DeGeM

DIGITALLY ENABLED GENOMIC MEDICINE

FUSING REVOLUTIONS IN GENOME-ENABLED BIOMEDICINE AND THE USE OF THE INTERNET AND INFORMATION TECHNOLOGY IN SOCIETY

PROJECTS AND PEOPLE AT THE UNIVERSITY OF CALIFORNIA, SAN DIEGO AND CAL-(IT)^2

Cal-(IT)^2's mission is simple: Extend the reach of the Internet throughout the physical world. Cal-(IT)^2 teams University of California, San Diego (UCSD) and University of California, Irvine (UCI), faculty, students, and research professionals with leading California telecommunications, computer, software, and applications companies to conduct research on the scientific and technological components needed to bring this new Internet into being. Institute applications researchers are conducting their studies in "living laboratories" to investigate how this future Internet will accelerate advances in environmental science, civil infrastructure, intelligent transportation and telematics, genomic medicine, the new media arts, and educational practices.

Cal-(IT)^2's research program is organized conceptually into vertically interlocking and collaborating layers *(www.calit2.net/research/layers.html)*. Research projects serve as vertical, multidisciplinary "convective currents" to draw and mix participation from across several layers. This brochure focuses on the Digitally Enabled Genomic Medicine (DeGeM, *www.calit2.net/degem*), the application layer of Cal-(IT)^2 at UCSD, which is lead by John Wooley. Pierre Baldi is the layer leader of DeGeM for Cal-(IT)^2 at partner institution UCI. Highlighting these selected projects and associated researchers will define the breadth and excitement of DeGeM and identify opportunities for involvement for interested academic researchers, students, and industrial partners.

DeGeM: This acronym is pronounced "dee gem." We have chosen to spell the abbreviation DeGeM to reflect the desired pronunciation.

UCSD

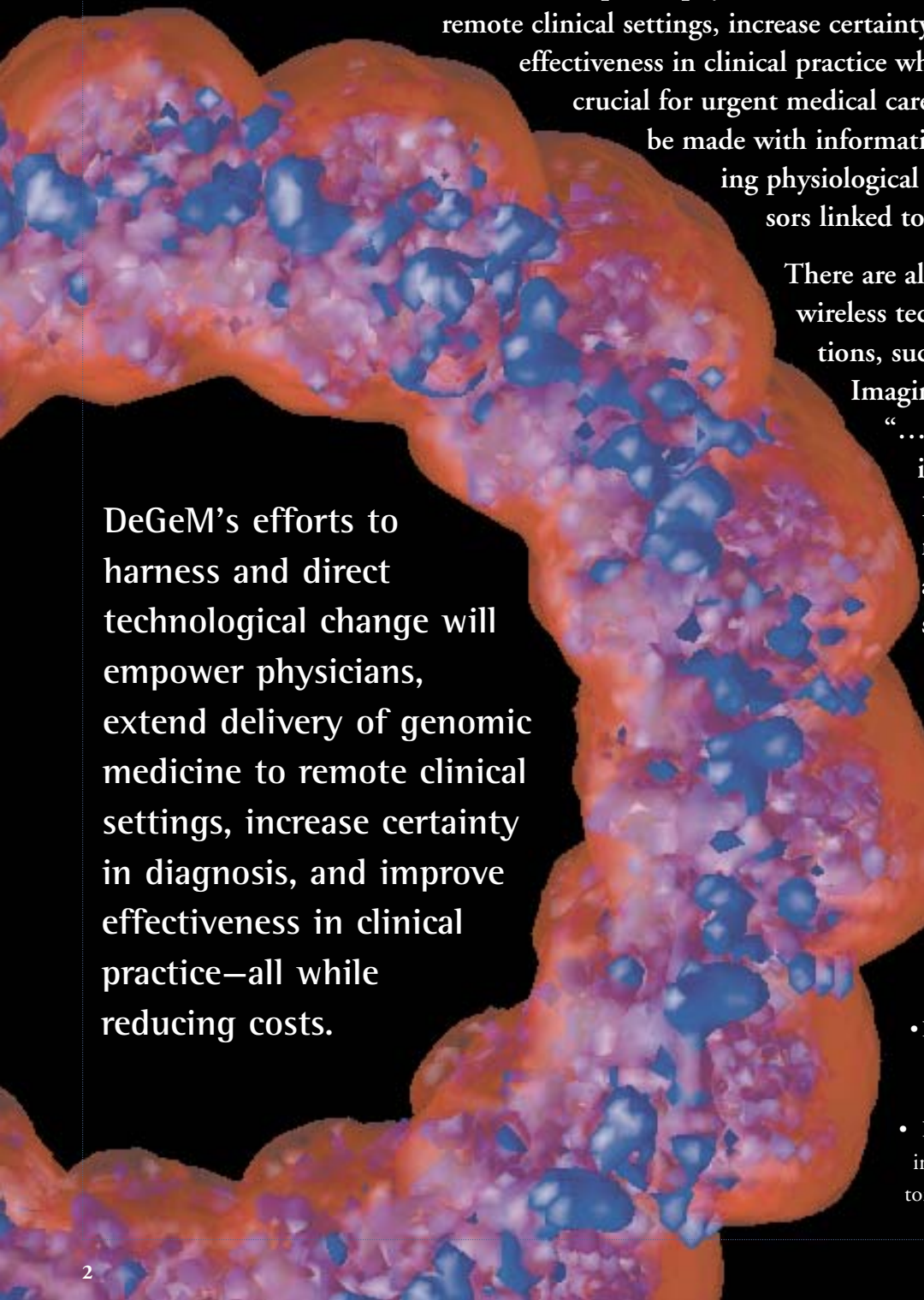
it^2

Contents:

Background Image: La Jolla Shores; Photo Credit: John Wooley, DeGeM Layer Leader, UCSD division Cal-(IT)^2.

Vision: THE VISION OF DIGITALLY

The combination of personalized (genomic) medicine and the application of wireless and sensing technology to health care delivery (digitally enabled medicine) will provide enormous benefits to the citizens and communities of California, the State's economy, and our society in general. High-speed, high-bandwidth wireless communication and sensing devices linked to powerful, interoperable clinical and biomedical databases will empower physicians, extend delivery of genomic medicine to remote clinical settings, increase certainty in diagnosis, and improve effectiveness in clinical practice while reducing costs. Decisions crucial for urgent medical care (e.g., cardiac monitoring) will be made with information updated on demand, including physiological data from non-invasive biosensors linked to wireless transceivers.



DeGeM's efforts to harness and direct technological change will empower physicians, extend delivery of genomic medicine to remote clinical settings, increase certainty in diagnosis, and improve effectiveness in clinical practice—all while reducing costs.

There are already examples of the power of wireless technologies in medical applications, such as in the video pill (Given-Imaging, *www.givenimaging.com*) "…to diagnose diseases of the small intestine including Crohn's Disease, Celiac disease and other malabsorption disorders, benign and malignant tumors of the small intestine, vascular disorders and medication-related small bowel injury"; the smart band-aid (PhiloMetron, *www.philometron.com*); or the wearable body monitoring devices for wellness (e.g., BodyMedia).

