes involved in the handling of calcium that were altered in the mutant mouse. These genes are important in the normal development and renewal of the skin."

Several other genes had unknown functions, but the gene chip study indicates that these genes play a role in skin and hair production and function. More biology experiments are being done to verify that these genes have a function related to skin and hair.

Michaud noted that the mice may now be used to determine the effects of various environmental agents on gene networks in the skin. "The skin often protects us from low doses of environmental agents such as ultraviolet radiation, microbes, and chemicals," he says. "Mice with mutations in skin genes, or humans with natural genetic variation, may be at increased risk from these same exposures, and the gene chips can help us to uncover the relevant genes."



This array of blue and pink dots shows the results of a gene chip experiment comparing the gene-expression patterns in mouse skin tissue obtained from a skin mutant and a normal mouse (wild-type control mouse). ORNL's bioinformatics group analyzed gene-expression ratios (ratios of genes turned on in mutant mice and genes turned on in normal mice).

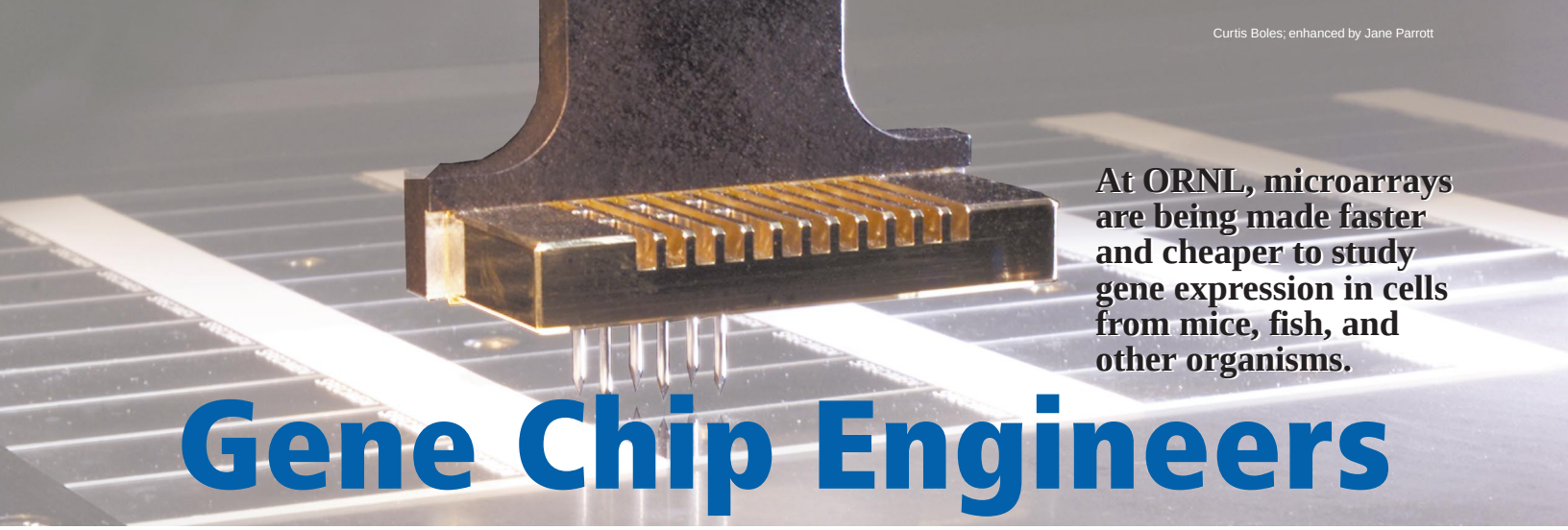
The structure of the proteins produced by the expressed genes—both the normal and mutant ones—is determined by LSD's Gerry Bunick using X-ray crystallography. Doktycz's group is also planning to develop a protein microarray to pinpoint protein-DNA interactions.

"In our study of complex systems biology, we are focusing on genetic variation in the skin and increased susceptibility to disease from environmental exposures," Michaud says, "but this approach for determining which genes are interacting can be applied to the study of any organ, ranging from the heart to the pancreas to the nervous system and brain. For example, we have a collaboration with Russ Knapp, leader of the Nuclear Medicine Group in LSD, in which we are examining gene expression patterns in diabetic mice treated with new insulin-sensitizing drugs."

An important goal of complex systems biology is to understand the functions and interactions of the estimated 35,000 human genes, of which very few are known. Fortunately, in this post-genomic era, new analytical tools, including mutagenesis techniques, gene chips, and supercomputers, are available to help scientists find meaning in the rough drafts of DNA sequences that have been completed for the mouse and human.

Ed Michaud observes the activity of normal and mutant mice in large beakers at the Mouse House.

Tom Cerniglio, enhanced by Gail Sweeden



At ORNL, microarrays are being made faster and cheaper to study gene expression in cells from mice, fish, and other organisms.

Gene Chip Engineers

What does a gene do in a mouse, fish, or some other organism? One technology that allows biologists to spy on gene activity is the microarray, a microscope glass slide dotted with an orderly array of DNA sequences.

Mitch Doktycz, Peter Hoyt, and their colleagues in the Life Sciences Division (LSD) specialize in designing and using microarrays, or gene chips, to help determine which genes are expressed as a result of specific diseases or exposures to environmental toxins. They also have developed technologies to speed up and reduce the costs of preparing DNA probes for gene chips, as well as genetic material from mice, fish, and bacteria.

"We look at hundreds to thousands of samples of biological liquids simultaneously on our microarrays," Doktycz says. "We are evaluating expressed messenger RNA (mRNA) isolated from various mouse tissues, including skin, liver, lung, brain, muscle, kidney, fat, heart, pancreas, spleen, gut, and testes.

In collaboration with Ed Michaud's group in LSD (see previous article), Doktycz's group has analyzed gene expression in samples collected from mice afflicted with skin disease to figure out which genes are altered in the diseased mice compared with equivalent genes in normal mice. A gene that is altered produces abnormally high or low levels of mRNA, which eventually results in altered levels of protein.

Doktycz is also collaborating with ORNL's Russ Knapp and Ed Michaud to determine the genes that are altered in mouse models of obesity and diabetes. Specifically, these experiments use gene chips to determine how new anti-diabetic drugs treat the disease and affect expression patterns of these genes.

Working with Mark Greeley of ORNL's Environmental Sciences Division (ESD), Doktycz and his colleagues have developed a "zebrafish tox-chip microarray." It is used to determine which genes are turned on in zebrafish embryos exposed to hormone-mimicking chemicals.

Besides applying microarrays to gene expression and genome studies, these ORNL researchers are developing more economical ways to prepare samples—extracting mRNA from cells for gene expression studies and attaching various DNA probes to microarrays.

"When we complete development of automated sample processing, hundreds of tissue samples will be processed in an afternoon," Doktycz says. "Using conventional techniques, it can take three to four days to analyze 12 samples. With our method, 96 samples can be prepared in parallel in just a few hours. This sample productivity is needed because we can make more than 100 gene chips in less than a day. Currently, we can print several thousand DNA spots on each gene chip."

Hoyt has developed an inexpensive, high-throughput liquid-handling method of extracting mRNA from tissues. A snippet of mouse skin or other tissue is broken up into cells by simultaneously homogenizing the samples. After the cells are placed in a microtiter plate, mRNA is isolated from them using an automated procedure. Hoyt is now beta testing a new Packard Bioscience robotic instrument for "walk-away" automated processing of mRNA samples. The same instrument is used to prepare the complementary DNA test samples placed on the gene chip (where they will pair with the matching mRNA samples).

Doktyz has been making a mark in the field of spotting technologies. He recently helped devise a commercial inkjet technology for dispensing microscopic drops of biological fluids at high speeds, a technology that could hasten the development of new therapeutic drugs. He worked on this project with Rheodyne, a California company that makes high-end valves, under a cooperative research and development agreement. The resulting "hybrid valve" is now produced commercially by Innovadyne Technologies, Inc., a Rheodyne spin-off company.

Gene chips and related technologies are revolutionizing biological studies. To help meet the need for faster, better, and cheaper ways to spy on genes, Doktycz and his colleagues are being ingenious.

Peter Hoyt checks the operation of the Packard multiprobe liquid-handling system and mechanical gripper. This state-of-the-art robotic tool is used to prepare samples (e.g., cloned DNA) for use on microarrays to analyze gene expression.

Searching for Mouse Models of Human Disorders

Mutant mice are tested by ORNL researchers and their collaborators to determine if these mice have diseases similar to those that afflict humans. Therapies tried on mouse models could lead to new medical treatments.

The lowly mouse is highly regarded as a key to solving a number of medical mysteries. The mouse has long been known to be genetically similar to the human. The recently obtained drafts of the human and mouse genomes suggest that the two genomes are 85% identical: The differences involve a few hundred of the approximately 35,000 genes in both organisms. From an economic standpoint, mice and rats are small and inexpensive to maintain, so it is not surprising that they are used in 90% of the research involving animals.

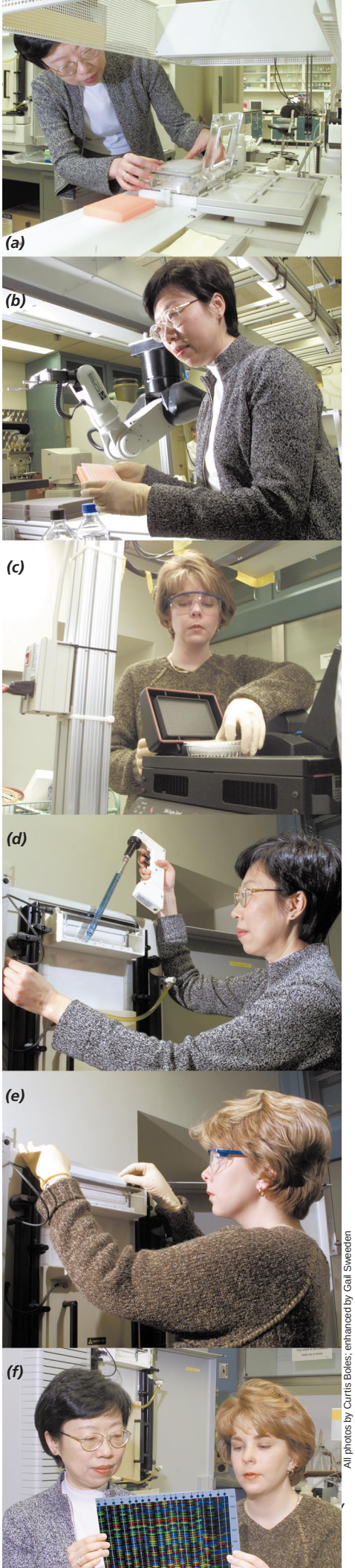
Many mice born with mutations in at least one gene are good models for human diseases. For example, some mutant mice produced by former ORNL geneticist Ray Popp have sickle-cell anemia. They are being studied as models of the human disease at Meharry Medical College, a historically black institution and a participant with ORNL in the Tennessee Mouse Genome Consortium (TMGC).

When a mouse is a model for a human disease, different treatments can be tested on it. Treatments found to control or cure the disease in the mouse could lead to the development of a therapy that works in humans with a similar disease.

Not all mouse models of human disease are perfect yet they may still be useful, according to Dabney Johnson, head of the Mammalian Genetics and Genomics Section in ORNL's Life Sciences Division. "Children with cystic fibrosis die of lung problems, whereas mice with CF die of intestinal blockages. Mice with the CF gene can survive if put on a liquid diet. Perhaps mice have a gene that makes a protein that enables them to compensate for the lung disorder caused by the CF gene. If so, knowledge of the structure and function of this protein could be the key to developing a drug that could benefit humans with CF."

To determine whether ORNL's mutant mice are good models for human diseases of the central nervous system (CNS), Johnson is collaborating on projects with the partners of TMGC. The TMGC recently received a $12.7 million grant from the National Institute of Mental Health (NIMH) to create 25 to 50 new strains of mutant mice that will be used to study neurological disorders. The partners include the University of Tennessee at Knoxville, UT-Memphis, St. Jude Children's Research Hospital, Vanderbilt University, and Meharry Medical College. The TMGC is interested in mutations leading to new genetic information about neurological conditions ranging from Alzheimer's and Parkinson's diseases to depression and addiction.

"We produce mutant mice by treating males with the powerful mutagen ENU (ethylnitrosourea), mating them with females with particular genetic characteristics that help trace the newly induced mutations, and screening their descendants for disorders in the brain and central nervous system," Johnson says. "The screening is done using several tests. For example, mice are placed on a spinning rod to see how well they can maintain their balance there before falling off. Mice having certain mutations lack the coordination and balance of normal mice and fall



Two researchers characterize normal and mutant genes for ORNL biologists. (a) Tse-Yuan Shen Lu, a research associate with ORNL's Life Sciences Division, loads cultures of bacterial cells containing mouse DNA in a Qiagen BioRobot, which extracts DNA from the cells into a liquid. (b) Shen Lu operates a liquid dispenser for simultaneously filling 96-well plates with DNA. (c) Melissa York, a research assistant with the University of Tennessee's Graduate School of Genome Science and Technology, places a 96-well plate in a polymerase chain reaction (PCR) machine for making thousands of copies of each DNA sequence in the plate. (d) Using a pipette, Shen Lu loads running buffer into the PE Biosystems automated DNA-sequencing machine. (e) York inserts into the machine a loading comb on which fluorescent DNA has been absorbed. (f) The researchers show an image of the sequence of DNA bases labeled with dyes that fluoresce when exposed to laser light. The sequence was obtained by gel electrophoresis in the DNA-sequencing machine at ORNL.

All photos by Curtis Boles; enhanced by Gail Sweeden

off this rotor-rod more quickly. We can detect whether a mouse is depressed by observing its behavior in a swimming test. If it tends to float rather than swim vigorously to try to get out of the water, we classify it as a depressed mouse.

"An activity test is used to determine if a mouse is underactive or overactive as a result of a CNS mutation. In this test to gauge a mouse's activity in a box, a photobeam sensor counts the number of times per minute that a mouse interrupts light beams sent across the box."

Johnson and her associates also use this test to measure a mouse's anxiety level. Mice are, by nature, anxious creatures. "A normal mouse stays near the wall where it is more protected rather than going to the middle of the box where it would be out in the open and feel more exposed to predators," she says. "A less anxious mouse, one that is calmer than the normal mouse and, therefore, likely to have a CNS mutation, ventures forth into the open space."

Johnson's group also uses an array of tests to measure learning abilities and memory in mice to screen for CNS mutations. For example, the researchers administer a mild foot shock and play a sound at the same time. A day later, when a normal mouse hears that sound, it will freeze in fear that the unpleasant shock may occur again. Mice with certain CNS disorders will ignore the sound and continue their activity.

As a part of that same test, a normal mouse returned to the same box 24 hours later will recognize the box as the site of the shock and freeze. However, a mouse with a CNS disorder will be just as active in the box as it was 24 hours before, prior to the administration of the shock. The two parts of this test measure two different kinds of memory.

Another CNS test is the startle test. "A normal mouse will have a measurable startle response when it hears a loud noise," Johnson says. "It will flinch, and this action will be detected by a load sensor. But an abnormal mouse may not startle at all, perhaps because it is deaf. Or a mouse with a CNS disorder could startle too much rather than simply jump or flinch."

The TMGC partners help ORNL screen mice for new mutations and analyze confirmed mutations in more detail. One new mutation recently discovered by ORNL's Eugene Rinchik causes the mouse born with it to have continuous seizures. Mice with this mutation have been sent to consortium researchers who then conduct studies to determine if the cause of the seizures is neurochemical or neurophysical. They will try to determine if this "seizure" mouse is a good model for some form of human epilepsy.

At ORNL, researchers run automated analyses of blood and urine samples from mice, measuring their white and red blood cell counts and hemoglobin concentrations. The dip stick urine test is used to measure

Tom Cerniglio, enhanced by Gail Sweeden

Dabney Johnson watches a mutant mouse afflicted with "behavorial despair" (left) make no effort to escape the water in a tank while the normal mouse swims rapidly in an effort to get out of the water during this swim test.

Mouse Models for the Human Disease of Chronic Hereditary Tyrosinemia

When a section of mouse chromosome 7 containing the coat color c gene is deleted by exposing mice to radiation, "albino" mice are born with a white, hairless coat. ORNL researchers found that a nearby gene that is also knocked out by irradiation causes chronic hereditary tyrosinemia in mice, a disease that also afflicts humans. Patients with this disease can develop a wide range of liver and kidney problems, as well as problems affecting the eye, the skin of the feet and hands, and the central nervous system.

Using the powerful mutagen ENU to alter a single DNA base pair in this gene, ORNL biologists recently produced two mouse models that more closely mimic the gene mutation that causes chronic hereditary tyrosinemia in humans.

Normal mice and people metabolize tyrosine, an amino acid available in food, to make melanin, a type of pigment produced in large amounts by dark-skinned people. But people and mice with the disease lack a normally functioning protein (enzyme) to carry out one step of the tyrosine metabolism process, which involves a series of enzymes. So, unless people with this disease are put on a special tyrosine-free diet, a substance that is not broken down because of the absence of a normally functioning enzyme will build up to toxic levels in the liver and kidneys, a fatal condition in mice.

By sequencing the same chromosomal region from both normal and abnormal mice and comparing the sequences, ORNL biologists identified new mouse models that carry mutations in this enzyme in the tyrosine breakdown pathway. "Mice with this disease die of poisoned livers," says ORNL biologist Dabney Johnson. "A by-product of the botched metabolism process is succinyl acetone, which accumulates in the liver and is excreted in urine where it serves as a diagnostic indicator of the disease. Because mice entirely lacking this enzyme die right away, we exposed male mice to ENU to produce mice with this enzyme in crippled form, rather than entirely missing, so we would have a live mouse model of a disease that some humans have. In this way, interested researchers could assess the effect of therapy on mice with chronic hereditary tyrosinemia."

Using X-ray crystallography, ORNL's Gerard Bunick, along with co-workers Joel Harp and David Timm, determined the structure of the enzyme that is responsible for hereditary tyrosinemia in the mouse. The mouse enzyme serves as an easily studied model for the same disease in humans.

"I found that this enzyme folds into a three-dimensional shape that has never been seen before," says Bunick. "On the basis of our structural observations, we were able to propose how the enzyme works, allowing us to identify the key amino acids that lead to dysfunction of the enzyme when mutated, resulting in tyrosinemia. We also identified the location of several sites of known mutation in the structure that could cause the protein to fold into an incorrect three-dimensional shape, which would also cause dysfunction of the enzyme."