Members of the DeGeM layer want to "improve health care by 3 PM," namely by

- Personalized Medicine-applying understanding about genomic information to tailor solutions to individuals
- Preventive Medicine-creating and improving technologies to provide constant monitoring of patients

ENABLED GENOMIC MEDICINE

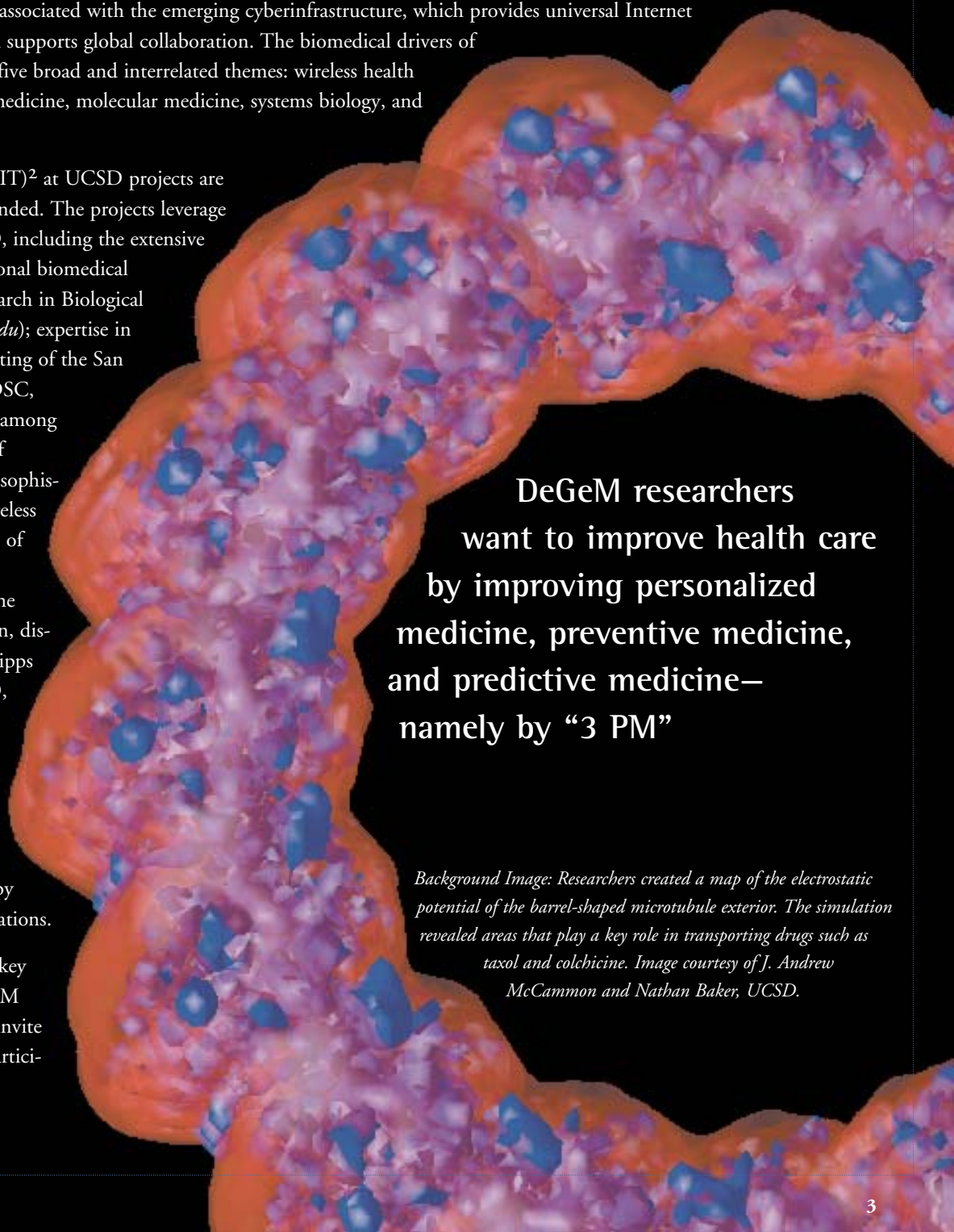
- Predictive Medicine-combining the genomic information and improved monitoring and data integration tools to forecast and anticipate health care needs for each person

Implementing the DeGeM Vision of Cal-(IT)² at UCSD: Building from experience to create laboratories of the future for the untethered health care provider: In our first three years, we have focused on ways to enhance the intellectual, analytical, and investigative output of the world's scientific community by harnessing developments in information technology, integration of information, and knowledge discovery that are associated with the emerging cyberinfrastructure, which provides universal Internet access to distributed resources and supports global collaboration. The biomedical drivers of our development activities reflect five broad and interrelated themes: wireless health care delivery, telescience and telemedicine, molecular medicine, systems biology, and bioinformatics.

Within the DeGeM layer of Cal-(IT)² at UCSD projects are being initiated, renewed, and extended. The projects leverage the strength and history of UCSD, including the extensive results of multi-scale, multi-locational biomedical research from the Center for Research in Biological Structure (CRBS, *www.crbs.ucsd.edu*); expertise in high- performance parallel computing of the San Diego Supercomputer Center (SDSC, *www.sdsc.edu*); clinical experience among physicians of the UCSD School of Medicine *(medicine.ucsd.edu)*; the sophistication of bioengineering and wireless research at UCSD's Jacobs School of Engineering *(www.jacobsschool.ucsd.edu)*; and the 100 years of research in exploration, discovery, and explanation at the Scripps Institution of Oceanography (SIO, *sio.ucsd.edu*) where researchers are mining proteins from the sea for use in biomedical research.

Research projects drive development of telecommunications and information technologies, pulled by the demands of biomedical applications.

The following pages elaborate on key projects and research of the DeGeM layer of Cal-(IT)² at UCSD. We invite you to explore opportunities to participate with us.



DeGeM researchers want to improve health care by improving personalized medicine, preventive medicine, and predictive medicine— namely by "3 PM"

Background Image: Researchers created a map of the electrostatic potential of the barrel-shaped microtubule exterior. The simulation revealed areas that play a key role in transporting drugs such as taxol and colchicine. Image courtesy of J. Andrew McCammon and Nathan Baker, UCSD.

Projects: UCSD ACTIVITIES

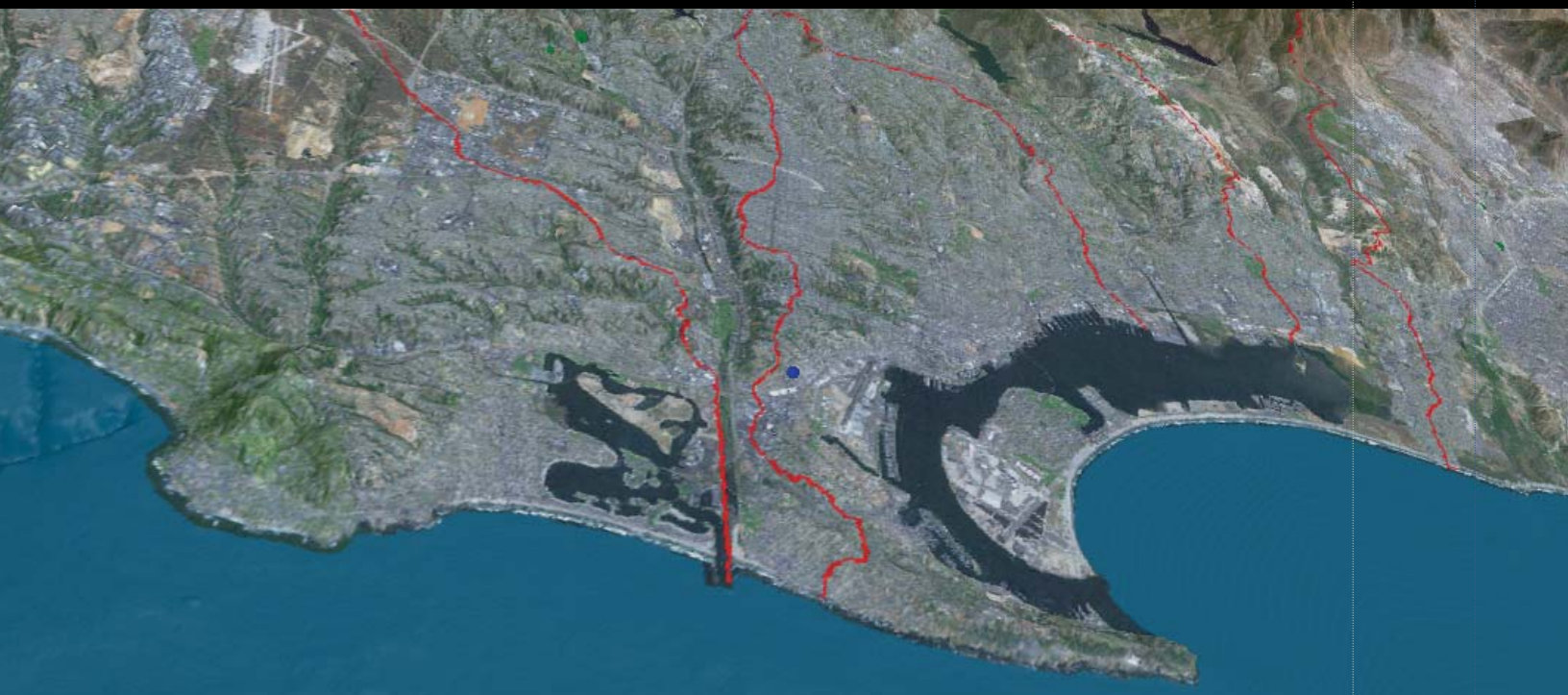
The following projects represent key DeGeM activities at UCSD. Some, like the Biomedical Informatics Research Network (BIRN), focus on data integration, a necessary step to fulfill the vision of DeGeM. Others, like the Wireless Internet Information System for Medical Response in Disasters (WIISARD), give us experience in emergency response-allowing DeGeM to push technological development in wireless networking during periods of intense use.

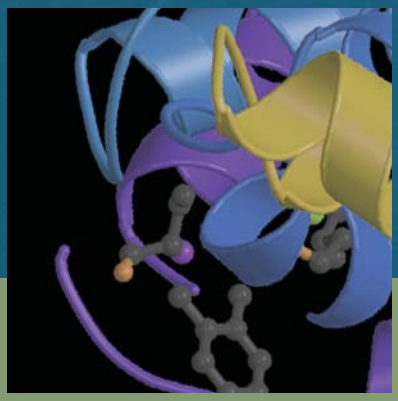
All projects bring together multidisciplinary teams of researchers focused on biomedical issues and push the development of information and telecommunications technologies. This combination enables new technical capabilities and gives rise to new research issues that need to be addressed.

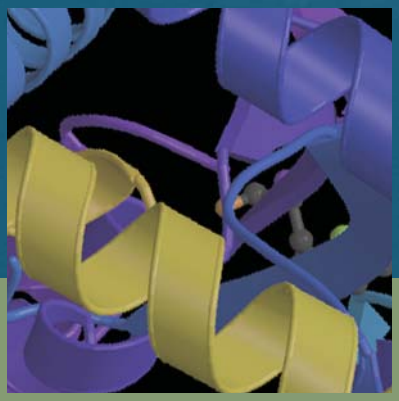
While DeGeM and Cal-(IT)^2 are focused on improving life for the citizens of California, all projects are likely to have a much broader impact, nationally and internationally. Their prominence sets national and international directions for research and brings recognition to the talent gathered under the Cal-(IT)^2 umbrella. In addition, DeGeM's reputation is helping to draw ever-more talented students and faculty to campus, thereby enriching the local intellectual environment.

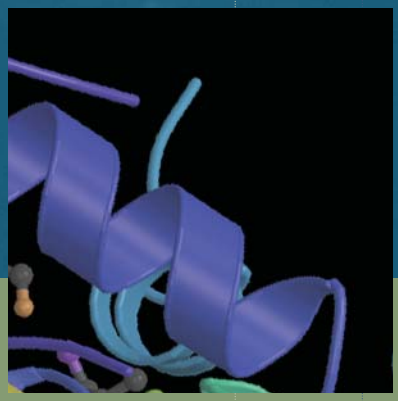
Background Image: As part of the OptIPuter project (www.siovizcenter.ucsd.edu/optiputer), and in collaboration with the Regional Workbench Consortium (RWBC: www.regionalworkbench.org), a 3-D regional canvas of California was created using multiple data streams. These data included high-resolution imagery draped on topography, juxtaposed with local seismicity (spheres hidden from view) and water sheds in the San Diego region (red lines). San Diego Bay and Mission Bay are viewable in the foreground. Through Cal-(IT)^2 collaborations, spatial display technologies such as this one are being applied in mapping areas of the brain.

IN SUPPORT OF CAL-(IT)^2 DEGEM









Background Image: High-resolution section of a rat cerebellum stained to highlight various cells. Green represents a calcium channel that plays an important role in proper neuronal function & is highly enriched in neurons called Purkinje cells. Blue represents the cell's nucleus. Red cells or glia provide physical & nutritional support for neurons. Developed by NCMIR. Studies on model animals gives insight into basic biology in humans.




Biomedical Informatics Research Network (BIRN) *Pioneering the Use of Grid Infrastructure for Medical Research and Patient Care (www.nbirn.net)*

BIRN is a Cal-(IT)^2 application-driven living laboratory. Established in 2001, BIRN developed and continues to design the hardware, software, and protocols necessary to share and mine data to support basic and clinical research. Central to the project is a scalable cyberinfrastructure consisting of advanced networks, federated distributed data collections, computational resources, and software technologies that are integrated to meet the evolving needs of collaborative testbed investigators. By pooling domain expertise, specialized research facilities, instrumentation, applications, and regional information, these investigators are tackling disease studies of greater scope and complexity than are possible by investigators working in isolation from one another.