A paper on this research has been accepted for publication in the journal *Structure*. A paper on the new mouse models for the disease has been published in the *Proceedings of the National Academy of Sciences*. A detailed understanding of the enzyme may lead to a drug to treat hereditary tyrosinemia in humans.

Brett Pate

Mutant mice lacking the coordination and balance of normal mice fall off the rotor-rod more quickly.

for sugar concentration and excess protein. This information tells the researchers whether mice are anemic or diabetic or suffer from infections, leukemia, or blood-clotting problems typical of hemophilia.

"Copies of our mouse mutants go to UT at Memphis, which screens mice for mutant genes that cause addiction to alcohol and drugs such as cocaine," Johnson says. "For example, our mice are given the two-bottle test in which one bottle contains water and the other, alcohol. A mouse with a genetic predisposition for alcoholism might drink from the alcohol bottle when thirsty, but a normal mouse drinks only from the water bottle. A normal mouse injected with alcohol falls off the rotor-rod quickly, and it loses its inhibitions and shows less anxiety—like most people who have too much to drink."

Mice suspected of having CNS disorders, based on ORNL tests, are sent to the UT Memphis Health Sciences Center. The mice are sacrificed and their brains are sliced, stained, and studied to determine if they have an abnormal anatomy. The eyes of these mice are also examined to determine whether the retina, neural connections, and other components are formed normally. Researchers also check eye samples for signs of macular degeneration and other predictors of impaired vision.

For many people, mice can be a nuisance, but the results obtained from research using mice could give victims of some diseases a new lease on life.

Obesity-related Gene in Mouse Discovered at ORNL

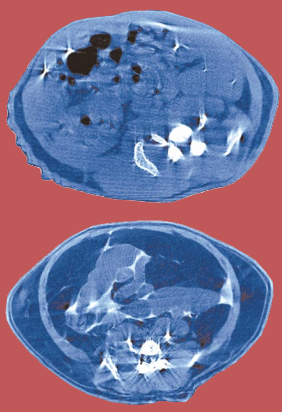
Some mice born at ORNL have grown dangerously fat, even though they have been on a low-fat diet since birth. Although they do not appear overweight, these mice have a mutated gene that plays a strong role in causing obesity in the form of internal fat deposits that are hazardous to their health. The gene was discovered by Madhu Dhar, postdoctoral fellow from the University of Tennessee who works in the laboratory of Life Sciences Division's Dabney Johnson. Dhar's research was funded by the National Institutes of Health, and the findings were published in a paper in a recent issue of *Physiological Genomics*.

"We have found that the normal mouse has a gene on chromosome 7 that probably plays a role in the transport of fat from the blood into fat cells, where fat is stored as a source of energy to keep the body healthy," Johnson says. "If a mutant form of this gene is inherited from the mother in certain genetic backgrounds, the offspring grow 35 to 50 percent fatter by middle age than does a normal mouse, even though they are eating food low in fat."

Unlike some known mouse obesity genes that can act all alone to cause excessive body fat, the ORNL researchers have shown that the chromosome 7 gene must act in concert with other genes involved in maintaining the body's energy balance. If female mice possessing this mutated gene are mated with males having different genetic backgrounds, the offspring may not become obese, suggesting they have genes that code for proteins that suppress fat accumulation. Humans and mice are genetically similar and produce similar proteins. In humans, chromosome 15 is similar to chromosome 7 in the mouse.

The ORNL group has been focusing on chromosome 7 in the mouse for some time. In the 1950s ORNL geneticists Bill and Liane Russell irradiated mice and observed that some of their offspring who survived had a pinkish coat and pink eyes instead of normal gray-brown fur pigmentation and dark eyes. It was later determined that the radiation knocked out DNA from a coat-color gene, called the pink-eye marker (*p*), on mouse chromosome 7. Neighboring genes also were deleted or altered in some irradiated mice, and the defects were passed on.

Thanks to the availability of improved technologies, ORNL researchers have defined the small region on mouse chromosome 7 that contains the *p* gene and its neighbors. Using recombinant DNA techniques, they identified and characterized genes from this small region in the normal mouse. And in mutant mice they have identified behavioral oddities, internal defects, and disorders such as epilepsy and obesity caused by the absence or alteration of certain neighboring genes on chromosome 7. From this information on genetic material gone awry, they can determine the function of the normal genes in this chromosome region.



MicroCAT images showing fat deposits in a normal mouse and a mouse with the obesity gene on chromosome 7.

Johnson says that, like humans, mice deposit fat in their bodies in different patterns that are genetically controlled. The distribution of fat deposits in mice with the obesity-related gene has been observed in three dimensions by Mike Paulus and Shaun Gleason in ORNL's Instrumentation and Controls Division (see article on p. 11). They have imaged internal fat deposits in mice using the MicroCAT miniature X-ray computerized tomography system they developed. Using special software, they have determined the size and weight of those fat deposits.

Madhu Dhar shows how a mouse is placed in a MicroCAT X-ray computerized tomography system. This device has been used to produce images showing the locations of life-threatening fat deposits in the body of a mouse with the newly discovered obesity-related gene. Mice with this gene have fat pads in known regions of the body. Thanks to biomedical imaging using the MicroCAT scanner, it is no longer necessary to sacrifice fat mice and dissect them to measure the size and weight of fat pads.

Curtis Boles; enhanced by Gail Sweeden

MicroCAT "Sees" Hidden Mouse Defects

ORNL's X-ray computed tomography system allows internal defects and organ changes in small animals to be mapped.

The white mouse sleeps on a narrow wooden trough outside a new high-resolution X-ray computed tomography (CT) system at ORNL called a MicroCAT scanner. The anesthetized mouse is gently inserted into the instrument, where it is scanned for about 20 minutes.

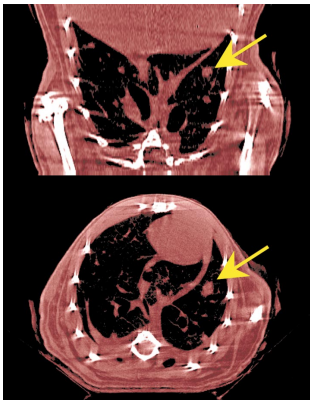


Image of a tumor in the lung of a mouse injected with lung cancer cells. This MicroCAT image correlates reasonably well with histology data. The results were published by ORNL's Steve Kennel in the May 2000 issue of Medical Physics.

Mike Paulus and Shaun Gleason, researchers in ORNL's Instrumentation and Controls Division who developed the MicroCAT scanner, note that the scans show spots indicating the locations of lung tumors of 1 to 2 millimeters in diameter. The mouse is then returned to its cage. The mouse is scanned several times over the next few days to get a profile of tumor growth in its lungs.

The mouse has tumors because it has been injected with lung cancer cells. Steve Kennel, a researcher in ORNL's Life Sciences Division (LSD), creates mice with lung tumors to test how well they respond to a special kind of radioimmunotherapy. When Kennel decides it is time to treat the white mouse, he injects it with a monoclonal antibody tagged with radioactive bismuth-213 (a decay product from ORNL's stockpile of uranium-233). This monoclonal antibody is like a smart bomb because it homes in on the tumor's blood vessels, where the radioactive bismuth-213 parks. The alpha radiation from the radioisotope destroys the nearby tumor tissue.

The mouse is scanned repeatedly over a period of days both during and after treatment. The MicroCAT pictures reveal that the tumors are progressively disappearing, indicating that the treatment works. (See following article.)

The MicroCAT scanner, which was developed using internal funding from ORNL's Laboratory Directed Research and Development Program, has been used to provide three-dimensional images for other small-animal studies:

- Measurement of the size, weight, and distribution of fat deposits in mice found to have an obesity-related gene on chromosome 7 (in collaboration with Madhu Dhar, as described in the article on p. 10);
- Imaging of tumors on the prostate glands of mice produced by researchers at Baylor University;
- Imaging of rat bones to measure their resistance to the passage of X rays and their bone wall thickness as an indication of bone mass and amount of calcium in bone, as part of a study of osteoporosis in animals;
- Imaging the uterus of a pregnant mouse repeatedly over time to determine the number and time of death of individual fetuses (as a result of genetic disorders);
- Validation of ORNL veterinarian Charmaine Foltz's body condition scoring system, in which she presses her thumb against the body of each mouse to rank the animal's health.

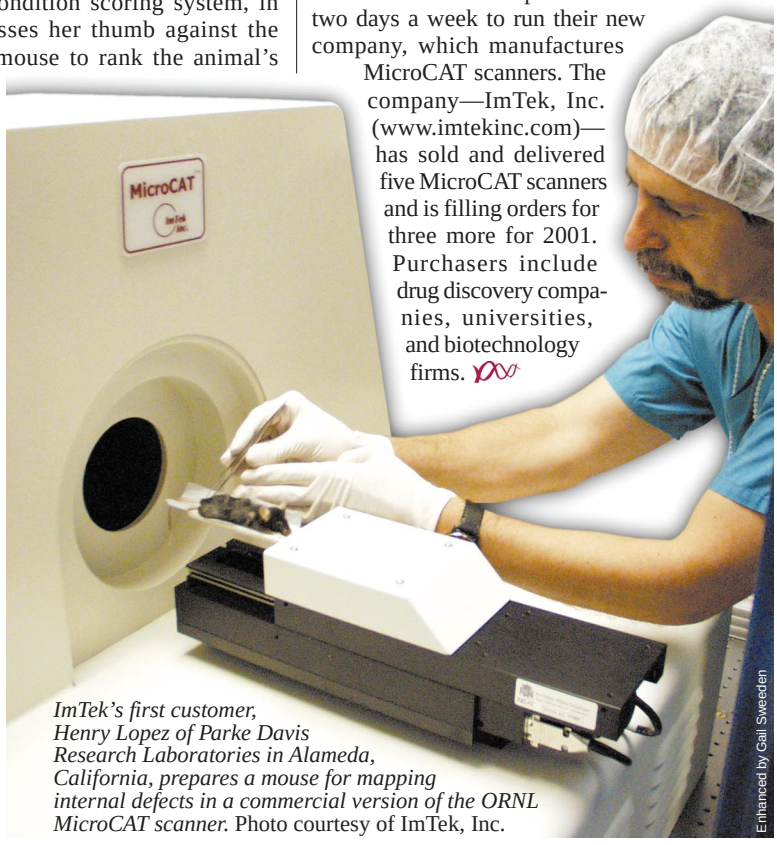
A second-generation MicroCAT instrument is being used for high-throughput screening of mutant mice for internal defects in ORNL's Laboratory for Comparative and Functional Genomics (Mouse House). The device saves time and money because biologists can screen mice individually for internal mutations in 7 to 20 minutes, without sacrificing and dissecting the animals.

To improve accuracy and speed in assessing mouse organs for defects and damage, Gleason has developed an automatic organ-recognition algorithm for CT images of the mouse. Gleason has shown that in the CT scan, the software can zero in on the mouse's kidneys and analyze differences in kidney texture that would show up in mice with polycystic kidney disease. He has also used the algorithm to calculate the approximate size of the lung and heart and evaluate the lung's level of fluid or amount of scar tissue, as an indication of its health or disease state.

Paulus has also used the MicroCAT scanner for research collaborations that do not involve small animals. For example, he has worked with plant geneticist Gerald Tuskan of ORNL's Environmental Sciences Division to image three-dimensional details of wood cells and cell wall thicknesses in samples of loblolly pine (see article on p. 28). Paulus is writing an algorithm that will enable automatic measurement of the size and shape of wood cells.

Paulus and Gleason have developed a new instrument that combines X-ray imaging with single-photon-emission-computed tomography (SPECT). This MicroCAT SPECT scanning instrument could be used to image lung tumors in mice and detect radioactivity from the treated tumors, to map their precise location.

With support from UT-Battelle, Paulus and Gleason are on entrepreneurial leave two days a week to run their new company, which manufactures MicroCAT scanners. The company—ImTek, Inc. (www.imtekinc.com)—has sold and delivered five MicroCAT scanners and is filling orders for three more for 2001. Purchasers include drug discovery companies, universities, and biotechnology firms.



ImTek's first customer, Henry Lopez of Parke Davis Research Laboratories in Alameda, California, prepares a mouse for mapping internal defects in a commercial version of the ORNL MicroCAT scanner. Photo courtesy of ImTek, Inc.

Curing Cancer in Mice

ORNL researchers have shown that a radioisotope-bearing antibody can target the blood vessels of lung tumors in mice, destroying the tumors.

One injection can give a mouse cancer but a second can cure it. That was one of the conclusions made from recent experiments on mice conducted by Steve Kennel and Saed Mirzadeh, both of ORNL's Life Sciences Division.

Dozens of mice were injected with lung cancer or breast cancer cells. The injected cells lodged in the mouse lungs and grew there. Several of the mice were later anesthetized, and the tiny, solid tumors in the lungs were imaged by high-resolution X-ray computed tomography, using the ORNL-developed MicroCAT scanner (see previous article).

The cancer-stricken mice were then injected with a special monoclonal antibody chemically hitched to a radioisotope produced at ORNL. In this protocol, the antibody targets the blood vessels of each tumor like a smart bomb; the antibody, which is a protein, docks with proteins found in lung blood vessel cells. The radiation from the parked radioisotope destroys the tumor cells around the blood vessel but, remarkably, leaves the vessel intact. Using the MicroCAT scanner, ORNL's Mike Paulus took images of the mice every few days; the images showed the gradual disappearance of the tumors."

"We demonstrated that this technique cures mice of lung tumors," Kennel says. "When we can deliver the radioisotope to the blood vessel that serves the tumor, we can kill the tumor. As a result of this radioimmunotherapy, the life spans of the treated mice are extended dramatically."

By contrast, the mice with implanted cancer cells that were not treated died in 15 days. The mice that received too low a dose of radiation also eventually died of the lung tumors.

Interestingly, the mice that received higher radiation doses were cured of cancer, but died earlier than normal, healthy mice. "These mice lived much longer than the other mice in the experiment, but they eventually died of pulmonary fibrosis," Kennel says. "The reason is that their lungs exhibited an inflammatory response as debris-collecting white blood cells were recruited to the lung to remove the damaged and dead cells."

Other types of experiments in radioimmunotherapy have been performed for years in mice, but the success rate has been low. "Previous approaches for radioimmunotherapy of solid tumors does not work using labeled antibodies that bind directly to tumor cells, because antibody stays in the blood and only a small fraction reaches the cells in the solid tumor," Kennel says. "Our approach has been to select or make a radiolabeled antibody that targets and parks in the blood vessels in the solid tumor. The type of radiation used is an alpha particle emitter that kills every cell within 100 microns."

The researchers found that a targeted antibody labeled with either radioactive bismuth-213 (^{213}Bi) or astatine-211 was most effective. Bismuth-213 is an alpha emitter obtained at ORNL; it is a decay product from the Laboratory's stockpile of uranium-233 left over from its molten-salt reactor experiments in the late 1960s. The astatine-211 used comes from the National Institutes of Health.

Recently, Kennel and a postdoctoral scientist, Sandra Davern, used a bacterial virus engineered to display antibody-like molecules. The bacterial virus was injected into mice with lung tumors. The researchers found that some of the virus stuck to tumor blood vessels. These engineered antibodies were separated by molecular biology techniques from the virus and replicated and amplified in bacteria for use in radioimmunotherapy experiments in mice.

These ORNL successes in eradicating lung tumors in mice could provide important clues about how to achieve effective human cancer cures.

Search for Signs of Inflammatory Disease

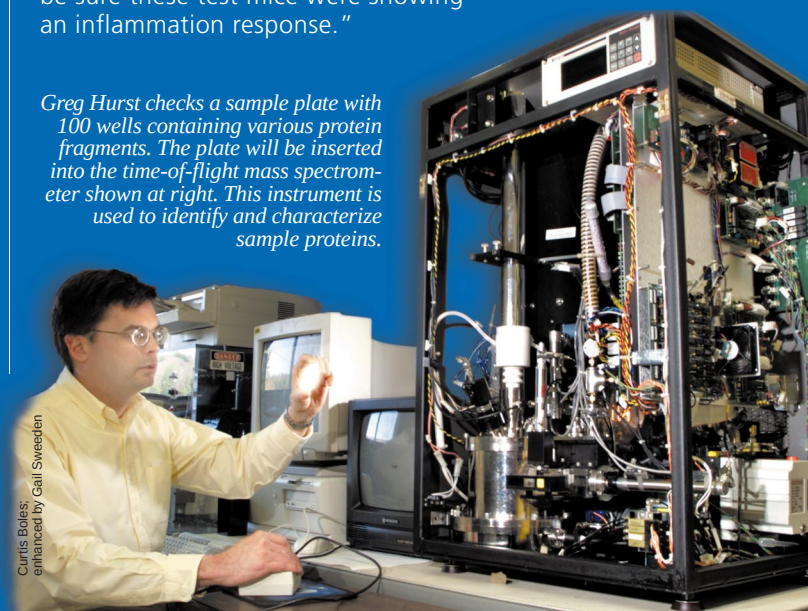
You fall on your shoulder and tear some cartilage, causing bone to rub against bone. Your shoulder becomes inflamed and begins to hurt because cytokine, a small signal protein secreted by your immune system, has recruited white blood cells to clean up the damage.

ORNL biologist Steve Kennel recently worked on a project to measure levels of a specific cytokine in samples from mice. The goal was to screen for mice that are likely to develop inflammatory diseases similar to those in humans, such as arthritis, psoriasis, and inflammatory bowel disease.

In a study supported by the Laboratory Directed Research and Development Program at ORNL, Kennel collaborated with Greg Hurst, a mass spectrometry expert in ORNL's Chemical and Analytical Sciences Division. They developed a technique that combines affinity chromatography with mass spectrometry to separate three specific cytokines out of the 100 or so different cytokines in mouse blood or fluid extracted from mouse cells.

"I obtained antibodies specific to different cytokines and attached them to beads," Kennel says. "These beads were tossed into a soup of proteins to fish out specific cytokines for analysis by matrix-assisted laser desorption ionization mass spectrometry.

"We were able to identify each cytokine. But because we had only small amounts of blood from a mouse, we were unable to detect unusually high levels of our target protein—the tumor necrosis factor alpha cytokine. So, we could not be sure these test mice were showing an inflammation response."



Greg Hurst checks a sample plate with 100 wells containing various protein fragments. The plate will be inserted into the time-of-flight mass spectrometer shown at right. This instrument is used to identify and characterize sample proteins.

Curtis Boles; enhanced by Gail Sweeden

Surprises in the Mouse Genome

In the live organism, not all mouse and human genes have predictable functions, and proteins with similar structures can have different functions.

Like a mouse darting across the hearth and disappearing behind the home entertainment center, the mouse genome can surprise even the most seasoned geneticist. For Eugene Rinchik of ORNL's Life Sciences Division, the "unexpected discovery" has been a theme in his career of producing and studying mutant mice.

"We find a lot of wonderful surprises in this research," he says. For example, in the early 1990s, Rinchik and ORNL scientist Bem Culiat were studying a form of inherited cleft palate, a facial deformity and birth defect, in mice. They found that newborns affected with cleft palate were missing a certain neurotransmitter receptor gene. By adding a rat gene that codes for this neurotransmitter receptor to a fertilized mouse egg lacking the mouse gene, Culiat corrected the disorder, and the resulting mouse was born without a cleft palate.

"The 'surprise' was that a reasonable prediction for the function of this gene would not have included effects on the palate during fetal development," says Rinchik. "If we used only the known biochemical function of this receptor as a guide, we would expect to find neurological dysfunction, not cleft palate, as the primary defect."

In another example of surprising findings, Rinchik frequently tells audiences about four different genes in mice that encode proteins that share several structural characteristics. Computational gene modelers have classified these proteins as cell-signaling molecules. In general, such molecules instruct cells to divide, grow, step up their metabolism, or die, for example.

"Scientists have studied mice with mutations in each of these four genes," says Rinchik. "One mutant gene results in a defective protein that causes mouse embryos to die in the uterus. Another mutant causes mice to be born with cleft palates (for an apparently different reason than that discussed above). The third mutant results in inflammatory disease in young animals, and the fourth causes the mouse to be born with slight cartilage abnormalities that show up as shortened ears in an otherwise healthy animal.

"The point is that although these proteins belong to the same general family of cell-signaling molecules, they have different functions in the mouse. The evidence gained from mouse-breeding experiments improve understanding of what happens at the level of the organism and, therefore, add value to computational predictions about the biochemical functions of genes in the mouse genome."

Rinchik and his colleagues continue to look for new dominant and recessive single-base gene mutations in the descendants of mice exposed to ethylnitrosourea (ENU). He was inspired to use ENU by long-time ORNL geneticists Liane and Bill Russell, who pioneered its use for producing mouse mutants that could be models for human disease, making ORNL a world leader in this area. For example, recently, Rinchik was pleasantly surprised to find a new mouse mutation that could shed light on a human disease. Some of his mutant mice were found to have seizures continuously for a few weeks until they died. These mice may be models for the human disorder epilepsy. The mice are currently being characterized by Lisa Webb, a graduate student working in Dabney Johnson's group at ORNL.

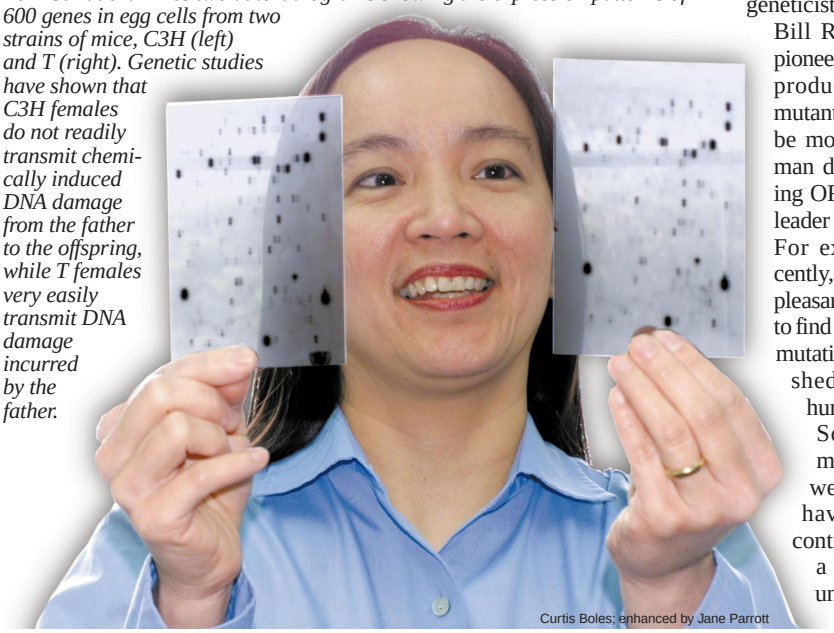
In 1979 ORNL's Bill Russell discovered that the chemical ENU causes mutations in mice. He and his wife Liane pioneered ENU's use for producing mouse mutants that could be models for human disease.

"Now that the DNA sequence of the mouse genome is being completed by the public sector, we should be able to locate and identify a mutated gene more rapidly," Rinchik says. "This can be done by comparing a DNA sequence with an altered base from a mutant mouse with the normal DNA sequence from the mouse genome map."

Rinchik sees more collaborations in the future between ORNL mouse geneticists and human geneticists. There is already a model for such interactions.

"In 1993, in a collaboration with Rob Nicholls, a human geneticist who is now at the University of Pennsylvania, we identified the human version of the mouse pink-eye dilution gene, which leads to a pigmentation defect in mice," Rinchik says. "Subsequently, human geneticists found mutations in this gene to be responsible for albinism in black Africans. They are born with little pigment and are light-skinned as a result."

Culiat, meanwhile, is using newly available molecular tools to explain a surprising finding by retired biologist Walderico Generoso. He found evidence suggesting that eggs from some female mouse strains can correct damage in sperm from male mice exposed to a toxic chemical, reducing the percentage of embryo deaths. By studying gene expression profiles, Culiat hopes to determine whether these eggs repair damaged DNA or have an extracellular filter that lets in normal sperm and keeps damaged sperm out. Indeed, the results could be surprising.



Curtis Boles; enhanced by Jane Parrott

Bem Culiat examines two autoradiograms showing the expression patterns of 600 genes in egg cells from two strains of mice, C3H (left) and T (right). Genetic studies have shown that C3H females do not readily transmit chemically induced DNA damage from the father to the offspring, while T females very easily transmit DNA damage incurred by the father.

Protein Identification by Mass Spectrometry

ORNL researchers are improving mass spectrometry tools to speed up protein identification and to screen for disease-causing proteins and bacteria.

Proteins are large, complex molecules that carry out the tasks of life. They direct our bodies' activities, organize our thoughts, and defend us against infection, keeping us healthy. But in their mutant forms and as coats on disease-causing microbes, proteins can help make us ill and threaten our health.

Each protein is initially formed as a string of amino acids whose identity and order are dictated by a gene according to the sequence of its DNA bases. The gene's instructions—carried by messenger RNA—also call for this string to be folded into a three-dimensional molecule that has an intricate shape, ranging from a saucer to a dumbbell to a corkscrew.

The amino-acid composition and sequence, as well as the molecular weight, of a protein produced by a certain type of bacteria are different from those of the protein forming the coat of a particular virus. Such a coat enables this molecular terrorist to break through the protective membrane of a cell and command it to produce more virus. Each type of microbe produces unique proteins, providing a characteristic protein signature and allowing identification of the microorganism. Thus, if the signature proteins in anthrax spores and botulism toxins could be accurately detected, it would be possible to provide an early warning about the proximity of biological warfare weapons.

The ability to identify proteins is also important because it allows researchers to determine whether an organism has a genetic disease. A genetic disease is often caused by a mutant protein, which has a composition slightly different from that of the normal protein it replaces.

One of the most powerful tools for detecting and identifying proteins is mass spectrometry, a technique that has been improved and used for a variety of research projects for many years at ORNL. A mass spectrometer sorts out charged particles according to their masses, allowing analysis of the elemental composition of complex molecules. A mass spectrometer produces a spectrum consisting of peaks and valleys that indicate the identity and number of different atoms making up the molecule being analyzed.

The mass spectrometer is an ideal instrument for identifying amino acids—the building blocks of proteins—and determining the order in which they are arranged. The mass difference, or distance in atomic mass units between the peaks along the spectrum, allows each amino acid (e.g., alinine, arsenine, glycine, or lysine—four of the 20 possible amino acids) to be identified.

NEW IONIZATION METHODS AID PROTEIN ANALYSIS

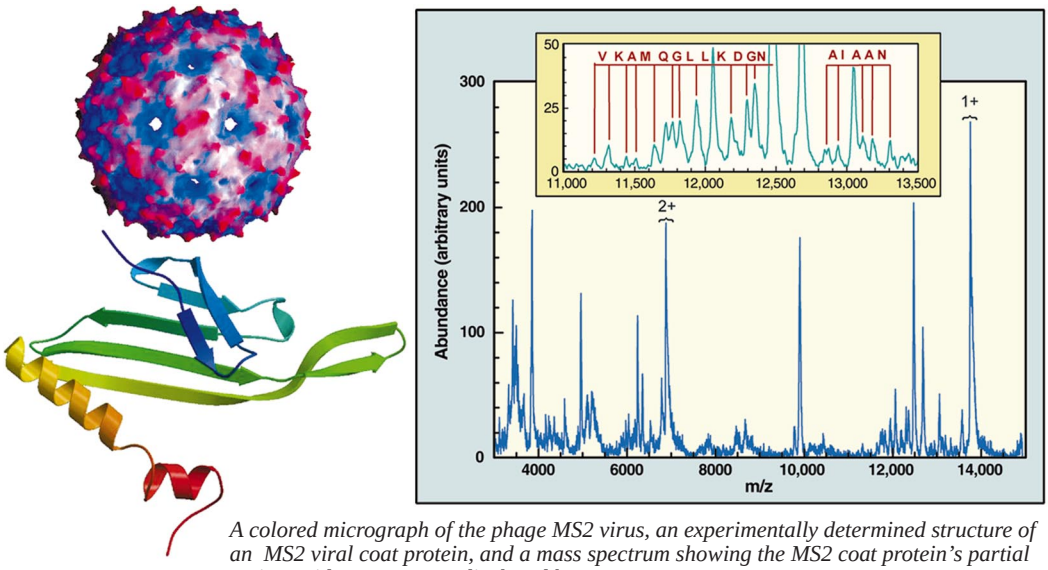
With the discovery in the past decade of two powerful methods that can be used to produce ions for analysis in a quadrupole ion trap mass spectrometer, progress in biological mass spectrometry has been nothing short of revolutionary. Mass spectrometry has become, in just a few years, an important tool for protein identification, peptide sequencing, identification and location of post-translational modifications of proteins, analysis of modified deoxyribonucleic and ribonucleic acids (DNA and RNA), and many other biological applications. These advances have become possible only through the capability to form ions from large biomolecules and through the research community's growing understanding of the chemistry of biological ions.

During the mid-1980s, the quadrupole ion trap was beginning to emerge as a mass analyzer with interesting characteristics for tandem mass spectrometry. The quadrupole ion trap operates on the principle that ions can be stored within an oscillating electric field. With appropriately shaped electrodes, an oscillating electric field (usually a quadrupole field or a variation thereof) can be created that stores ions in three dimensions. Furthermore, the amplitude (strength) of the electric field can be changed so that ions of different mass-to-charge ratios are ejected from the ion trap and into a detector. In this way, the ion trap serves as a mass spectrometer.

In the latter part of the 1980s, the Organic Mass Spectrometry Group in ORNL's Chemical and Analytical Sciences Division (CASD)—one of the leading groups in the world today in this area of research—began to study electrospray

Greg Hurst (center) and Jim Stephenson (right) watch Bob Hettich prepare to use a pipette to dilute a sample of bacterial proteins in a vial. The sample will be placed in the Fourier transform ion cyclotron resonance mass spectrometer (in foreground) for analysis.

Curtis Boles; enhanced by Gail Sweeden



A colored micrograph of the phage MS2 virus, an experimentally determined structure of an MS2 viral coat protein, and a mass spectrum showing the MS2 coat protein's partial amino acid sequence are displayed here.

ionization. It is one of the important new ionization methods for mass spectrometry because it forms gaseous ions from polar and nonvolatile molecules in solution, without the addition of heat. Gary Van Berkel, Scott McLuckey, and Gary Glish, all of CASD, were the first to couple the electrospray ionization technique with the ion trap mass spectrometer.

"In electrospray, a tiny drop of protein-containing solution is injected into a thin glass needle," says CASD's Jim Stephenson. "The needle is held at a potential of several thousand volts, providing energy and adding multiple charges to the protein. As the solution leaves the needle, it evaporates, forming a fine mist of charged droplets from which ions emerge. These ions in a gaseous state are introduced to the mass spectrometer."

Another relatively new ionization method used at ORNL to analyze biological samples is matrix-assisted laser desorption (MALDI) mass spectrometry. For protein analysis, a protein is dissolved in solution and added to a matrix. The mixture is placed on a probe tip, which is illuminated by an ultraviolet light beam from a nitrogen laser. The laser beam desorbs the protein off the surface of the probe tip. Protons are then transferred between the protein and matrix, leaving the protein as a negatively charged ion.

BIOLOGICAL APPLICATIONS OF MASS SPECTROMETRY

Since coming to ORNL in 1995 as a postdoctoral researcher, Stephenson, now a staff scientist in CASD's Organic Mass Spectrometry Group, has been participating in biological research using mass spectrometry. He has been working with other CASD researchers, particularly Keiji Asano, Doug Goeringer, Bob Hettich, Greg Hurst, Rose Ramsey, Gary Van Berkel, and Michelle Buchanan, now director of CASD.

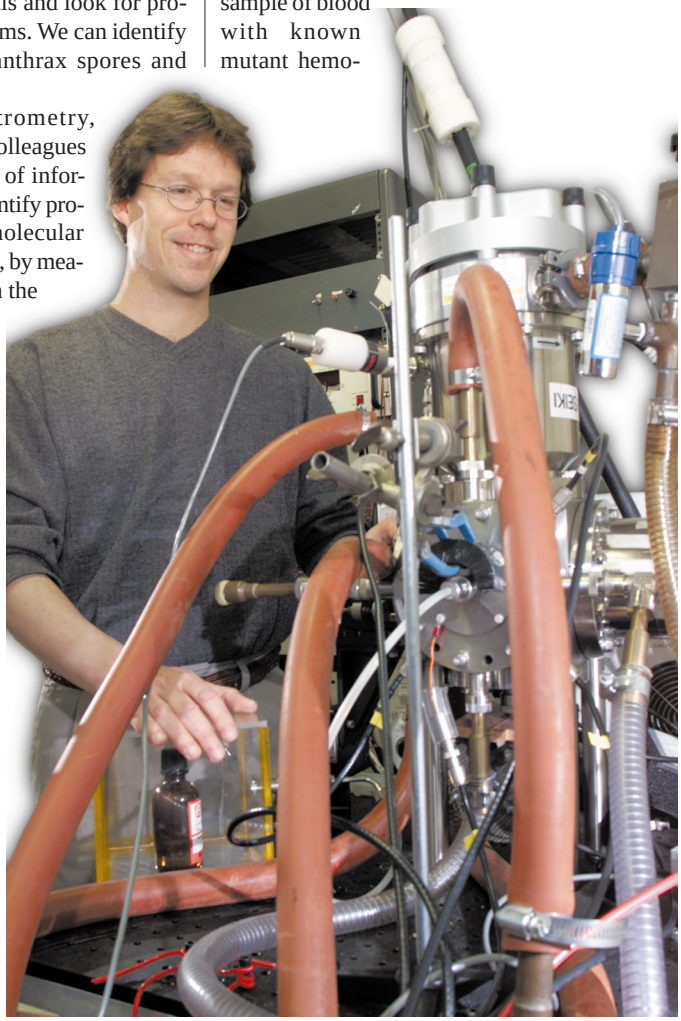
"We use mass spectrometry to identify the characteristic protein signatures of various bacteria and viruses, such as the tobacco mosaic virus," he says. "We extract proteins from *Escherichia coli* bacteria cells and look for proteins unique to these organisms. We can identify the signature proteins in anthrax spores and botulism toxins."

Using mass spectrometry, Stephenson and his ORNL colleagues obtain two important pieces of information that allow them to identify proteins. First, they get the molecular weight of the protein. Second, by measuring the distances between the peaks in the spectrum, they can figure out which amino acids are present and in what order to get the "sequence tag." They also can deduce the amino-acid sequence by the way the protein falls apart when charges and energy are added to it.

"With this information, we can go to the protein sequence database and identify the protein," Stephenson says. "For example, we might find six tobacco mosaic virus proteins in the database that have the same sequence tag but different molecular weights. Only one protein in the database matches the molecular weight and sequence tag we found."

Stephenson and his colleagues have been analyzing the disease state of mutant mice in a project with mouse geneticist Gene Rinchik of ORNL's Life Sciences Division. For this experiment, Rinchik exposes male mice to the powerful chemical mutagen ethylnitrosourea (ENU) and mates them with female mice from the same strain. As a result, many different mutations are produced in otherwise identical offspring because ENU can alter a single base pair in a gene. Stephenson and Rinchik are interested in using the mass spectrometer to detect mutant proteins that result from these inherited genetic changes. As a proof of principle, they are initially looking for mice born with a mutant hemoglobin protein.

Red blood cells carry oxygen to all parts of the body by using a protein called hemoglobin. A small sample of blood is easy to obtain from a large number of mice, so hemoglobin is a good model to test whether or not the mass spectrometer could find small inherited protein changes in the progeny of mice treated with the ENU mutagen. Rinchik was able to obtain a sample of blood with known mutant hemo-



Curtis Boles; enhanced by Gail Sweeden

Jim Stephenson adjusts an electrospray ionization ion trap mass spectrometer.

globins from Ray Popp, retired from ORNL's former Biology Division, and Stephenson found that the mass spectrometer could easily identify the changes. "Now," Stephenson says, "we can use mass spectrometry on blood samples from the offspring of mutagen-treated mice to detect new mutations. If we are successful in recognizing hemoglobin protein variants, we can use the technique to identify new, inherited mutations in any protein, increasing the efficiency and reducing the cost of finding new inherited variants in mice."

The ORNL group also uses mass spectrometry to detect post-translational modification (PTM) proteins. In a cell, a protein can be modified in different ways by other proteins or by exposure to a pollutant. For example, in a process called phosphorylation, a phosphokinase enzyme can attach a phosphate to a protein to activate it or remove a phosphate to inactivate the protein. Mass spectrometry can be used to confirm the presence or absence of a phosphate in a protein.