The BIRN cyberinfrastructure is developed by the BIRN Coordinating Center and driven by the requirements of three research test beds:

- Function BIRN: 11 universities studying regional human brain dysfunctions related to schizophrenia.
- Morphometry BIRN: Nine research institutions investigating whether structural differences in the human brain distinguish diagnostic categories in unipolar depression, mild Alzheimer's disease, and mild cognitive impairment.
- Mouse BIRN: Four institutions studying animal models of disease at different anatomical scales to test hypotheses associated with human neurological disorders including schizophrenia, attention-deficit hyperactivity disorder, multiple sclerosis, and Parkinson's disease.

BIRN works as part of UCSD's Center for Research on Biological Systems. As additional biomedical research and clinical care test beds evolve, BIRN will continue to expand the boundaries of information technology infrastructure, enriching the Global Grid movement by providing application pull from these new domains. BIRN leverages other activities of Cal-(IT)^2 by involving researchers from SDSC's Knowledge and Data Engineering Lab, the Lambda Grid living lab and OptIPuter project (also profiled in this brochure), and the Interfaces and Software Systems layer. Furthermore, BIRN integrates activity across the two Cal-(IT)^2 partners, UCSD and UCI, with UCI's Steven Potkin leading the Function BIRN. BIRN is supported by an award from the National Center for Research Resources of the National Institutes of Health, which extended the project in 2004. The principal investigator of the BIRN Coordinating Center is Mark Ellisman.

Wireless Internet Information System for Medical Response in Disasters (WIISARD) *Bringing Cutting-Edge Wireless Internet Technologies from the Hospital to the Field Treatment Station (www.wiisard.org)*

The difficult-to-accept truth of our present era is that the United States faces a future clouded by the threat of terrorist actions involving nuclear, biological, and chemical agents. Use of such weapons by terrorists will produce large numbers of civilian casualties that will overwhelm existing health care facilities, jeopardizing the lives of victims and health care providers.

The WIISARD project is using wireless Internet information system technologies to address recognized shortcomings in the field care of victims at mass casualty events. WIISARD is designed to be deployed at the site of a nuclear, biological, or chemical attack or a natural disaster to support the care of large numbers of victims for a period of hours to days, while national medical resources are being marshaled to aid delivery of definitive care.

WIISARD is an integrated application that is bringing cutting-edge wireless Internet technologies from the hospital to the field treatment station. It has components that enhance the situational awareness of first responders, facilitate recording of medical data, aid in the monitoring of severely ill patients, and facilitate communication of data to hospitals.

The WIISARD project has brought together first responders from the County of San Diego Metropolitan Medical Strike Team (MMST) (including fire, paramedic, public health officials, hazardous materials and law enforcement personnel) with engineers and computer scientists from Cal-(IT)^2 and physicians from the faculty of UCSD's School of Medicine. These teams will evaluate WIISARD throughout the three-year contract, beginning with controlled studies of individual components and culminating with a randomized trial conducted during a simulated chemical weapons attack. The first hand-on opportunity to test aspects of wireless technology took place in a May 2004 simulated event of a "dirty bomb" explosion inside a Carlsbad office building. This was one of the largest-ever emergency response drills in San Diego County—dubbed *Operation Moonlight*—to test preparedness among first responders and emergency-relief agencies, and was conducted with the MMST and WIISARD researchers and students (see *www.calit2.net/research/labs/features/5-17-04_wiisard.html* for details).

The lessons learned within WIISARD are likely to lead to new approaches for care and monitoring of patients in hospitals and clinics as well as improvements in wireless network robustness and security.

WIISARD was launched in 2003 and is funded by the National Library of Medicine (NIH). PI: Leslie Lenert, MD, UCSD; Co-PI: Ramesh Rao, Ph.D., UCSD.

Joint Center for Structural Genomics (JCSG)

Introducing Advanced Automation Technology in Protein Structure Determination (www.jcsg.org)

JCSG is developing new technologies for high-throughput structure determination and testing them by solving the structures of hundreds of proteins from a thermophilic bacteria *Thermotoga maritima.*

Structural genomics, which builds on genome sequencing efforts, focuses on the characterization of proteins, the products of genes. The growing impact of structural biology on biomedical research has prompted efforts to determine protein structures on a large scale.

The objective is to make these structures widely available for clinical and basic studies that will expand the knowledge of the roles of proteins in normal biological processes and in disease. For example, the structure of a disease-related protein can provide insight into how the protein works normally and how a faulty structure can cause disease. This same structure may reveal how to design drugs to treat that disease. Although structure determination techniques—chiefly X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy—have advanced dramatically in recent years, they are still time-consuming and labor-intensive. The research centers funded by NIH seek to streamline and automate these techniques, as well as every other task in structural genomics, ranging from selecting proteins for structure determination to analyzing the final data.

The JCSG five-year goal is to introduce advanced automation technology in protein structure determination. More than 100 protein structures have been solved by JCSG researchers. This number may reach 200 before the end of this phase of the project. The center is composed of three core groups, each handling distinct, but interconnected technological challenges. The groups are the Bioinformatics Core (BIC) located at UCSD, the Crystallomics Core (CC) located at The Scripps Research Institute (TSRI) and Genomic Novartis Foundation Institute, and the Structure Determination Core (SDC) located at Stanford University.

The Cal-(IT)^2 focus has been on BIC, which is responsible for target selection, sample tracking, information management, structure validation and deposition, and post-structural analysis. Through these functions, BIC provides the integrated informatics backbone required for the successful operation of JCSG. By developing tools for large-scale automated data collection, integration, and analysis, BIC is addressing a new challenge in biological sciences: how to integrate new high-throughput technology into traditional single-system-oriented biological research. As JCSG prepares for a competitive renewal for the next five years, its focus will shift from basic technology development to direct participation in biomedical research.

The JCSG was launched in 2001 and expanded in 2003. It is funded by the National Institutes of General Medical Science (NIGMS) of NIH. The PI is Ian Wilson of TSRI; the UCSD lead is John Wooley and the BIC Leader is Adam Godzik.

Protein Data Bank (RCSB PDB) *Providing the International Community with Persistent & Searchable Access to 3D Macromolecular Data (www.rcsb.org)*

A detailed description of the shape of a macromolecule can provide critical information to understand how it functions. The PDB is the sole international repository for 3-D structure data of biological macromolecules.