In addition, Stephenson and his colleagues are using electrospray ionization mass spectrometry (ESI/MS) to help determine the three-dimensional structure of proteins. Using a cross-linking chemical of a known length that attaches between two neighboring lysines in a polypeptide chain, the group can measure the molecular distance between these amino acids. This information is of value to the Computational Protein Structure Group, led by Ying Xu, which is part of LSD's Computational Biology Section. This group uses a protein-threading computer model to predict the structure of proteins. (See article starting on p. 20.)

SPEEDING UP PROTEIN IDENTIFICATION

In a project supported by internal funding from the Laboratory Directed Research and Development Program, Stephenson and his colleagues further developed ESI/MS so that it could analyze proteins much faster than the conventional method he describes below:

"Traditionally, we take a purified protein and break it into smaller pieces by digesting it with a proteolytic enzyme that selectively cleaves the protein at specified amino acid sites. The resulting products from the proteolytic digestion are then separated on a liquid chromatography column and then are transferred to the mass spectrometer directly. Sequence tags are then generated by adding energy to the protein pieces via collisions with helium atoms. These protein pieces fall apart into the individual amino acids of the protein. From these data we can figure out the sequence of the sequence tag and identify the protein by checking the sequence tag and molecular weight against protein data in the database. This approach takes about a day."

By making improvements in the ESI/MS technique and eliminating liquid chromatography, Stephenson and his associates could analyze a single protein in just a few minutes, not a day. "The problem then was how to use ESI/MS to identify many different proteins in a complex mixture at once. You could present one protein at a time to the mass spectrometer after separating the protein mixture by liquid chromatography or gel electrophoresis. But this approach is time consuming. So we designed an ion-ion reaction instrument to separate out a target protein or allow the mass spectrometer to look at one protein at a time."

Stephenson and Ben Cargile, a former graduate student at the University of Tennessee, developed a protein identification algorithm based on their discovery of how intact proteins fall apart when energy is added. This algorithm can be used to take sequence-specific data from intact proteins and identify them through a database search.

Speeding up protein identification and the collection of information on the compositional differences between normal and mutant proteins and the measurements of distances between protein building blocks could lead to the rapid development of more effective therapeutic drugs. ORNL's capabilities in combining computational analysis with data obtained by ORNL's Organic Mass Spectrometry Group could result in increased protection and improvement of human health.

Rapid Genetic Disease Screening Possible Using Laser Mass Spectrometry

In 1993, using their new ability to detect 35 to 110 base pairs in a single DNA strand, Winston Chen and his associates at ORNL were the first to demonstrate the use of laser desorption mass spectrometry (LDMS) to detect a mutant gene responsible for cystic fibrosis (CF). Their demonstration used clinical samples custom prepared by Drs. Karla Matteson and Lan-Yang Chang, both of the University of Tennessee Medical Center (UTMC).

CF is an inherited, fatal disease in which mucus buildup promotes digestive disorders and bacterial infections in the lungs. Because each person with CF is the child of parents who both carry defective forms (alleles) of a particular gene, there is interest in large-scale screening to let people know their chances of having a child with CF.

"Our technique could be used to rapidly screen many people for a specific defect in a gene on chromosome 7 that causes 70% of all CF cases," Chen says. "The defect is the absence of three base pairs of DNA in both alleles that control production of CFTR, a protein that prevents mucus buildup. CF carriers have a single defective allele that may be passed on to their offspring, and people born with CF have two defective alleles.

The ORNL group was the first to show that LDMS can diagnose a genetic disease via DNA analysis. The technique, which is not yet used commercially, screens for CF in minutes, not hours, making it 10 times faster than conventional gel electrophoresis. Also, it does not use toxic chemicals or radioactive materials, which require costly methods of disposal.

In their continuing collaboration with Drs. Matteson and Nick Potter, also of UTMC, Chen and his colleagues demonstrated in 1996 that a single-base mutation (replacement of the right base with the wrong one) can be detected by LDMS.

More recently, he and Dr. Potter showed that LDMS can detect neurodegenerative diseases that result from dynamic mutations, such as Huntingdon Disease. The normal Huntingdon gene has 9 to 24 repeats of the GAC sequence. The mutant gene causing Huntingdon disease has more than 25 repeats. The researchers showed that LDMS can identify the HD gene because of its greater molecular weight, resulting from the many additional repeats.

"Laser desorption mass spectrometry," says Chen, "is emerging as a new tool for screening populations for various genetic diseases."

Lab on a Chip Used for Protein Studies

ORNL's lab on a chip is being used commercially to identify proteins and shows promise for drug discovery and disease screening.

Ten years ago, ORNL's Mike Ramsey built the first lab on a chip. Now, improved versions of this miniature chemistry lab are being shipped all over the world. A toaster-sized, computerized device containing four matchbox-sized protein identification chips modeled after Ramsey's invention is being mass-produced by Caliper Technologies, Inc., in California. It is being sold to biotechnology firms by Agilent Technologies.

"Proteins, like DNA, pose a massive chemical measurement problem," Ramsey says. "But we have learned how to use a lab on a chip to measure molecular weights of proteins in much smaller samples and in much shorter times than are required by conventional methods."

Caliper Technologies is developing microchips for drug discovery. Such a device would help pharmaceutical firms rapidly identify compounds effective in inhibiting the activity of disease-causing proteins. "We believe that an automated device with massively parallel microfluidic chips can work with sample volumes that are 1/10,000th the volumes analyzed in conventional benchtop devices, at 10 to 100 times the speed or more," Ramsey says.

The current drug discovery chip contains four channels—thinner than human hair—that connect reservoirs, all of which are carved into a rectangular glass plate, using microfabrication technologies. A disease-related enzyme is introduced into a chip channel. Because of pressure differences, the enzyme and a modified substrate flow through the channel network, mix, and react. The reaction product is fluorescent when exposed to a laser beam. The amount of fluorescence is a measure of the reaction rate.

When a test compound is introduced into the chip through another channel, it typically reacts with the enzyme, blocking out the substrate so less of the fluorescent product is produced at a time. The reduced fluorescence signal indicates the effectiveness of the test inhibitor compound. By introducing different test compounds to the device every 5 seconds, it is possible to rapidly compare reaction rates to identify potentially effective drugs. In a recent demonstration at Caliper, nearly a million compounds were screened while using less than 1 microgram of enzyme (usually a very valuable material).

In ORNL's Chemical and Analytical Sciences Division, considerable research is being conducted by Ramsey's group on developing improved lab-on-a-chip technologies for biological, environmental, forensic, and defense applications. The lab on a chip has been honored by *R&D* magazine as one of the 40 top innovations that have come about since the magazine began its R&D 100 competition in 1963. It also has been recognized by a panel of citizens as one of the top 23 technologies developed using Department of Energy funding.

In 1998 at Caliper Technologies, Rose Ramsey (Mike's wife) and a colleague there first demonstrated that the lab on a chip could complete a two-dimensional (2D) separation of peptides in under 10 minutes. The 2D chip separation uses two types of separations to resolve the peptides. In one channel, they are separated by differences in migration speed and in another channel by differences in peptide charge and size in response to an electric field (capillary electrophoresis). By contrast, it takes 24 to 48 hours to do this separation using conventional 2D gel electrophoresis. More recently, Stephen Jacobson, Chris Culbertson, and Norbert Gottschlich contributed to a newly designed 2D chip that works even faster.

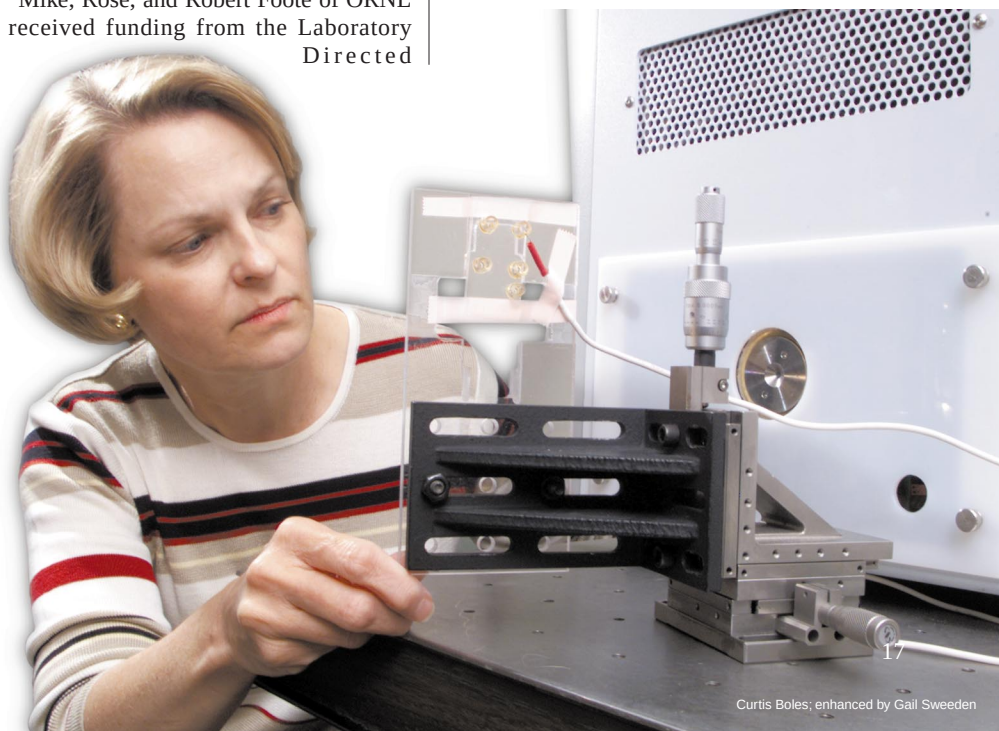
Mike, Rose, and Robert Foote of ORNL then received funding from the Laboratory Directed Research and Development Program at ORNL to analyze proteins by combining the lab on a chip with an electrospray ionization time-of-flight mass spectrometer (ESI/MS). Rose, who conceived the idea, and Iulia Lazar, a post-doctoral fellow at ORNL, showed that the procedure could be used to rapidly analyze hemoglobin, the protein that makes blood cells red and transports oxygen from the lungs to the body tissues.

In the chip, the hemoglobin from a drop of blood is reacted with the enzyme trypsin, which cleaves the blood protein at the scattered sites of two amino acids. The fragments of the hemoglobin are electrosprayed directly from the chip as ions into the mass spectrometer.

"The pattern of the fragments gives a fingerprint for the protein that can be compared to amino-acid sequences in the protein database, allowing us to identify the protein," Rose says. She used this technique to sequence over 70% of human hemoglobin from a drop of blood. She demonstrated that the technique can rapidly distinguish sickle-cell hemoglobin, whose fingerprint is different from that of normal hemoglobin. This work was recently published in *Analytical Chemistry*. This technique could also be used to rapidly screen for other hemoglobin variants.

The lab on a chip may be small but its potential for advancing medical diagnosis and treatment is quite large.

Rose Ramsey checks the lab on a chip developed for use with a time-of-flight mass spectrometer (right). Proteins can be separated and digested chemically on the chip. The protein fragments are "sprayed" as ions from the chip into the mass spectrometer for further analysis and identification.



Curtis Boles; enhanced by Gail Sweeden

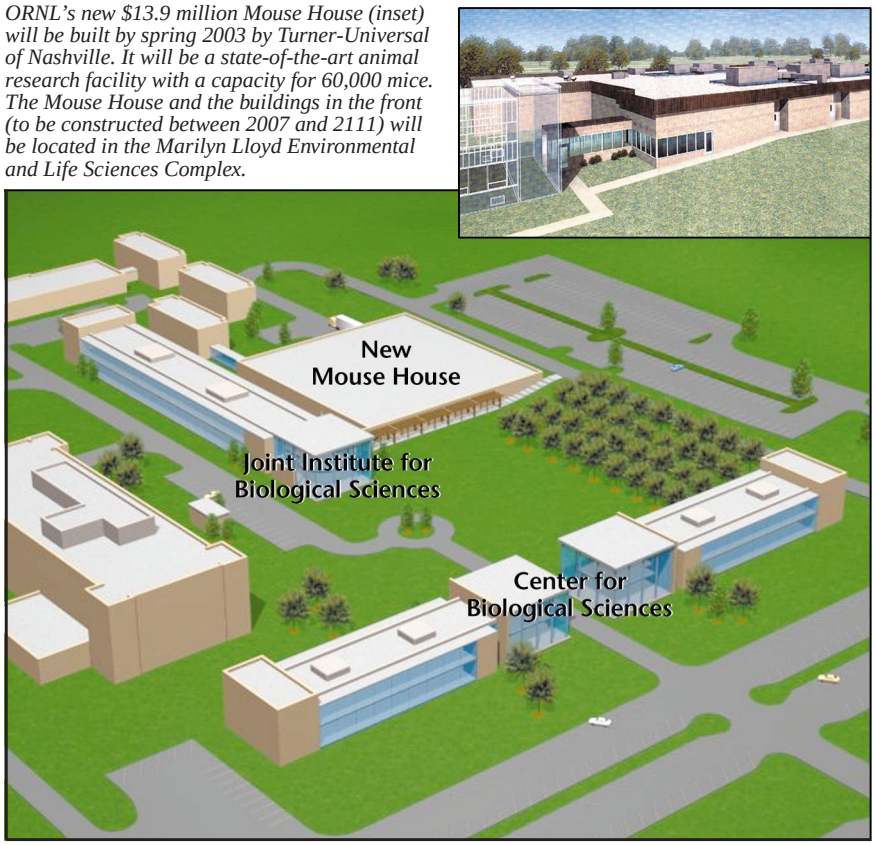
The Mouse House: From Old to New

While some ORNL mice are allowed to grow old for studies of aging, mutant mouse embryos are being frozen, awaiting birth after the new Mouse House is built.

In the spring of 2003, ORNL's new Mouse House will open, replacing the old Mouse House that dates back to 1948. The new Mouse House (officially called the Laboratory for Comparative and Functional Genomics) will hold 60,000 mice. It will be located in ORNL's new Marilyn Lloyd Environmental and Life Sciences Complex.

The mice in the old Mouse House, including those allowed to grow old there, will not be moving to the new Mouse House. The mice in the new Mouse House will indeed be new. But the genetic heritage of many of the newly bred mice will be traceable back to the old Mouse House.

Using cryopreservation techniques, some of which were pioneered by former ORNL biologist Peter Mazur and his colleagues, staff in ORNL's Life Sciences Division (LSD) are freezing and preserving eight-cell mouse embryos, sperm, and egg-containing ovaries. "We already have 450 stocks of frozen mouse embryos," says LSD's Eugene Rinchik. "Genetic characteristics and other information on each mouse stock have been entered into a computerized database (http://www.lsd.ornl.gov/htmouse). If needed for new experiments, any of these embryos can be thawed out and implanted into new female mice brought to the new Mouse House. The mice born to the surrogate mothers will have the same mutant genes as mice now in the old Mouse House."



ORNL's new $13.9 million Mouse House (inset) will be built by spring 2003 by Turner-Universal of Nashville. It will be a state-of-the-art animal research facility with a capacity for 60,000 mice. The Mouse House and the buildings in the front (to be constructed between 2007 and 2111) will be located in the Marilyn Lloyd Environmental and Life Sciences Complex.

The current ORNL Mouse House is a valuable research resource. Its colony of mice is of special interest because they represent a variety of mutations to genes, including those that cause obesity, diabetes, skin and stomach cancer, leukemia, cleft palate, polycystic kidney disease, chronic hereditary tyrosinemia, neurological dysfunctions, seizures, and a wide variety of birth defects. All of these diseases and disorders are similar to human afflictions.

For research on the effects of aging, some mutant mice are being allowed to grow old. "In our lab, we are putting four males and four females of some of our mouse families on the shelf and letting them age to 18 months, which is the equivalent of 75 years in humans," says Dabney Johnson, head of LSD's Mammalian Genetics and Genomics Section. "As these mice approach old age, we screen them for late-onset diseases.

"So far our tests of non-mutant older mice in mazes show that they have less agility with age but that activity levels and memory are the same as in younger mice. If we find abnormal older mice, we will determine which gene is associated with the later-onset abnormality."

"We are also working with our partners in the Tennessee Mouse Genome Consortium to examine older mice, which carry mutations induced at ORNL, for late-onset diseases and disorders," Rinchik says. "The diseases of interest include, but are not limited to, cancer, diabetes, obesity, memory problems similar to Alzheimer's disease, neurological problems similar to Parkinson's disease, and inner-ear or other neurological problems that cause loss of balance, lack of coordination, or even deafness.

"We are trying to find out if we can recognize mutations that cause effects in older animals but not necessarily in younger ones. Most new inherited disorders are currently identified in younger animals, and we are trying to extend that observation period to older mice carrying mutations. One goal is to determine if we can recognize good models for old-age diseases in our mutant mice."

There are various ways to produce mutations in mice. One way is to expose them to radiation and chemicals, including a powerful mutagenic chemical called ethylnitrosourea (ENU), first used by Bill and Liane Russell at ORNL. A newer way is to use DNA recombinant technology or ENU to alter genes in embryonic stem cells from mice.

In one of Rinchik's approaches to breeding mutant mice, numerous males are exposed to ENU and mated with females that have a known mutation that results in an easily identified abnormality (e.g., unusually hairy ears). Such "visible markers" allow geneticists to track the mutated chromosomes through subsequent generations. Various tests are then performed on the descendants of these mice, to determine if they have recessive or dominant mutations that result in changes in behavior, biochemistry, physiology, or anatomy, or that render them more susceptible to other diseases.

Using mice for medical research is an old idea, but the benefits of this kind of research to humans are likely to improve as resulting research at ORNL and elsewhere yields new discoveries. The new Mouse House will allow ORNL researchers to carry on the Laboratory's long tradition of mouse genetics.

Human Genome Analyzed Using Supercomputer

A computational analysis of the human genome by ORNL and UT researchers provides insights into what our genes do.

The human genome has 100,000 genes. One gene makes one protein. Humans and bacteria have entirely different genes.

These common beliefs were shattered earlier this year by the findings of the International Human Genome Sequencing Consortium, which includes the Department of Energy Joint Genome Institute (JGI) to which ORNL contributes computational analysis. On February 15, 2001, three days after a major announcement, the consortium published the paper "Initial Sequencing and Analysis of the Human Genome" in the journal *Nature*. The paper states that the human genome has "about 30,000 to 40,000 protein-coding genes, only about twice as many as in worm or fly"; each gene codes for an average of three proteins; and it is possible that hundreds of genes were transferred from bacteria to the human genome.

Ed Uberbacher, head of the Computational Biology Section in ORNL's Life Sciences Division, was one of the hundreds of contributors to this landmark paper. He and his ORNL colleagues performed computational analysis and annotation of the DNA data produced by JGI to uncover evidence of the existence of genes about which little or nothing was known.

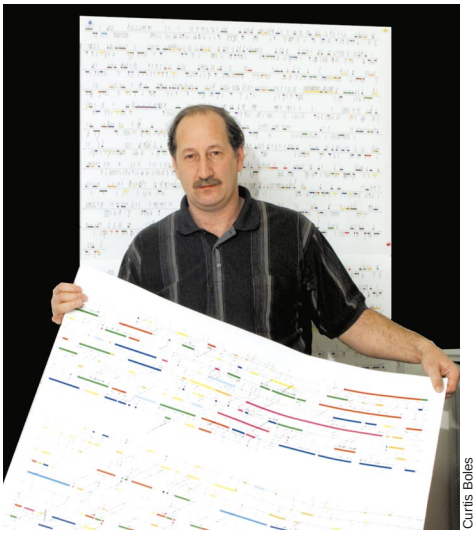
Uberbacher and his colleagues also performed an analysis of the complete, publicly available, human genome. The analysis, funded by DOE, was performed by ORNL, University of Tennessee, and University of Pennsylvania researchers using three computational methods, the GenBank database, and the IBM RS/6000 SP supercomputer at DOE's Center for Computational Sciences at ORNL. One of the computational methods used was the latest version of the Gene Recognition and Analysis Internet Link (GRAIL), which was developed by Uberbacher and others at ORNL in 1990 and rewritten as GrailEXP for parallel supercomputers.

"We have found experimental and computational evidence for some 35,000 genes," says Uberbacher. "We have also provided information on how many genes are expressed in different tissues and organs of the body. For example, we determined that more than 20,000 genes are expressed in the central nervous system. About 10% of all human genes are expressed only in the brain."

The researchers found 728 cell-signaling genes that tell cells when to divide and when to grow. They identified "zinc fingers"—regulatory proteins that bind to DNA bases composing genes to turn them on or off. These cell-signaling genes and zinc fingers are unique to the human genome.

GrailEXP located almost 2600 genes exhibiting "alternative splicing"—the ability to produce two or more proteins by combining the gene's dispersed protein-coding regions (exons) in different ways.

Each human gene contains multiple exons separated by noncoding regions called introns. Cellular machinery called a spliceosome strips out all the introns and joins the exons together. Sometimes certain exons are skipped.



Curtis Boles

Ed Uberbacher shows a printout, spread over a conference table, that depicts human chromosome 20's genes as bars of different colors and lengths. In the background are genes from microbes.

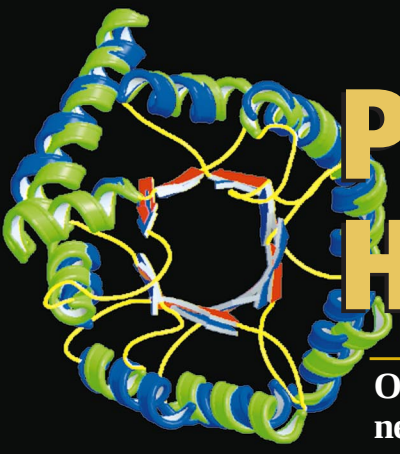
"We found a gene with 10 exons, but in different human tissues different individual exons are not read, so part of the code is left out that directs the cell to make a protein," Uberbacher says. "This gene could have 10 different protein products."

Some of the genes are known, and detailed information on their sequences is found in GenBank. Other genes are less well characterized but are similar to genes found in model organisms, such as the mouse. Still other genes are inferred based on expressed sequence tags (ESTs). An EST is a unique stretch of DNA within a coding region of a gene that can be used to identify full-length genes. ESTs were used in computational predictions to locate additional genes and predict the makeup, structure, and function of the proteins they encode.

In addition to genes, the researchers found many DNA sequences that are repeated in the human genome. This "junk" DNA may have a purpose: It lowers the probability that random mutations in DNA strike the coding sections of important genes. "Although the human and mouse genome are about the same size," Uberbacher says, "we found longer stretches of repeated DNA sequences, making up 40 to 48% of the human genome, separating clusters of genes like vast deserts between metropolitan areas."

Enhanced by Gail Sweeden

The IBM RS/6000 SP supercomputer at ORNL was used for human genome analysis.



Protein Prediction Tool Has Good Prospects

ORNL ranks high in its ability to computationally predict protein structures. The next step is to speed up predictions to facilitate the search for effective drugs.

The international competition to predict the three-dimensional (3D) structures of 43 proteins, using computational tools, was intense. Of the 123 groups competing in the fourth Critical Assessment of Techniques for Protein Structure Prediction (CASP-4) competition, which was held from June through September 2000, an ORNL group placed sixth, putting it in the top 4%. In fact, ORNL placed ahead of all other Department of Energy national laboratories in the contest.

The actual structures of the 43 target proteins were determined experimentally by nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography, and the data were unpublished at the time of the competition. "The computational groups were provided with the identity and order of amino acids making up each protein and the length of the one-dimensional amino-acid sequence," says Ying Xu, leader of the Computational Protein Structure Group in the Computational Biology Section of ORNL's Life Sciences Division. "From this information our team predicted protein structure." Other team members were Dong Xu, Oakley Crawford, and Phil LoCascio.

Motivating this competition is the search by biologists to discover the function of individual proteins that work together in "protein machines" to form an organism or keep it alive. They also want to understand how these functions are performed at the molecular level. Proteins often do their work by docking with another protein. Because the function of a protein is related to its shape, it is essential to learn the 3D structure of each protein. Using the details of a protein's shape, a chemical compound can be custom designed to fit precisely in the protein, like a hand in a glove, blocking or enhancing the protein's activity. In this way, a highly effective drug with no side effects could be created for each individual.

"The demand for rapid protein structure determination will grow drastically because information that could be used for rational drug design is becoming available rapidly," Xu says. "Traditional experimental methods for determining protein structure may not be able to keep up with the pace at which amino-acid sequences are being generated. Computational techniques, in conjunction with experimental methods, could more rapidly determine protein structures on a genome scale."

For the CASP-4 competition, the ORNL researchers used a computer package that they developed and continue to improve. It is called the Protein Structure Prediction and Evaluation Computer Toolkit (PROSPECT) and is one of only a few dozen protein-threading computer programs in the world.

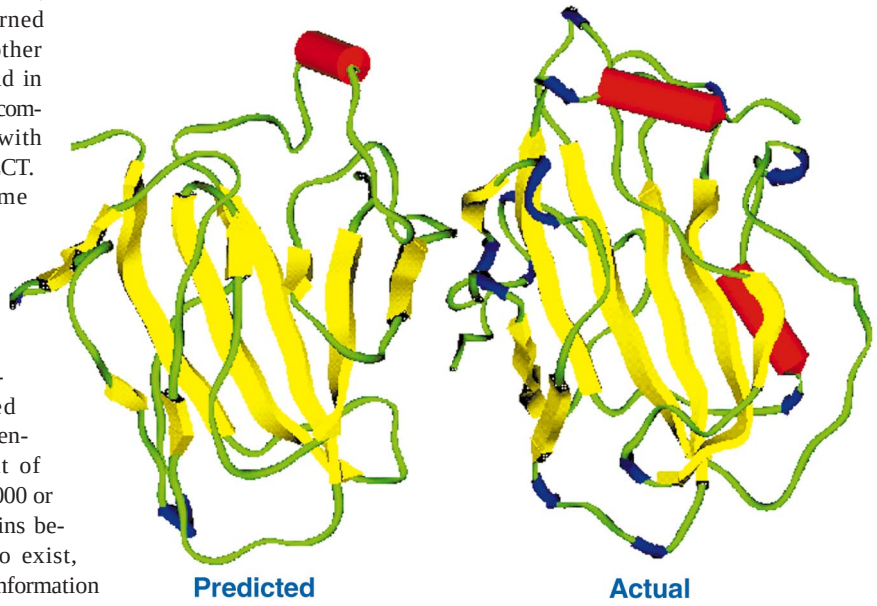
"In the CASP competition, you get a 0 if you fail to identify the correct structural template," Xu says. "You get a 4 if your alignment between the target protein and the template is perfect. You get scores of 1 to 3 depending on how close you are to being correct. The scores are added up for all 43 proteins. We recognized two-thirds of the correct templates, which is the most among all the competing teams, and one-third of our alignments were off."

Recently, the ORNL team attended a conference in Asilomar, California, and learned how other teams did in CASP-4 compared with PROSPECT.

"Some 10,000 protein structures have been determined experimentally out of the 100,000 or so proteins believed to exist, and the information is stored in the Protein Data Bank," Xu says. "To keep up with the production rate at which protein sequences are being generated by the genome projects, computational methods are clearly needed. Structure predictions have been made by the conventional *ab initio* technique in which a supercomputer is used to predict how an amino-acid chain can fold itself into a final shape based on first principles of physics and chemistry. Unfortunately, it takes weeks to months to predict the structure of even the smallest protein using this approach and the prediction reliability is poor."

The ORNL group uses template-based methods of protein structure prediction. These methods rely on experimentally determined 3D structures in the Protein Data Bank. The ORNL group uses PROSPECT to do "protein threading," in which a string of amino acids is computationally aligned along different protein templates—like an embroidery thread drawn through a printed design—to determine which template gives the best fit. In a perfect alignment, the amino-acid atoms are at their preferred lowest energy levels and are compatible with neighboring atoms and the protein's environment.

High-priced bows for birthday presents? No, these images show the actual structures of several proteins and their predicted structures using PROSPECT software at ORNL.



The ORNL group also uses homology modeling to fine-tune the predicted structure. In this technique, if two amino-acid sequences are similar and one sequence has a known structure, researchers can use this information to help determine the structure of the unknown protein sequence. By calculating the detailed forces between atoms and adjusting the final predicted structure to minimize the atoms' energies, the researchers computationally tweak the predicted structure of the target protein to make it energetically more favorable.

"It is believed that about 1000 unique protein structural folds exist in nature and that many proteins share each of these unique structural folds," Xu says. "Some 600 unique protein structures have been determined experimentally. Once the 1000 unique structural folds are determined by NMR and X-ray crystallography, the rest of the 100,000 protein structures can be accurately modeled computationally."

Xu's group, which has four staff researchers and two postdoctoral scientists, is involved in the National Institutes of Health's Structural Genome Initiative, which is dedicated to finding the structures of 100,000 human proteins. As part of this effort, NIH has funded seven pilot centers for experimentally determining protein structures. They include centers at DOE's Argonne, Brookhaven, Lawrence Berkeley, and Los Alamos national laboratories. At Lawrence Berkeley, David Eisenberg, a pioneer in protein threading, is trying to determine the structures of proteins in the genome of the rod-shaped bacterium that causes tuberculosis.

"A new trend in structure prediction is the incorporation of partial experimental data as constraints in the computation process, to make structure prediction closer in accuracy to the experimental structures," Xu says. "PROSPECT is ideally suited for incorporating data from local researchers and from these pilot centers. The data include measurements of distances between amino acids and information on which amino acids tend to be found on the surface of a protein and which don't."

Recently, the ORNL group modeled a protein complex using PROSPECT and experimental data provided by Cynthia Peterson, a University of Tennessee researcher who has identified a number of disulfide bonds between amino acids in certain parts of the protein. PROSPECT is being used to incorporate experimental data provided by Greg Hurst and Jim Stephenson of ORNL's Organic Mass Spectrometry Group. They are using electrospray ionization ion trap mass spectrometry and a cross-linking chemical to determine the distances between two amino acids—both lysines—in a protein. Their initial studies found that the lysines linked by this chemical of a known length are 4 angstroms apart. (See article starting on p. 14.)

Because NMR data provide distances between amino acids in a protein, the ORNL group will gladly accept partial NMR data from the pilot centers, which otherwise will not be used because the information is insufficient to determine a whole protein structure. "This amount of data is good enough to help PROSPECT reliably predict a protein structure," Xu says. He notes that the level of confidence, or uncertainty, in knowing a protein structure with complete accuracy is within 1 to 1.5 angstroms for X-ray crystallography, within 2 to 2.5 angstroms for NMR, and within 4 angstroms for computer modeling.

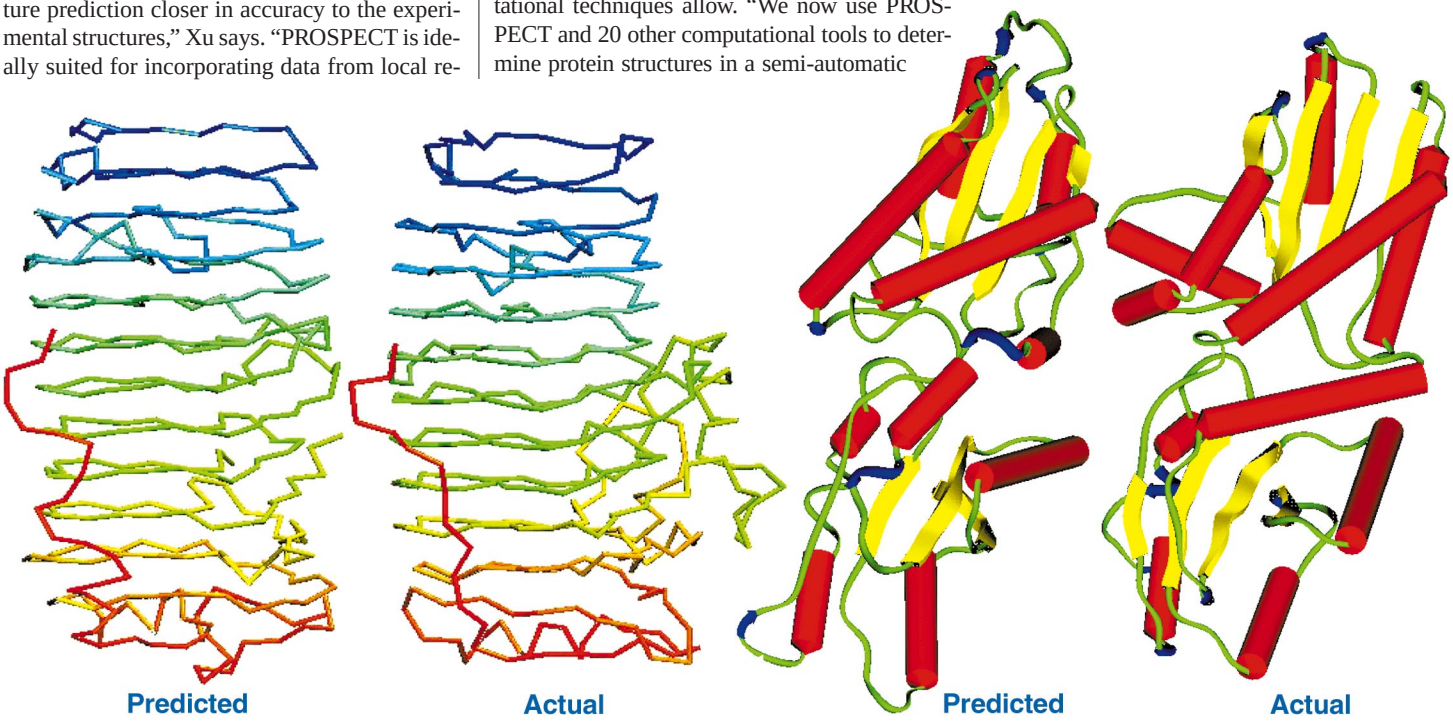
The ORNL group is focused not only on predicting protein structures more accurately but also on doing it much faster than current computational techniques allow. "We now use PROSPECT and 20 other computational tools to determine protein structures in a semi-automatic fashion," Xu says. "Using the IBM RS/6000 SP supercomputer at DOE's Center for Computational Sciences at ORNL, we can now thread 100 or more proteins a day against 2000 possible template structures. We are seeking funding to develop software to build an expert system and an automated protein structure pipeline to run on the IBM supercomputer. The expert system will mimic the human decision-making process to automate the computational tools. Our goal is to predict about 100 protein structures a day.

"If the proposal is funded, our first project will be to predict the structure of proteins of the *Prochlorococcus marinus* genome, a bacterium with about 1600 genes. We hope to show that we can predict these protein structures."

In the next three years, the ORNL group expects to do some computer simulation that will be of interest to the pharmaceutical industry. Their work could enable more rapid design of drugs that are safe and effective.

"To do this, you have to know whether a ligand, which is a group of molecules typical of a new drug, will dock with a particular protein to inhibit or stimulate its activity," Xu says. "We will be doing computer modeling to determine whether and how a ligand binds with various proteins to cause a healing or harmful effect."

PROSPECT is a copyrighted computer program. It is being used by over 20 academic organizations, including MIT, Columbia University, the University of Michigan, and the University of Texas. Millennium Pharmaceuticals is interested in licensing the program. ORNL's Technology Transfer and Economic Development Directorate seeks to license this computer toolkit for commercial use because its recent successes suggest it has very good prospects.



Microbe Probe: Studying Bacterial Genomes

ORNL researchers are using gene chips, mass spectrometry, and computational analysis to understand what microbe genes do.

In 1956 a new microbe was discovered in a can of spoiled ground beef thought to have been sterilized by radiation. The bacterium was named *Deinococcus radiodurans* because it can survive doses of radiation thousands of times higher than would kill most organisms, including humans. The Department of Energy is interested in studying *D. radiodurans* because of its efficiency in repairing its own radiation damage and because it can reduce, or add electrons to, uranium, iron, chromium, and technetium. Thus, the bacterium might be a strong candidate for remediating mixed wastes—combinations of radioactive materials and toxic metals—found at DOE sites. For example, it could be used to convert uranium in storage ponds from a soluble to an insoluble form so that it sinks into the sediments instead of remaining in water that might flow off-site.

ORNL researchers are interested in knowing more about the bacterium's genes that repair radiation-induced damage to its DNA. Knowing how these DNA repair genes work together in a network could help scientists better understand how living organisms evolved and how microbes that may have lived on Mars or other extraterrestrial sites tolerated extreme radiation environments. This knowledge may lead to gene therapy in which DNA repair genes are inserted into the body to prevent radiation-induced cancer or to treat the disease. DNA repair genes discovered in *D. radiodurans* could be transferred to other bacteria that reduce toxic metals, to make them better able to survive radiation as they treat mixed wastes.