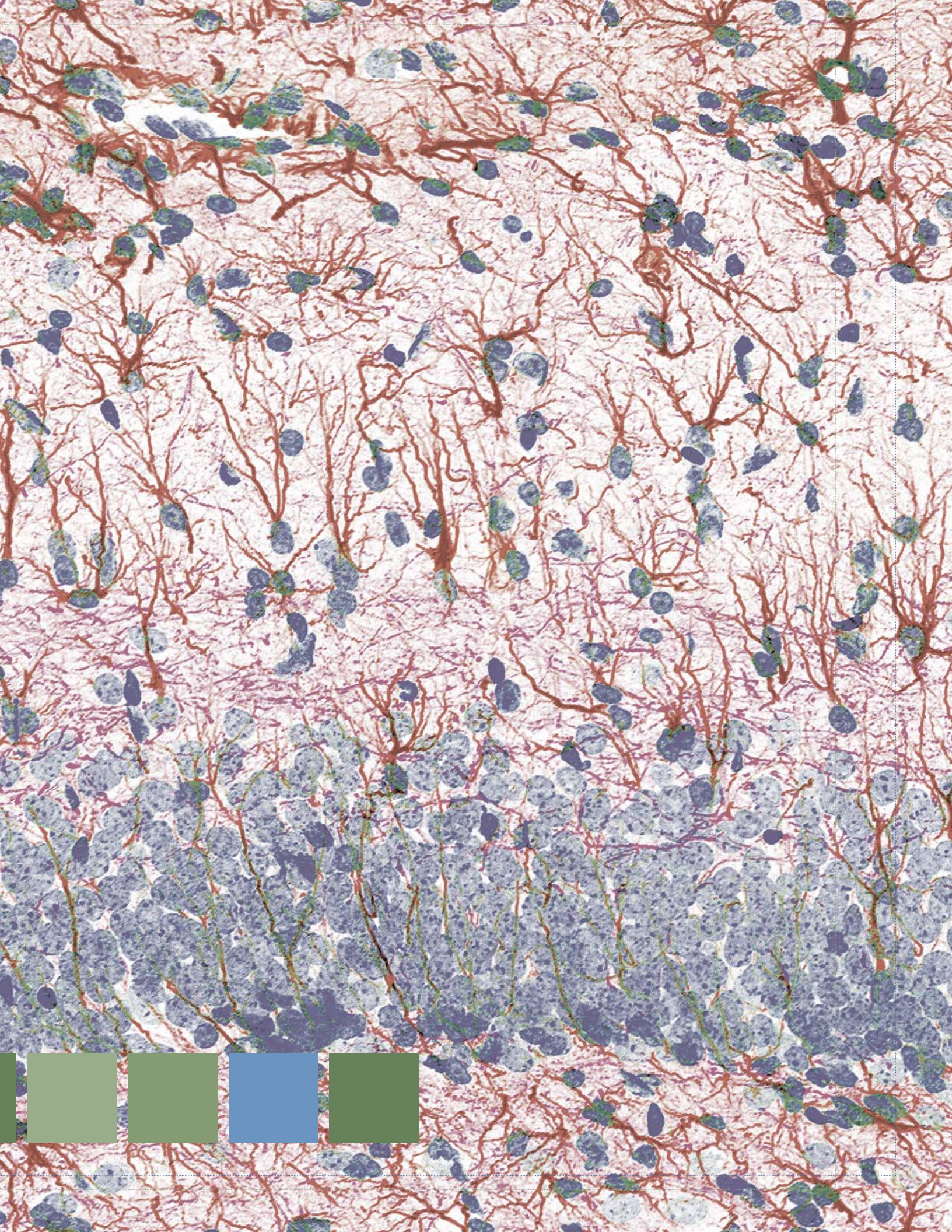
The data archived in the PDB assists the pharmaceutical and biotechnology industries in developing new drugs. Medical researchers use the structures of proteins and other biological macromolecules to understand diseases and unlock their therapeutic potential. Information contained in the PDB also helps researchers understand the chemistry and biochemistry of natural processes.

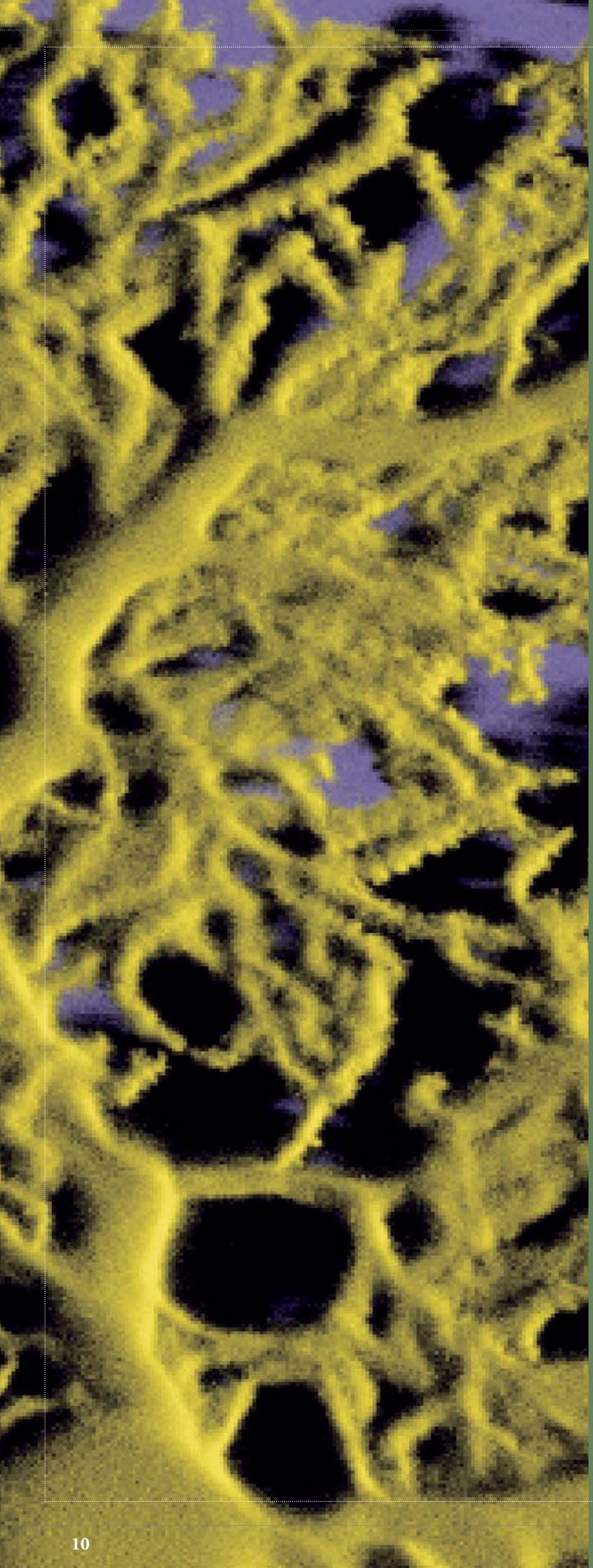
The PDB is working hand in hand with structural genomics efforts to ensure that the technology and capabilities needed to facilitate the collection, validation and distribution of data generated by these efforts will be available. This key, internationally available data resource provides the basic scientific knowledge to allow DeGeM researchers to take first steps toward personalized, preventive, and predictive medicine.

The PDB was founded in 1971 at Brookhaven National Laboratory. Their initial release contained seven structures. Since July of 1999, the PDB has been managed by three member institutions of the Research Collaboratory for Structural Bioinformatics (RCSB) and, as of March 2004, the archives contained almost 25,000 structures with more than 4,600 structures added in 2003 alone. The mission of the RCSB PDB is to provide consistent, well-annotated data in a timely manner so as to enable biological research. All data in the PDB archives are freely available. The RCSB PDB is managed by Rutgers, the State University of New Jersey; the San Diego Supercomputer Center at the University of California, San Diego (SDSC/UCSD); and the University of Maryland Biotechnology Institute/National Institute of Standards and Technology's Center for Advanced Research in Biotechnology (CARB/UMBI/NIST).