Thanks to advanced technologies such as microarrays and mass spectrometry and internal funding from the Laboratory Directed Research and Development Program, six ORNL researchers are improving the understanding of *D. radiodurans*. One of their goals is to identify which genes in *D. radiodurans* are expressed during exposure to high levels of radiation, resulting in the production of DNA repair proteins. Another goal is to discover the regulatory genes that control other genes involved in radiation resistance. The researchers are Jizhong Zhou, Bob Burlage, and postdoctoral scientists Dorothea Thompson and Alex Beliaev, all in the Environmental Sciences Division (ESD); Bob Hettich of the Chemical and Analytical Sciences Division, and Randy Hobbs of the Research Reactors Division.

To identify the *D. radiodurans* genes involved in DNA repair, Zhou and his ESD colleagues use microarrays, also known as gene chips. (See p. 5 for an explanation of how gene chips work.) They will place the entire complement of genes from *D. radiodurans* on the same gene chip and then use the chip to identify differences in global gene expression between *D. radiodurans* cells exposed to high radiation and those not so exposed. (The expressed genes produce messenger RNAs that can be detected.) Genes expressed only in the irradiated bacteria are likely to play specific roles in DNA repair.

The ORNL researchers will then develop genetic vectors for generating deletion mutants—bacteria from which key regulatory genes involved in radiation resistance are removed. These regulatory genes code for proteins that may activate other genes involved in DNA repair. If the mutant bacteria die when exposed to high radiation doses, then the results of this experiment suggest that the deleted genes play a key role in radiation resistance.

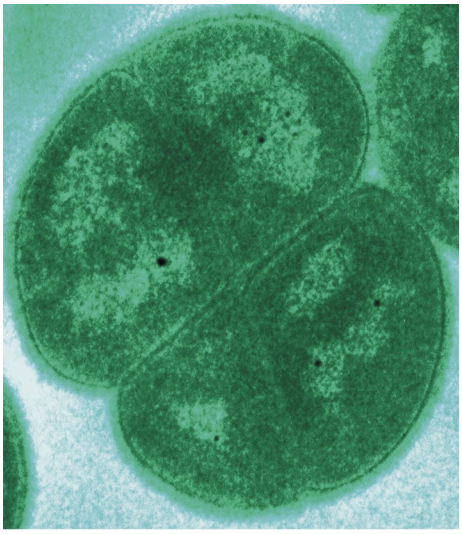
To understand the cellular response of the *D. radiodurans* bacteria to radiation, it is necessary to identify protein expression profiles and protein-DNA interactions. Zhou and his colleagues will culture the bacteria under both non-irradiated and irradiated conditions. The bacteria cells then will be lysed and the proteins will be extracted, purified, and finally characterized by Hettich, using high-resolution electrospray mass spectrometry.

"The capability of this technique for accurate measurement of the protein masses, as well as identification of sequence tags (short amino-acid sections from the protein), provides unparalleled information for unambiguous protein identification," Hettich says. Such information will be compared to that obtained from more conventional two-dimensional gel electrophoresis. The data should provide insights about how the microbe's proteome—its total protein profile—is altered in response to DNA damage, providing important details about gene activity.

"Knowing when and where a gene is expressed often provides a strong clue to its biological function," Zhou says. "Gene expression patterns and protein profiles in a cell can provide detailed information about its state." The researchers' goal is to identify pathways of interactions between DNA and proteins that result in DNA repair of radiation damage.

Another group of researchers interested in *D. radiodurans* is the Computational Biology Section of ORNL's Life Sciences Division. Using gene-recognition algorithms and other tools, Frank Larimer and his colleagues analyze completed DNA sequences of microbes and compare the identity and order of these DNA bases with known sequences in databases. They identify the known genes contained in these microbial genome sequences and predict the amino-acid sequences, molecular sizes, and possible functions of the proteins these genes encode. They also predict the makeup and functions of unknown genes and proteins.

"We have studied the completed sequence of *D. radiodurans* and can recognize many of its DNA repair genes," Larimer says. "We inferred the function of their proteins based on their



Deinococcus radiodurans *is a microbe of particular interest to DOE because of its ability to thrive in radiation levels thousands of times higher than those that would kill most organisms, including humans. It also may prove useful in bioremediation of toxic waste. Shown here is a group of four microbe cells.* Courtesy of Uniformed Services University of the Health Sciences.

False-color image (opposite page) of a microarray used to analyze messenger RNA levels in Shewanella oneidensis MR-1 *cells under different growth conditions. Two-color fluorescence detection allows two different biological samples to be analyzed simultaneously on a single array. Genes highly expressed under aerobic growth conditions are shown in green (Cy3), whereas genes expressed specifically under anaerobic conditions in the presence of iron [Fe(III)] appear as red (Cy5) spots.*

degree of similarity with proteins in the database associated with repair genes.

"The majority view is that there is a 'eureka' DNA repair gene to explain why *D. radiodurans* is so resistant to high levels of radiation. I take the minority view that, although the bacterium's DNA repair genes look like the repair genes in other organisms, *D. radiodurans* does the repair job more efficiently. It is extremely adept at aligning and recombining its broken chromosome parts in the proper order."

D. radiodurans is one of the millions of microbes that have been evolving on the earth over the past 3.8 billion years. Many of the 1% of the microbes known to humans survive and thrive in extremes of radiation, heat, cold, pressure, salinity, and acidity, often where no other life forms could exist. DOE is interested in identifying and harnessing the talents of some of these microbes for cleaning up hazardous wastes, producing energy (e.g., methane), and sequestering carbon.

Since its establishment in 1994, DOE's Microbial Genome Program has focused on determining the sequences of the genomes of selected bacteria and other microbes that do not cause disease. The ORNL group provides computational and other bioinformatic analysis of microbial sequences obtained by DOE's Joint Genome Institute (Lawrence Berkeley, Lawrence Livermore, and Los Alamos national laboratories) at their production facility in Walnut Creek, California. Larimer and his colleagues "annotate" these microbial DNA sequences and others provided by academic groups. That is, they add to the database "biological footnotes" about the genes, the coding and noncoding regions of the genes, and the possible structure and function of the proteins encoded by individual genes. For each microbe, the researchers translate roughly 2 billion DNA bases into meaningful information.

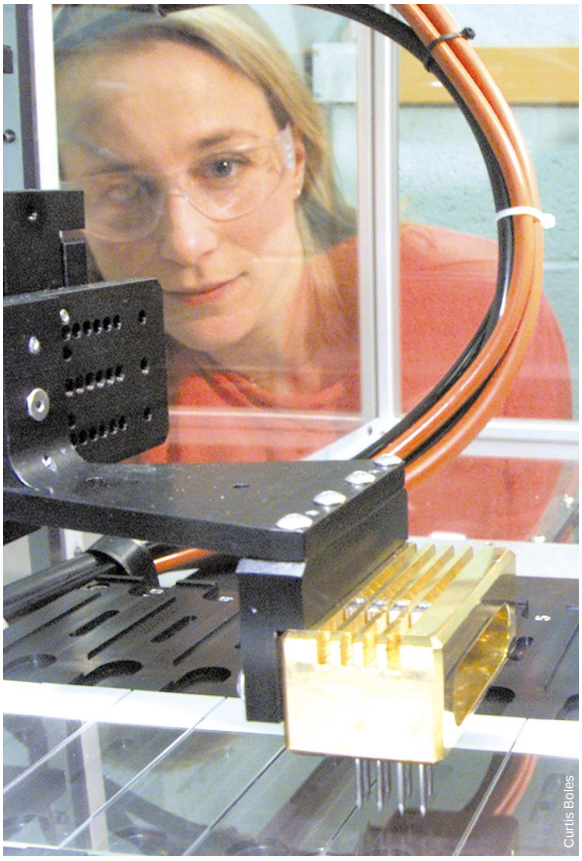
Using the IBM RS/6000 SP supercomputer at DOE's Center for Computational Sciences at ORNL, the group runs GrailEXP, the version of the ORNL-developed Gene Recognition and Analysis Internet Link written for parallel supercomputers, along with other gene modeling programs to determine the correct DNA sequences in the microbial genes that code for proteins. "Ninety percent of bacterial DNA encodes protein compared with less than 2% for humans," Larimer says. "Microbes are very efficient this way. So we try to distinguish between what portions of the bacterial sequence are genes and what is a statistical anomaly."

The typical bacterium has 2000 genes (each microbial gene has about 900 bases). The ORNL group tries to find gene signatures in bacteria—unique combinations of genes that indicate each bacterium's identity.

"What is startling is that we can assign functions for 50% of the genes in a microbial gene, which might be 1000 genes. The other half of the genes have no known function. Half of these genes are unique—we've never seen them before. There is a lot of genetic biodiversity in microbes. Microbes acquire sets of genes from each other and throw away large blocks of genes they don't need as they evolve."

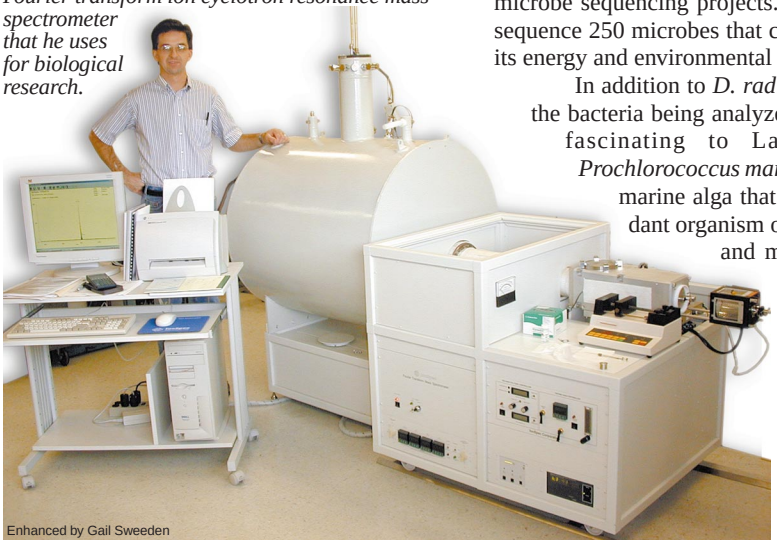
To date, 50 microbial genes have been completely sequenced. Larimer's group has been annotating them and working with 200 ongoing microbe sequencing projects. DOE's goal is to sequence 250 microbes that could be useful for its energy and environmental missions.

In addition to *D. radiodurans*, some of the bacteria being analyzed are particularly fascinating to Larimer. One is *Prochlorococcus marinus*, a blue-green marine alga that is the most abundant organism on earth in number and mass. These ocean algae, whose individual cells are each less than 1 micron in diameter, together fix more carbon dioxide than all of the terrestrial biosphere. *Prochlorococcus marinus* is found 100 meters deep in the ocean yet it absorbs sufficient blue-green photons penetrating to that depth to get enough energy to fix carbon.



Julia Stair watches the Cartesian Technologies gene-chip maker go through a test run. It is used for studies of gene expression in bacteria of environmental significance. Stair, a post-masters researcher with Oak Ridge Associated Universities, prepares microarrays for studies of D. radiodurans *bacteria.*

Bob Hettich shows the new high-performance Fourier transform ion cyclotron resonance mass spectrometer that he uses for biological research.



Another interesting microbe that fixes carbon dioxide is *Nitrosomonas europa*, which gets its energy not from sunlight but from oxidizing ammonia. Using ammonia fertilizer, it is responsible for putting nitrogen into the soil, making it fertile for plants.

Zhou is fascinated by gene-chip and mass spectrometry studies of the metal-reducing bacterium *Shewanella oneidensis*. "We compared gene expression patterns in *Escherichia coli* bacteria and a mutant of *Shewanella oneidensis* that we generated," Zhou says. "We found that the functions of about 75% of their genes are similar. But there are differences. In *E. coli* one gene activates nitrate reduction, but the counterpart gene in mutant *Shewanella oneidensi* represses nitrate reduction.

"Using bioinformatics, scientists often assign gene functions based on the similarity of sequences of two different bacteria. This approach may not correctly define all gene functions. To get the complete picture, bioinformatics must be supplemented by experimental approaches that analyze expression in microbes, using microarrays and mass spectrometry."

SNS and Biological Research

Three world-class biological instruments are being designed for the Spallation Neutron Source. They will help biologists determine the atomic-level structure of proteins and other signaling compounds that allow cells to communicate and coordinate activities across an organism. The research could lead to safer, more effective drugs.

Some drugs that combat AIDS—the deadly disease caused by the human immunodeficiency virus (HIV) that kills or damages the body's immune system cells so the body can't fight infections—have been discovered by accident. For example, the protease inhibitors used to treat AIDS were originally tested as medications for reducing blood pressure. These drugs resemble pieces of the protein chain that the HIV enzyme, called protease, normally cuts. By gumming up the protease scissors, HIV protease inhibitors prevent the protease from cutting long polyprotein chains into the shorter structural proteins HIV needs to assemble a cocoon of protection around its RNA genome. Although the virus can still invade other cells, without the protective cocoon, the virus genome is exposed to host enzymes, which easily destroy the invading viral RNA and prevent further replication.

Finding the right chemical structures to halt the replication of HIV and then bringing the drugs to market can take years and cost millions of dollars. As a result, millions of people who need these medications to survive cannot afford them. However, there is cause for hope. Recent work in biological neutron crystallography and new biological instruments planned for the Spallation Neutron Source (SNS) at ORNL could provide rapid insights into macromolecular structures that could lead to safer, more effective, and less expensive drugs.

Determination of protein structures by neutron diffraction will be useful for drug design. HIV protease, a target of AIDS drug designers, belongs to the general class of protein-digesting enzymes called aspartic proteases. These enzymes (including the stomach's pepsin) are named for aspartic acid, which is present at their active sites. A solvent molecule bound tightly to aspartate carboxyl groups is presumed to take part in the catalytic mechanism that enables the enzyme to break a protein chain. The best currently accepted mechanisms are largely based on X-ray images of inhibitor structures, but the active-site hydrogen atoms cannot be definitively located by current X-ray analyses, leaving an incomplete picture of enzyme catalysis.

Jonathan Cooper of the University of Southampton in England and Dean Myles of the European Molecular Biology Laboratory Outstation in Grenoble, France, published a report in 2000 on the neutron diffraction structure of the fungal aspartic protease endothiapepsin. This work represents the largest protein solved by neutron diffraction methods to date. The success of neutron diffraction in determining the positions of catalytic hydrogens reveals a route to the development of more effective inhibitors to aspartic proteases. It also suggests that this method could play a significant role in rational drug design.

Similar and better atomic-resolution images of biological substances are anticipated from the SNS after its completion in 2006 when it is operating at 2 megawatts. That's because the SNS will have a much higher neutron flux; it will offer 10 times the number of neutrons now available at any existing neutron research facility. The intrinsic time structure of the pulsed neutron source makes SNS ideal for atomic-resolution protein crystallography studies.

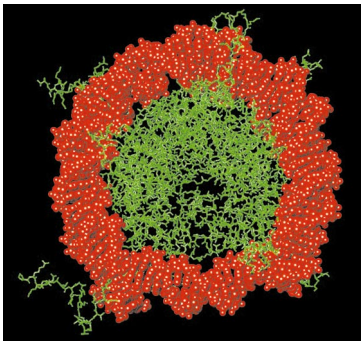
"If the SNS has an instrument that can do high-resolution crystallography on protein crystals with 100-angstrom repeats," says Gerard Bunick of ORNL's Life Sciences Division (LSD), "we could contribute crucial molecular information that could lead to more effective drugs against HIV, for example.

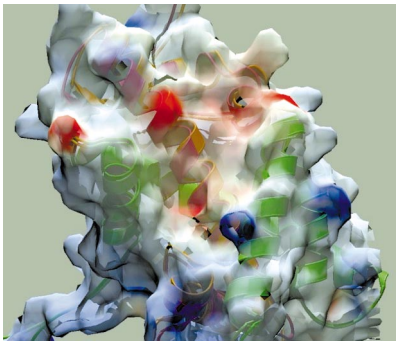
"X rays give excellent high-resolution images of heavy atoms in protein crystals, but neutrons also see lighter atoms such as hydrogen, which makes up half of all the atoms in proteins. Only one in 100 proteins crystallizes well enough to get the resolution needed to see hydrogen atoms using the bright X rays from synchrotrons, according to a survey of structures in the Protein Data Bank."

Only about a dozen protein structures have been solved using neutron diffraction. But that could change with the opening in 2001 of the new Protein Crystal Station at the Los Alamos Neutron Scattering Center and the later operation of the SNS. Bunick's colleague Leif Hanson from the University of Tennessee says that within 10 years, 50 to 100 protein structures will be determined annually using neutron diffraction. Neutron diffraction studies of protein structures are also being facilitated by the availability of improved neutron detectors that speed up data collection and the ability to grow larger crystals of proteins.

"It is now easier to grow larger protein crystals in space and on earth," says Bunick, who set up crystal-growing experiments that were run on the U.S. Space Shuttle Columbia in 1995 and on the Russian Mir space station in the late 1990s. Nearly perfect crystals of nucleosomes were grown in the microgravity environment of these vehicles orbiting around the earth. Nucleosomes are the building blocks of chromosomes; they each consist of a core of histone proteins around which approximately two turns of double-stranded DNA are wrapped.

ORNL's Gerry Bunick and his colleagues plan to use the SNS to better understand interactions between DNA and proteins and between proteins. Below are color renderings of a nucleosome core particle (DNA in red, proteins in green) and detail from an exposed surface of the protein core. The ribbon (right) depicts the superposition of symmetry-related proteins in the structure. The transparent surface shows basic and acidic domains in blue and red respectively. The image shows that the acidic domains are where the symmetry-related proteins are not identical.





Artist's conception (opposite page) of the Spallation Neutron Source. Rendering by John Jordan.

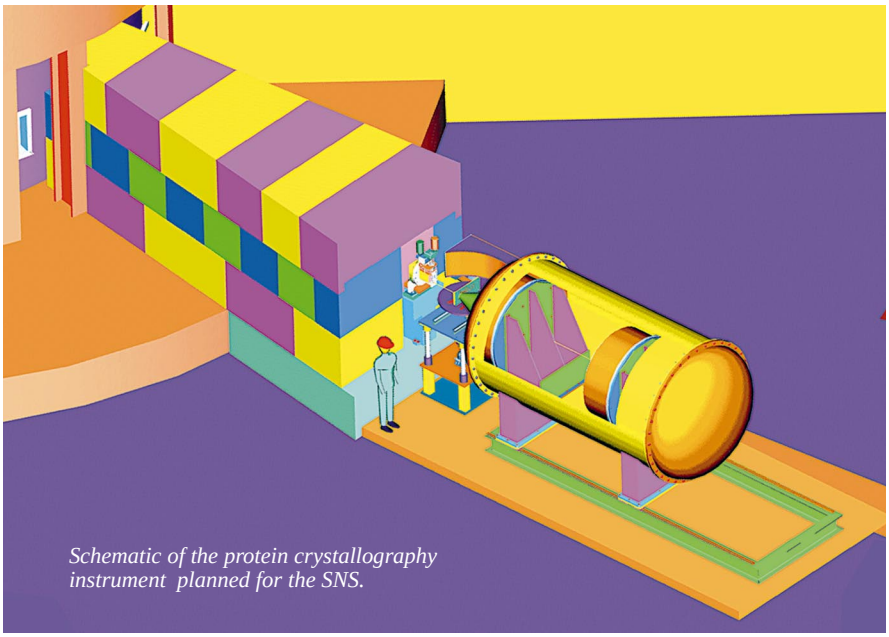
"We have proposed to NASA to grow large protein crystals for neutron studies in the microgravity environment of the International Space Station," Bunick says. "And better ways to grow large protein crystals on earth have been developed as a result of microgravity crystal growth research sponsored by NASA."

Bunick, Hanson, and Joel Harp, all of LSD; Chris Dealwis of the University of Tennessee; and Jinkui Zhao of the SNS organized and co-hosted an SNS workshop December 18, 2000, in Knoxville. They were pleased that the workshop participants recommended that two instruments be designed to do high-resolution protein crystallography studies using SNS neutrons. The first instrument would be used for high-resolution protein crystallography on crystals with up to 100-angstrom (Å) repeating motifs. The second instrument, if funded, would be located in the long-wavelength target station. It would be used to study protein complexes with 200–250 Å repeating motifs in the crystals, other large macromolecular complexes at lower resolution, and biological membrane systems.

According to Thom Mason, ORNL's associate laboratory director for the SNS, biologists are interacting with engineers to design 3 of the 12 world-class scientific instruments planned for the SNS. These instruments will complement the capabilities of DOE's Center for Structural and Molecular Biology at ORNL. They will help biologists determine the atomic-level structure of proteins, amino acids, hormones, peptides, and other signaling compounds that allow cells to communicate with each other and coordinate their activities across the organism.

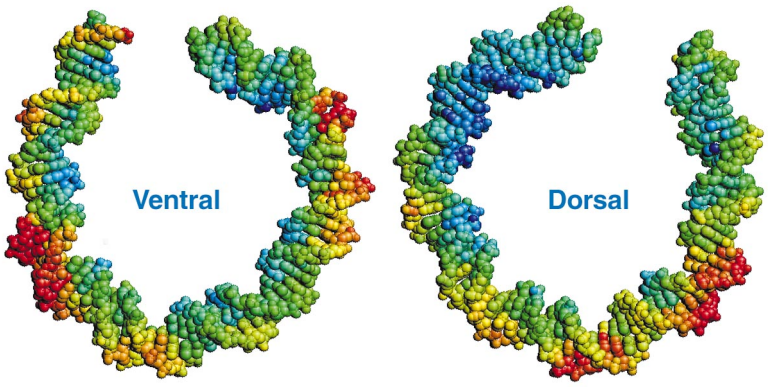
Besides the protein crystallography instrument, biologists plan to use a liquids reflectometer to study changes in surfaces, interfaces, and layered structures in biological materials. "It could help scientists study how proteins affect the structure of membranes in cells," Mason says.

A small-angle neutron scattering (SANS) instrument is also planned. It will "see" a target with a length scale ranging from thousands of angstroms to less than an angstrom. "It's a little like having a camera with a zoom lens," Mason says. "You zoom in on the subject to see the small details. Then you pull the zoom lens back to get the big picture, but you can't see the smallest details. With the SANS at the SNS, we will be able to zoom in on the protein's details at the interatomic level, what we call large Q. Then we can pull back and see the whole protein, what we call small Q. We can span this wide range in a single experiment, something that can't be done easily with X-ray crystallography or existing neutron instruments."



Schematic of the protein crystallography instrument planned for the SNS.

The two turns of DNA from the nucleosome core particle, separated at the dyad axis for clarity, rendered as a CPK model. The colors depict relative mobility of the atoms, with blue and red being the least and most mobile respectively. Neutron diffraction data from the SNS will help reveal how water molecules affect the structure of DNA and its interactions with proteins.



Zhao, a neutron scientist who does biological studies at the High Flux Isotope Reactor (HFIR), says that the biological SANS instruments that will be installed at both HFIR and the SNS will be used to study proteins and other biological molecules in their natural solution. Both instruments will be superbly suited to study the shape, conformational changes, and dynamics of proteins. "The difference will be that the HFIR SANS will look at large length scales, whereas the SNS SANS will look at a range of length scales simultaneously, such as protein-membrane interactions, with proteins on the large-length-scale side and membranes on the small-scale side."

"To study the interactions of proteins and other biological molecules in solution, we must use contrast matching," Zhao says. In this technique, some hydrogen atoms in the sample are replaced with heavier hydrogen, or deuterium, atoms. Deuterium and hydrogen atoms scatter neutrons very differently. Changing the ratios of hydrogen and deuterium in the solvent water mixture changes the visibility of proteins to neutrons. It's like adding red dye to a glass of water containing red and yellow balls so that all you see are the yellow balls."

"Scientists mask out the portion of the protein they are not interested in and make highly visible the active part of the protein that does interest them," says Mason. "The ratios of hydrogen and deuterium in water mixtures can be changed in various samples to meet researcher needs."

The SNS can also be used for inelastic neutron-scattering experiments for the study of protein dynamics. Inelastic scattering of neutrons results from an "inelastic collision" in which the total kinetic energy of neutrons colliding with target atoms is not the same after the collision as before.

"Measurements of neutron-scattering energies will provide information on the collective motion of proteins, which can be correlated with protein function," Zhao says. "For example, when a catalytic protein is active, its motion is random. But when its motion becomes harmonic at low temperature, its catalytic function stops."

Bunick and his colleagues could use an SNS protein crystallography instrument to help them better understand the origin of Rett Syndrome (RS), a debilitating neurodevelopmental disorder that causes loss of speech and profound mental retardation in young girls. Studies at the SNS might improve understanding of the structure and interactions of native and mutant DNA-binding proteins, enabling the production of drugs to reverse the deleterious effects of the defective protein.

When research comes to life at the SNS, we can expect some early discoveries to jump-start the design of drugs to improve human life.

Accessing Information on the Human Genome Project

DOE's Human Genome Management Information System at ORNL communicates Human Genome Project information to the public.

If you want news on the new biology, whom should you call? Try the dedicated ORNL group that maintains the Human Genome Management Information System (HGMIS) for the Department of Energy. This group issues a number of publications (including *Human Genome News*, published about twice a year) and maintains a popular Web site (www.ornl.gov/hgmis/). The HGMIS folks have become an important source of information on the human, animal, plant, and microbial genomes for the news media, the trade press, biology researchers, and teachers and students.

The HGMIS is DOE's educational outreach arm of the Human Genome Project, which is supported by both DOE and the National Institutes of Health. The HGMIS Group, led by Betty Mansfield of ORNL's Life Sciences Division since 1989, is responsible for providing information about the project, its progress, its applications, and its impacts on society. In 1997, on behalf of the group, Mansfield received an Exceptional Service Award for Exploring Genomes, presented at the 50th Anniversary Symposium of DOE's Biological and Environmental Research (BER) Program; she was recognized "as founding and managing editor of *Human Genome News* and for outstanding success in communicating scientific information to the U.S. and international communities about the Department's BER Program."

In 2000 the HGMIS Group received inquiries from ABC's *Who Wants to be a Millionaire*?" and NBC's *Jeopardy* concerning the accuracy of answers to questions on human genetics and the human genome. The ORNL group has provided information and graphics to a number of U.S. and foreign news outlets, including CNN, NBC, ABC's *Good Morning America*, and ABC London.

The group also has provided answers to questions from reporters from the print media, including the *Boston Globe*, *Chicago Tribune*, *Los Angeles Times*, *Wall Street Journal* (interactive online edition), *Science Magazine*, and *Wired Magazine*. Marissa Mills, a trained journalist, marketing expert, and Web content developer in the HGMIS Group, has developed press kits and answered numerous electronic mail questions from the media related to human genome news events.

Mills provided assistance to two representatives of the Native American community seeking information about tracing Native American ancestry through DNA. She also has created and presented programs on careers in genetics for women, African Americans, and residents of rural Appalachian communities.

Denise Casey, a trained biologist and versatile writer in the HGMIS Group, was guest editor of and wrote an article for the special genes and justice issue of *Judicature*, a magazine for judges, published in November-December 1999. Recently, she and group member Judy Wyrick created a human chromosome landmarks poster, jointly sponsored by DOE and Qiagen, a company supplying reagents for genome research.

Other members of the HGMIS Group are Anne Adamson, Laura Yust, and Web architect Sheryl Martin. A creative lot, they provide valuable editing, publication layout, and Web design expertise for the group. In addition to being the group's primary editor, text manager, and facilitator, Adamson recently co-authored texts on the DOE Ethical, Legal, and Social Issues program and contributed to an invited article with program manager Dan Drell on how genomics may affect the outcome of medicine 20 years from now. Yust, a graduate student, does Web work, coordinates mailing and outreach, and maintains the extensive database of 15,000 newsletter subscribers.

The group is particularly proud of their Web site. The HGMIS Web site statistics show an average of 2.6 million user sessions per year; the average length of a user session is 12.5 minutes. The site has 9 million text file hits per year, and many more hits when graphic files are counted. The site has specialized pages, including pages on the benefits of genome research; explorations into the ethical, legal, and social issues surrounding the availability of personal genetic data; and primers that are widely used by researchers from other disciplines who wish to contribute to genomics research, teachers and students, genetic counselors, and biotechnology company personnel.

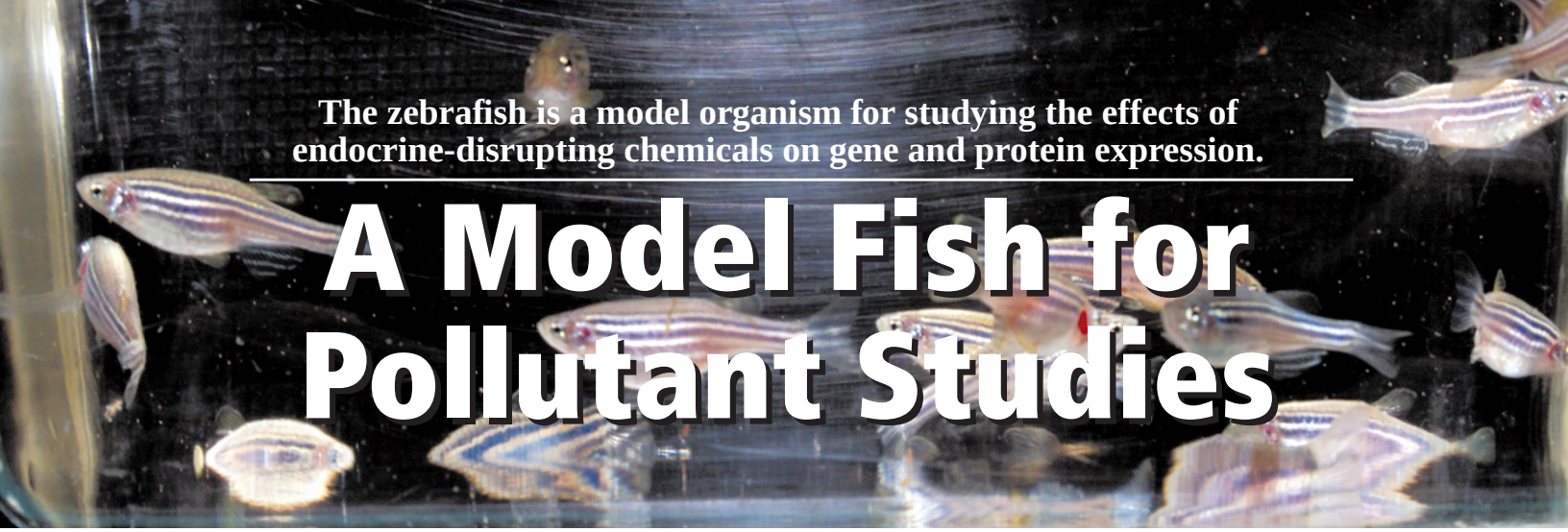
The Web site also brings in 2100 questions a year by electronic mail, including letters from parents of children with genetic diseases.

The work has its surprises. "When I gave a presentation on the Human Genome Project to a group of doctors," Mansfield says, "I was first asked not about the medical implications but rather for advice on investing in biotech stocks."

A biologist by training, Mansfield's prior research included studies on changes in the proteins produced by abnormal red blood cells in leukemic mice after chemicals were added to make these cells normal. "I was doing proteomics research back when it wasn't cool," Mansfield says. Now, she heads a leading source of information on a field that is really hot.

Judy Wyrick (right) and Sheryl Martin examine the on-screen image of the cover of a document entitled Genomes to Life, *which the HGMIS group prepared for DOE. The draft document, available on the HGMIS Web site, describes the goals of the proposed program, including identifying multiprotein complexes that carry out the functions of living systems. See* www.DOEGenomesToLife.org.

Curtis Boles; enhanced by Gail Sweeden



The zebrafish is a model organism for studying the effects of endocrine-disrupting chemicals on gene and protein expression.

A Model Fish for Pollutant Studies

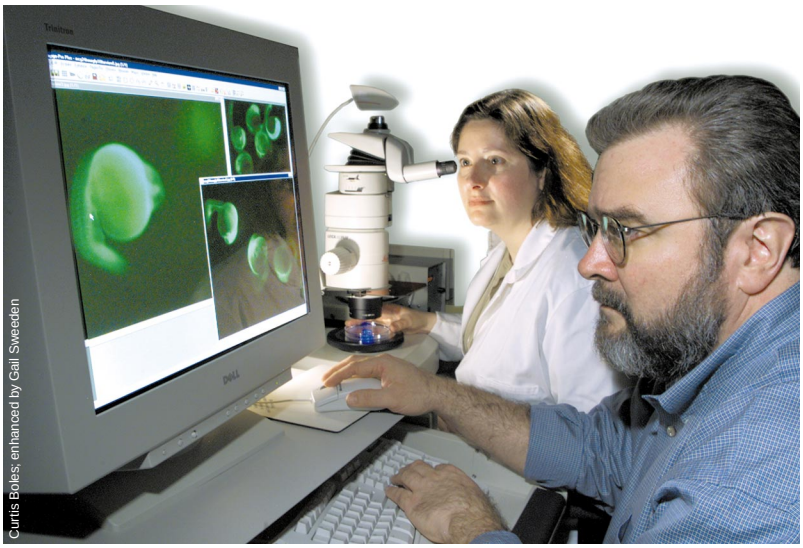
Mark Greeley admits the one-inch-long, brownish-gray zebrafish is a "rather drab aquarium fish." But he says the small striped fish "can be considered the mouse of the fish world, because it is arguably the best vertebrate model for development." Greeley then explains how to make a zebrafish look interesting: Turn it green.

Greeley, director of ORNL's Aquatic Toxicology Laboratory in the Environmental Sciences Division, is an ecotoxicologist who studies the effects of contamination on fish and relates the effects back to humans and to other wildlife, as well. He is particularly interested in endocrine-disrupting chemicals, artificial and natural substances that mimic or block the normal actions of hormones such as estrogen and testosterone. Endocrine-disrupting chemicals, including such common or well-known substances as organochlorine pesticides, plasticizers, dioxins, and PCBs, can cause significant health problems in wildlife and humans.

The risks of endocrine-disrupting chemicals are most apparent in fish and wildlife. "Downstream of sewage treatment plants, fish often have combined male and female sex organs from exposure to estrogenic compounds in pharmaceutical waste and other endocrine-disrupting chemicals," Greeley says. "Downstream of a paper and pulp mill, we have studied a fish population that consists entirely of males because of exposure to natural plant sterols and other hormone mimics in the wood-processing by-products. Endocrine-disrupting chemicals have been implicated in skewed sex ratios in birds, developmental abnormalities in frogs and other amphibians, deformed sex organs and other reproductive problems in alligators, and decreased fertility in a variety of wildlife species."

Endocrine-disrupting chemicals may also have subtle but no less significant health effects on humans. Although the evidence is less conclusive, these chemicals have been implicated in the development of human breast and testicular cancers, early onset of puberty in both boys and girls, lowered sperm counts, and decreased fertility in the developed world.

Greeley is using the zebrafish as a model to study the molecular toxicology of endocrine-disrupting chemicals. "The ultimate goal of this research," he says, "is to understand how these toxicants cause their adverse effects by discovering which genes are turned on and off and which proteins are produced or altered as a result of exposure."



Curtis Boles; enhanced by Gail Sweeden

Suzanne Garnmeister examines a zebrafish embryo in a fluorescence microscope while Mark Greeley looks at screen images of zebrafish embryos. Embryos were exposed to estrogen after being microinjected with a gene construct combining the vitellogenin gene with a bioreporter gene for the jellyfish green fluorescent protein.

To examine the behavior of specific genes as a result of exposure to endocrine-disrupting chemicals, Greeley and his associates—Suzanne Garnmeister, Kitty McCracken, and Zamin Yang—use a jellyfish gene that codes for a green fluorescent protein (GFP). The bioreporter GFP gene is first genetically fused to a hormone-responsive gene, such as the vitellogenin gene, which regulates the formation of yolk proteins in zebrafish and other non-mammalian vertebrate organisms. The resulting gene construct is then microinjected into fertilized single-cell eggs and incorporated into the zebrafish genome. Whenever the native hormone-responsive gene is expressed in these transgenic fish in response to natural hormones or from exposure to endocrine-disrupting chemicals, the bioreporter gene construct also turns on and glows with the typical intense blue-green fluorescent signature of the GFP gene. Because of the remarkable clarity of the zebrafish embryo, the researchers can determine exactly when and where the specific gene of interest is activated or turned off in response to toxicant exposure.

The simultaneous expression of potentially thousands of zebrafish genes during exposure to endocrine-disrupting chemicals is also being examined using a zebrafish DNA "toxchip" microarray currently being developed in collaboration with Mitch Doktycz and Peter Hoyt of the Life Sciences Division (see the article on their work on p. 7). "Results with an early version of this chip," Greeley says, "clearly demonstrate the enormous potential of this tool in toxicological research."