RCSB obtained responsibility for PDB in 1998, and its support was renewed in 2003. PDB is supported by funds from the National Science Foundation, the National Institute of General Medical Sciences, the Department of Energy, the National Library of Medicine, the National Cancer Institute, the National Center for Research Resources, the National Institute of Biomedical Imaging and Bioengineering, and the National Institute of Neurological Disorders and Stroke. The overall director of the RCSB PDB is Helen M. Berman (Rutgers). Other members of the RCSB PDB leadership team are Philip E. Bourne (SDSC/UCSD), Gary Gilliland (CARB/UMBI/NIST), and John Westbrook and Judith L. Flippen-Anderson (both of Rutgers).

Image (opposite page): Section of rat hippocampus stained with various cellular markers & imaged using 2-photon fluorescence microscopy. Developed at NCMIR.





LIPID Metabolites And Pathways Strategy (LIPID MAPS) *Developing Understanding & Interaction Networks of Lipids in Cell Function (www.lipidmaps.org)*

Sequencing of the human genome has made it possible and provided the impetus for building a comprehensive picture of a mammalian cell. Significant efforts are underway in the fields of genomics and proteomics to identify all genes and proteins in a given organism. The goal is a complete map of the genes, gene products, and their interaction networks in a functioning cell.

However, to establish a truly comprehensive picture of a cell requires tying the cell's metabolome into the rapidly developing genomic and proteomic maps, creating a complete map of the genes, gene products, and the metabolites they produce, as well as their interaction networks in a functioning cell.

A cell's metabolome is an enormous and complex entity. Therefore, the LIPID MAPS Consortium is focusing on the lipid components of the metabolome by developing an integrated metabolomic system capable of characterizing the global changes in lipid metabolites ("lipidomics"). Lipids are central to the regulation and control of cellular function and disease. Comprising an enormous fraction of cellular metabolites, lipids include such diverse classes as fatty acids, eicosanoids, phospholipids, neutral lipids, sphingolipids, glycosphingolipids, sterols, polyisoprenoids, and glycolipids. Because of similarities in the chemical characteristics of compounds, the lipid metabolic pathways are very complex.

The goals of the LIPID MAPS initiative are to separate and detect all lipids in a specific cell and discover and characterize any novel lipids that may be present; quantitate each of the lipid metabolites present and the changes in their levels and locations during cellular function; define the biochemical pathways for each lipid; and develop lipid maps, which define the interaction networks.

LIPID MAPS was launched in 2003 and is funded by National Institute of General Medical Sciences' Large Scale Collaborative "Glue" Grant. The overall PI is Edward A. Dennis (Editor-in-Chief of the *Journal of Lipid Research*), and the project involves researchers from 18 institutions. This activity is central to DeGeM, and takes us closer to the vision of personal, preventive, and predictive medicine.

Image: The image is a 3-D reconstruction of a Purkinje cell from the rat brain cerebellum. It was created at the National Center for Microscopy and Imaging Research from a series of light microscopic images. The fuzz covering the distal processes is the dendritic spines, small protuberances from the dendritic shaft that are the main sites of synaptic input. The Purkinje neuron has up to 150,000 of these spines covering its dendrites.

The HAP Project *Discovering the Genetic Basis of Human Disease (www.calit2.net/compbio/hap)*

With the explosion of genomic sequence data and the completion of the human genome project, much of the progress in understanding the genetic basis of disease relies on computational analysis of the genomic data. Some of the most useful data for this analysis is human variation data. This data consists of information on the variation in genes associated with a disease for a population of individuals. Understanding the relation between variation and disease is a fundamental challenge, which can shed light on the genetic basis and mechanisms of human disease. This challenge spans three research fields: genetics, bioinformatics, and medicine.
Understanding the genetic basis of disease involves two steps. First, we must determine the functional variants in each gene locus that is linked to the disease and the effect of functional variants on the regulation and gene products of the gene. Second, we must understand how these intermediate phenotypes affect disease outcomes. Using this information, we can identify subtypes of the disease that are candidates for different drug response.

Our lab's research focuses on building tools to modeling the function of variation in a gene locus, correlating the intermediate phenotypes to disease outcomes, and identifying subtypes of the disease based on genetic variants. The disease focus of this proposal is hypertension, and the tools will be applied to the large amount of data collected at UCSD through the pharmacogenomics project.

These new tools are building upon the foundation of existing tools developed in the lab already publicly available to the research community. The haplotype analysis methods are built on top of a phasing algorithm HAP. This method has been shown to be very effective over long regions and perform much faster than the current state-of-the-art methods. The method has been publicly available for over a year at *www.calit2.net/compbio/hap* and has phased more than 3,000 submitted datasets. We also build on MITRA (MIsmatch TRee Algorithm), a motif-finding method for the detection of transcription factor binding sites. The main idea behind MITRA is that we use a very efficient data structure, a mismatch tree, which allows us to traverse the space of all possible motifs very quickly. The motif-finding software has been available for two years and more than 1,500 data sets have been processed at *www.calit2.net/compbio/mitra*.

National Biomedical Computation Resource (NBCR) *Advancing Biomedical Research by Developing Tools to Provide Access to Advanced Cyberinfrastructure (nbcr.net)*

NBCR's mission at UCSD is to conduct, catalyze, and enable biomedical research by harnessing forefront computational and information technologies. To fulfill this mission, NBCR efforts are focused on four key activities:

- Develop and deploy advanced computational tools for modeling and simulation, query and integration of data resources, 3-D image processing, and interactive visualization
- Integrate computational, data and visualization resources in a transparent advanced grid environment to enable transparent access to distributed data, computational resources and instruments
- Implement and support advanced grid/cyber-infrastructure for biomedical researchers
- Train a cadre of researchers with interdisciplinary knowledge of both biology and the relevant computational technologies

The key aim of the resource is to provide transparent access *to the new and emerging grid infrastructure* that will deliver integrated compute, data, physical, experimental, and human resources to biomedical scientists investigating a wide range of medically important problems spanning scales of biological organization from small molecule drug design and comparative genomics to diagnostic brain imaging and cardiovascular disease

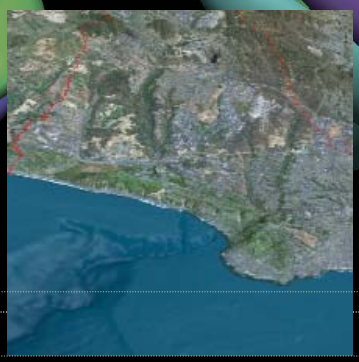
The technology research and development activities of NBCR involve collaborations among researchers at UCSD, SDSC, Cal-(IT)^2, and The Scripps Research Institute and Washington University, Saint Louis. Core research projects include

- Integrative Modeling of Subcellular Processes: Application to Synaptic Activity and Pharmaceutical Discovery
- Data Integration and Analytic Tools for Molecular Sequences
- Structurally and Functionally Integrated Modeling of Cell and Organ Biophysics
- Creating Visualization Environments for Multi-Scale Biomedical Modeling
- Grid Computing and Analysis for Multi-scale Biomedical Applications

NBCR draws upon the expertise of researchers throughout the US in a number of collaborative projects, including the Knowledge and Data Engineering Lab, the Lambda Grid living lab, the Interfaces and Software Systems layer and the BIRN living laboratory.

NBCR was launched in 1994 with funding from the National Center for Research Resources of NIH, and was granted a third 5-year award in 2004. It is also a part of UCSD's Center for Research on Biological Structure. The PI is Peter Arzberger. Co-investigators include Kim Baldridge, Chaitan Baru, Mark Ellisman, Michael Holst, J. Andrew McCammon, Andrew McCulloch, Anuska Mihailova, Philip Papadopoulos, and Wilfred Li, all of UCSD; Nathan Baker of Washington University; and Arthur Olson and Michel Sanner of The Scripps Research Institute.

Background Image: HIV-1 Protease: This enzyme, which is essential for HIV to reproduce, is the target of AIDS drugs known as protease inhibitors. Emerging mutant HIV strains are showing resistance to these drugs. The wild type protease is shown in green and the drug-resistant double mutant in purple. Analysis of the structural dynamics of these molecules, including the flaps that form the top of the binding site, and the "Ear to Cheek" region, have led researchers to a possible explanation for the mutant's drug resistance as well as a potential new target site for new drugs. Image credit: A. McCammon and A. Perryman, UCSD.



National Center for Microscopy & Imaging Research (NCMIR) Telescience Project

Creating the Framework for End-to-End Electron Tomography (ncmir.ucsd.edu and telescience.ucsd.edu)

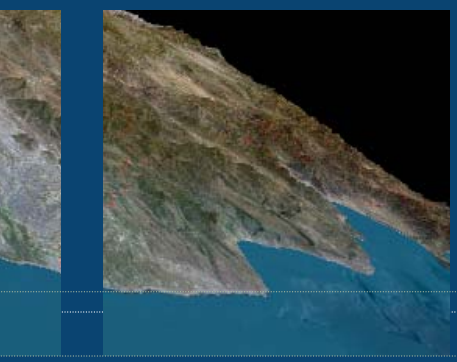
NCMIR developed Telescience as a comprehensive, platform-independent, Grid-enabled system that allows researchers to perform end-to-end electron tomography and large field light microscopy. Telescience is performed through the "Telescience Portal," a Web interface with a single username and password that allows biologists to access a suite of tools to manage this process. Advances in Telescience illustrate the benefits of a scalable and persistent infrastructure produced by integrating and sharing resources, technology, and experience of globally distributed partners. Users have access to

- Resource scheduling
- Remote instrumentation
- Parallel tomographic Grid-based reconstruction
- Visualization, segmentation, and image processing tools
- Heterogeneous, distributed file systems for data archiving
- Transparent deposition of data products into cellular structure databases (e.g., NCMIR's Cell Centered Database)
- Utilities for shared whiteboard image annotations and chatting among multiple sites

Telescience provides the biologist with the power of the computational Grid while masking the complexity of its administration.

Telescience has been selected as one of the driving applications for the Pacific Rim Applications and Grid Middleware Assembly (PRAGMA). This association has led to international collaborations with facilities that include

- Osaka University, which assists in the development of IPv6-enabled technologies to facilitate interactive remote control of the high-voltage electron microscopes in San Diego and the ultra-high-voltage electron microscope in Osaka
- Taiwan's National Center for High-Performance Computing (NCHC), which assists with the expansion of Telescience capabilities to visualize and analyze massive data
- Korea Basic Science Institute (KBSI), which is implementing Telescience technologies into their eScience program to facilitate remote use of their 1.25-million-volt, ultra-high-voltage electron microscope



The principal investigator is Mark Ellisman.

OptIPuter *A Powerful, Distributed Cyberinfrastructure to Support Data-Intensive Scientific Research and Collaboration (www.optiputer.net)*

With the advent of optical networking, the exponential growth rate in bandwidth has crossed over and will continue to exceed that of Moore's Law, enabling a new world view in which the central architectural element is not the computer platform but optical networking. As happened a decade ago in supercomputing, this transition is made possible by the use of parallelism. But this time, parallelism takes the form of multiple wavelengths of light, or lambdas, capable of traversing single optical fibers. Staying true to the analogy with supercomputers, the result is supernetworks.

In this context, the OptIPuter project is exploring a new architecture for distributed information infrastructure. This architecture is motivated by a number of emerging large-scale, shared, science and engineering systems. It aims to learn how, as George Gilder suggests, to "waste" bandwidth and storage so as to conserve "scarce" computing in this new world of literally inverted values.

OptIPuter, the name, is based on "opt" from optical networking, "IP" for Internet Protocol (from TCP/IP, a protocol enabling transmission of data between computers on the Internet), and "uter" leveraging the root of the word "computer."

OptIPuter is driven by close collaborations with two community systems: NSF's EarthScope and NIH's Biomedical Informatics Research Network (BIRN). Both are beginning to produce an accelerating flood of data stored in distributed, federated data repositories. The progress of their science has been hampered because the individual data objects (a 3-D brain image or terrain data set) are too large (Gigabytes in size) compared to the sizes that can be manipulated or visualized interactively over today's networks.

What these scientists require-and what OptIPuter is prototyping-are ultra-high-speed, predictable, "clear-channel" networks linking PC clusters, storage, and visualization systems, which will enable collaboration on massive amounts of previously uncorrelated data. We believe this will be possible by enabling lambda paths to form uncongested networks the full distance from source to destination.

OptIPuter was launched in 2002 and is funded by the National Science Foundation. This project is one of Cal-$(IT)^2$'s flagship research projects and includes the Biomedical Informatics Research Network as one of its driving applications. The PI is Larry Smarr, UCSD/Cal-$(IT)^2$. Co-PIs include Tom DeFanti and Jason Leigh, University of Illinois at Chicago; Mark Ellisman, UCSD; and Philip Papadopoulos, SDSC/Cal-$(IT)^2$.

Pacific Rim Applications & Grid Middleware Assembly (PRAGMA) and Pacific RIM Undergraduate Experiences (PRIME)

Building Sustained Collaboration & Advancing the Use of Grid Technologies in Applications (www.pragma-grid.net, prime.ucsd.edu)

PRAGMA is an open, institution-based organization, founded in 2002, to establish sustained collaborations and advance the use of grid technologies in applications among a community of investigators around the Pacific Rim. PRAGMA is based on several premises: the conduct of science is global; the grid promises to revolutionize science as much as networking has done to our daily activities; the grid is too difficult to use for most researchers. PRAGMA recognizes that constructing and using the grid to promote e-science is inherently a global, collaborative undertaking.

To accomplish its mission, PRAGMA conducts joint projects that develop and integrate grid middleware to advance applications, shares resources to create a testbed, and addresses scheduling and allocation issues across institutional and international boundaries. PRAGMA is committed to disseminating the results of its efforts to the broader community and works with regional and international groups to enhance overall grid infrastructure and promote global collaboration.

PRAGMA's accomplishments illustrate how the grid brings remote resources *(observational equipment, computers, data, and people)* together to one's local work environment. Examples range from controlling a microscope to understanding cell processes in the brain to monitoring the environment in national parks, from distributing computations that can lead to insights into drug discovery to moving files essential to high-energy physics experiments, and from conducting a global structural genomic experiment to rapidly deploying technology to assist the world in fighting the SARS outbreak.