Genes alone can't tell the whole story. Proteins produced as a result of gene expression actually mediate the toxic response.

"In collaboration with Brian Bradley of the University of Maryland–Baltimore County, we have demonstrated significant changes in protein expression in developing zebrafish following exposure to estrogen and estrogen-mimicking chemicals," Greeley says. "Proteins normally produced during early development were absent, but atypical proteins were produced. We expect the zebrafish to be an excellent model for studying the functional relationship between gene and protein expression in response to toxicant exposure.

"Because humans have many of the same genes, what we learn about gene and protein expression in zebrafish exposed to endocrine-disrupting chemicals and other toxicants will help us better understand what these chemicals can do to humans."

Zebrafish are becoming more fascinating now that they are giving scientists the green light to learn more about how toxins affect our genes.

Controlling Carbon in Hybrid Poplar Trees

ORNL scientists are helping to search for genes that could allow the creation of trees that store more carbon and offer higher-value products.

Using carbon dioxide from the atmosphere, as well as sunlight and water, hybrid poplar trees grow fast and tall, up to 12 feet per year. They also harbor a considerable amount of carbon in their stems, branches, leaves, and roots. Plant geneticists would like to design hybrid poplar trees that maximize the amount of carbon they store in their cell walls. These trees could then be used to more effectively sequester carbon dioxide, a greenhouse gas, through increased carbon storage in their roots and, after the roots decay, in soil. Alternatively, when harvested and digested microbially, these "designer" trees could offer an increased yield of commodity chemicals (e.g., polylactic acid, furfural, and acetic acid) and ethanol fuel.

In trees, carbon is "allocated" between aboveground stems, branches, and leaves and belowground roots. It is "partitioned," or divided, among three types of plant cell-wall components—cellulose, hemicellulose, and lignin. A plant could be designed to have an unusually high cellulose content above ground, if increased ethanol production is desired. In addition, if carbon sequestration is the goal, its roots could be designed to have unusually high lignin content, which is resistant to degradation by microbes, increasing the residence time of carbon in the soil.

To produce a hybrid poplar tree, flowers from the female, Populus trichocharpa, *are isolated and inoculated with pollen from the male,* Populus deltoides. *Hybrid progeny grow from the resulting seeds. The flowers and pollen from the hybrid trees can be crossed to produce descendants of the "grandparent" trees. The genomes of the grandparent trees and their progeny can be compared to a deck of cards that is shuffled and reshuffled for each tree. In the grandparents, the male has all black cards and the female has all red cards. In the next generation, the cards are half red and half black. In the third generation, the percentages of black and red cards vary greatly for each tree, producing genetic diversity that allows the linkage of genes to highly desirable traits.*

"In five years, we hope to determine which genes control carbon allocation and partitioning in hybrid poplar trees," says Gerald Tuskan, a plant geneticist in ORNL's Environmental Sciences Division (ESD). "Our research indicates that carbon allocation is controlled by a small number of regulatory genes, that separate genes controlling cell-wall chemistry operate independently above ground and below ground, and that genes controlling carbon allocation affect carbon partitioning."

Tuskan is working on a three-year project to enhance bioenergy conversion and carbon sequestration in woody plants with his ESD colleagues Stan Wullschleger, Tim Tschaplinski, and Lee Gunter; Brian Davison of the Chemical Technology Division; and several researchers from DOE's National Renewable Energy Laboratory. The team is studying wood tissue samples from some 300 hybrid poplars grown in Washington that are the progeny of trees from Minnesota and Oregon parents.

Tuskan and his colleagues are mapping the hybrid poplar genome by finding genetic "markers" unique to trees that have a desirable trait, such as higher-than-normal cellulose content above ground. A marker is a known DNA sequence associated with a particular gene or trait; in this study, it consists of two unique, non-repeating DNA sequences flanking simple sequence repeats, such as GAGAGAGAGA. Some 150 markers have been found so far; the project's goal is 400 markers.

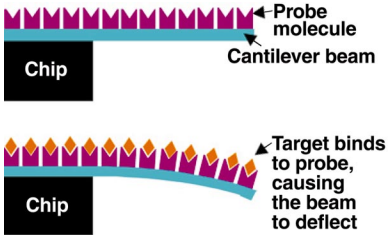
"Each hybrid poplar tree has a unique genetic fingerprint," Tuskan says. "We look for an association between markers unique to each tree and variations in the allocation and partitioning of carbon content. Once we find the marker that controls the trait we are interested in, such as high lignin content in the roots, then we will try to locate the genes responsible. Such genes could be used to design tree root systems that are high in lignin content."

Tuskan is also interested in finding the genes that control the size and thickness of a tree's cell walls, the substructure of wood that determines its usefulness and commercial value. "It's because of differences in cell sizes and wall thicknesses that oak floors are stronger than pine floors, maple furniture is more attractive than aspen furniture, and white oak rather than red oak is used to make barrels to store wine," he says. "Cell dimensions also determine whether a tree's wood is suitable for combustion or production of paper or ethanol."

Use of a light microscope or scanning electron microscope to determine wood cell dimensions in samples from various trees is expensive and time consuming. So, Tuskan sought help from Mike Paulus of ORNL's Instrumentation and Controls Division. Paulus is a co-developer of the high-resolution, X-ray-computed tomography system called a MicroCAT scanner. Although used mostly to image internal defects in small animals, the MicroCAT scanner also offers a faster, better, and cheaper way to measure the lengths and diameters of cell walls in wood. (See article on p. 11.)

"With the MicroCAT, we can get cell measurements from an intact block of wood, whereas for microscope studies, we have to slice wood into very small pieces," Tuskan says. "With the light microscope, we were getting 100-micron resolution, but with the modified MicroCAT, we get 10-micron resolution and may be able to get down to a resolution of one to two microns. The MicroCAT is a great tool for rapidly screening for wood-cell dimensions in the context of a large genetic mapping study."

Disease Detectives



ORNL researchers are developing two types of miniaturized devices for diagnosing diseases. These devices are based on cantilevers and biochips.

CANTILEVER DEVICES

Take a silicon chip as small as a grain of rice and carve out barely visible diving-board–like projections at one edge. Coat these cantilevers with gold. Attach thiolated single-stranded DNA to the gold-coated cantilevers. Allow single-stranded DNA of different sequences to come in contact with the DNA attached to the cantilevers. If a DNA sequence complements the DNA on a cantilever, they will bind together, or hybridize, to form double-stranded DNA. Thomas Thundat and Karolyn Hansen, both of ORNL's Life Sciences Division, use this recipe to distinguish between numbers of base pairs in DNA sequences on cantilevers.

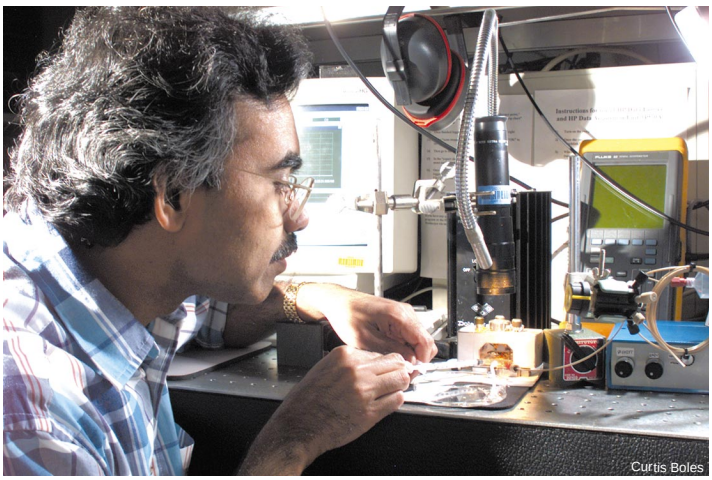
"We immobilized single strands of DNA containing 20 bases on a series of cantilevers," Thundat says. "The DNA bases on these cantilevers paired with 20, 15, 10, and 9 bases of single-stranded DNA introduced to the cantilevers. We found that the cantilevers all bend because of the changes in surface tension as a result of DNA hybridization." The more bases a cantilever holds, the more it bends, changing the angle of deflection of laser light bounced off the cantilever, as recorded in a detector.

Thundat believes that this technology could be used for DNA sequencing and that the approach would cost less and take less time than conventional techniques because it would avoid the step of adding fluorescent dyes to label the DNA bases. This technology has been licensed to Graviton, Inc.

Thundat believes that cantilevers can be used to detect defective genes that cause breast cancer, colorectal cancer, and cystic fibrosis. These mutant genes have one incorrect DNA base. ORNL experiments have shown that a DNA sequence in a liquid sample will hybridize with a complementary DNA sequence bound to a cantilever, even if the sample sequence has one wrong base, or a mismatch.

"We found that a mismatch causes the cantilever to bend up instead of down," Thundat says. "This change in bending direction could be used to detect defective genes that cause disease."

Thomas Thundat examines a cantilever beam (see drawing above) on which an antibody for prostate-specific antigen is attached. By bending when PSA binds to the probe antibody molecule, this device detects early signs of prostate cancer in serum samples with unusually high sensitivity.



The cantilever technology could also be used to detect prostate cancer. ORNL researchers have immobilized on a cantilever the antibody for prostate-specific antigen (PSA), the chemical signal for the disease. An ORNL collaboration with the University of California's Professor Arun Majumdar has shown that the cantilever bends when its antibody matches PSA in serum samples supplied by Majumdar. The sensitivity of this technology is 10 times higher than that of conventional techniques.

BIOCHIP DEVICES

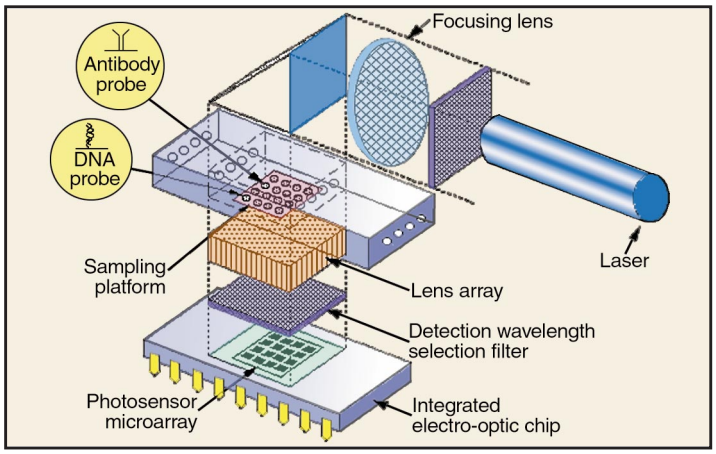
By detecting a mutant breast cancer gene, a doctor can predict that a patient will get breast cancer. By detecting a certain protein, a doctor can determine that the patient has adult-onset diabetes. Someday, physicians will be able to rapidly analyze both genes and proteins from a single drop of a patient's blood, using a palm-size device. At least that's the goal of Tuan Vo-Dinh and his co-workers David Stokes, Minoo Askari, and Guy Griffin, all of the Life Sciences Division, and Alan Wintenberg of ORNL's Instrumentation and Controls Division.

"We can do genomics and proteomics on a single platform using our multifunctional biochip," says Vo-Dinh. "The biochip is being designed to process up to 100 samples in 30 minutes." The multifunctional biochip is an advanced version of the group's DNA biochip, which contains only DNA probes. This technology has been licensed to HealthSpex, Inc., in Oak Ridge.

To sample a patient's blood for a DNA sequence that is a red flag for a genetic disease, the multifunctional biochip has a complementary DNA sequence to which this mutant sequence will bind. To sample for a specific disease-related protein, the biochip has a probe (e.g., an antibody or protein receptor) that will bind with this particular protein.

ORNL experiments have shown that the biochip can detect the tuberculosis bacterium, the HIV gene, a cancer suppressor gene, the anthrax bacterium used in biological warfare, and *Escherichia coli* found in contaminated food. Thus, the biochip could be used for medical diagnosis, defense, and food safety applications.

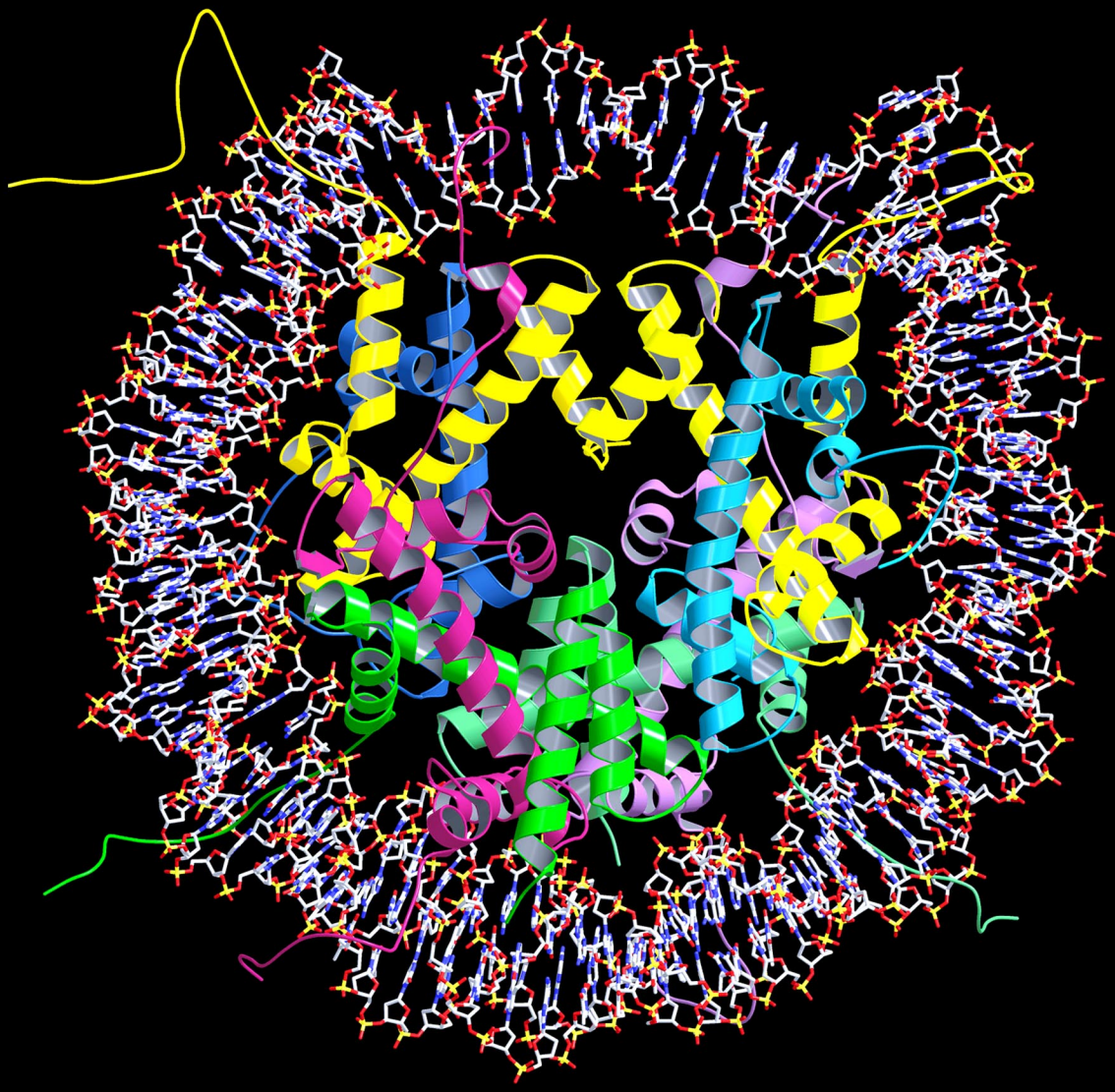
Small though they are, both the cantilever device and biochip could make a big contribution to health care.

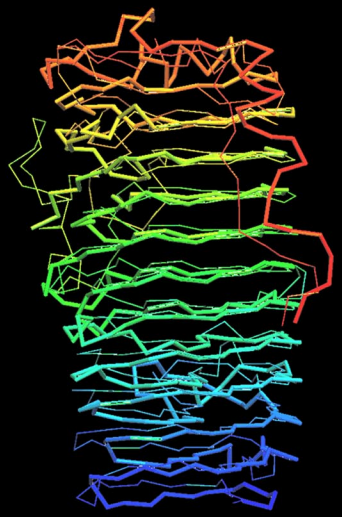


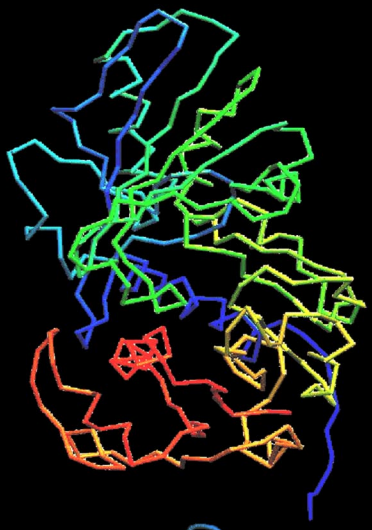
To make a multifunctional biochip work in the doctor's office, a patient's blood sample would be processed to separate DNA sequences and proteins, which are then labeled with a fluorescent dye. The chip would then be exposed to these blood fragments. Disease-related DNA and proteins will bind to the chip's complementary DNA and antibody probes. A miniaturized laser diode illuminates the array of sites, causing fluorescence at the sites of hybridization. The signals are collected and sent to a microprocessor chip, which stores the identity of the probes at each location. By matching signal locations to the probes known to be there, the microprocessor detects disease-related genes and proteins in the patient's blood and relays the diagnosis to the physician.

OAK RIDGE NATIONAL LABORATORY *REVIEW*
P.O. Box 2008, Oak Ridge, Tennessee 37831-6266

PRSRT STD
U.S. Postage
PAID
Permit No. 3
Oak Ridge, TN







The experimentally determined crystal structure of the nucleosome core particle is shown above in a color rendering. The protein is represented by ribbon models of the secondary-structure elements. The DNA is represented as a wire frame model, indicating base pairing between complementary strands. The image recently graced the cover of *Acta Crystallographica Section D, Biological Crystallography* (Volume 56, Part 12, December 2000). The work is a result of research by the Life Sciences Division's Gerry Bunick, UT's Joel Harp and Bryant Hanson, and David Timm of the Indiana University School of Medicine. The group hopes to use the Spallation Neutron Source to better understand interactions between DNA and proteins and between proteins. **(See article starting on p. 24.)**

The three images at the right are predicted protein structures. **(See article starting on p. 20.)**