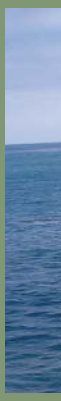
PRAGMA aims to inspire other international collaborations and promote new means to nurture, sustain, and expand those collaborations so that we as a global society can address critical issues and improve economic growth, quality of life, and the health of our planet.

PRAGMA is funded by an award from the Office of International Science and Engineering and the divisions of Shared Cyber infrastructure and of Biological Infrastructure of the NSF. The PI and co-PI are Peter Arzberger and Philip Papadopoulos, respectively.

The closely related Pacific RIM Experiences for undergraduates (PRIME: *prime.ucsd.edu*) activity began in April 2004 by major support from the NSF and contributing support from Cal-(IT)^2 to provide students a research experience at one of three PRAGMA sites: Cybermedia Center at Osaka, Japan; the National Center for High-Performance Computing in Hsinchu, Taiwan; and Monash University in Melbourne, Australia. Gabriele Wienhausen, Provost, Sixth College and Education layer leader is PI, and Linda Feldman, Academic Internship Program, and Peter Arzberger are co-PIs.

Interdisciplinary Bioinformatics Program

Preparing the Next Generation of Biomedical Researchers (bioinformatics.ucsd.edu)



Deciphering the genetic code of living organisms is dramatically changing our understanding of the natural world and promises to improve substantially the quality of human life. Recent advances in technology have led to the creation of a new interdisciplinary science-genomics. In simple terms, genomics is the reading and understanding of the blueprints for life. Understanding how genomes work requires sophisticated computer-based information-handling tools (bioinformatics) and new high-throughput technologies to understand the function of genes on a genome-wide scale (functional genomics).

The most pressing problem in the post-genome sequencing era is to understand the integrated functions of thousands of genes. Dealing with this problem requires an interdisciplinary research structure dedicated to developing intellectual and human capital in bioinformatics and genome science. The complexity of this new paradigm in biology, i.e., understanding the organization, evolution, and function of whole genomes rather than single genes, created the need for new sets of tools and human resources.

Future developments in genomics, and the applications that derive from genomics, will depend on scientific progress at the interface of three major disciplines: biology, engineering, and computer science. In addition to the scientific advances required to understand the functions of genomes, the accelerated growth of modern biology warrants revolutionary changes in academic curricula.

The Bioinformatics graduate program at UCSD draws upon the interdisciplinary expertise of affiliated faculty from seven participating departments and graduate programs:

- Bioengineering
- Biology
- Biomedical Sciences
- Chemistry and Biochemistry
- Computer Science and Engineering
- Mathematics
- Physics

UCSD offers an interdisciplinary specialization in Bioinformatics. Because the program's mission is central to DeGeM and its goal is to train the next generation of researchers, DeGeM has been supporting the development of the program by offering a fellowship each year to ensure UCSD and Cal-(IT)^2 jointly can attract the best students to this area. Shankar Subramaniam is the Director of the program. The program is also supported by an NIH graduate research training grant, for which Dr. Subramaniam is the PI and Director.

Background Image: The role of dynamics in the biological function of proteins and enzymes has been a topic of intense debate for decades. The use of vibrational spectroscopy, the most direct method to study dynamics, has been limited by the congestion of the protein vibrational spectrum. Floyd Romesberg, assistant professor of chemistry at The Scripps Research Institute, and his colleagues use cytochrome c as a model system to demonstrate that individual protein vibrations may be directly observed by the incorporation of an isotopically labeled amino acid, and that those labeled vibrations are sensitive probes of the protein environment. This volume-rendered image shows a model that was investigated with computational quantum chemical methods developed by Kim Baldridge in NBCR, for the Fe (II)- and Fe (III) cytochrome c prosthetic group, revealing the structure and specific vibrational properties of the FeP (DMS) complex, which were then compared to experimentally generated data.

People: AT CAL-(IT)²/UCSD

We organize our activities into projects, but we depend on people to conceive and execute them. Here are some of the key players who are achieving DeGeM's vision. Their research rides on the crest of the new wave of DeGeM science: integrative, multi-scale, and linked tightly with information technologies.

Other researchers associated with DeGeM can be found at *www.calit2.net/partners/academic_participants.html#biotech*, and their work will be highlighted in future brochures. We note the wide variety of expertise they represent, often engaged in other activities of Cal-(IT)².

Background Image: "Reconstruction of Neuronal Spiny Dendrite" Collaboration between NCMIR UCSD and Osaka University led to this imaging study of dendritic spines. Image created by combining electron tomography with high-resolution light microscopy.

John Wooley (DeGeM Layer Leader at UCSD)
medicine.ucsd.edu/pharmaco/jwooley.html

John Wooley is Associate Vice Chancellor for Research at UCSD, an adjunct professor in Pharmacology, Chemistry and Biochemistry, and a Strategic Advisor and Senior Fellow of SDSC. He received his Ph.D. degree in 1975 at The University of Chicago, working with Al Crewe and Robert Uretz in Biological Physics. Wooley created the first programs within the US federal government for funding research in bioinformatics and computational biology, and has been involved in strengthening the interface between computing and biology for more than a decade. For Cal-(IT)², Wooley directs the biology and biomedical layer or applications component, termed Digitally Enabled Genomic Medicine (DeGeM), a step in delivering personalized medicine in a wireless clinical setting. His current research involves bioinformatics and structural genomics, while his principle objectives at UCSD are to stimulate new research initiatives for large-scale, multidisciplinary challenges. He also collaborates in developing scientific applications of information technology and high-performance computing; creating industry-university collaborations; expanding applied life science opportunities, notably around drug discovery; and establishing a biotechnology and pharmacology science park on UCSD's health sciences campus.

Peter Arzberger
nbcr.ucsd.edu/~Arzberger.html

Peter Arzberger is the Director of the Life Sciences Initiative at UCSD and Deputy Leader for the Digitally Enabled Genomic Medicine layer. In addition, Arzberger is the Director of the National Biomedical Computation Resources *(nbcr.ucsd.edu)*, an NIH Research Resource that merges computing and information technologies to catalyze and facilitate biomedical research across a broad range of biological scales.

Arzberger is also involved with the Long Term Ecological Research (LTER: *www.lternet.edu*) community to facilitate interaction between ecologists and computational and information scientists. In addition, he is a member of a scientific advisory subcommittee for the Global Biodiversity Information Facility *(www.gbif.org)*.

Through these two efforts, Arzberger has become involved in international activities including policies on access to data produced from public funding *(dataaccess.ucsd.edu)*. Another international activity Arzberger helped co-establish is the Pacific Rim Application and Grid Middleware Assembly whose goal is to build sustained collaborations among researchers around the Pacific Rim by building applications on top of the emerging Grid hardware and software *(pragma.ucsd.edu)*.

Arzberger is former Executive Director of SDSC and NPACI, and a former NSF program officer of the Computational Biology Program in BIO and Program Director of the Statistics and Probability program in the Division of Mathematics Sciences, and served as Deputy NSF High Performance Computing and Communications Coordinator. He obtained his Ph.D. in mathematics in 1983 from Purdue University.

Kim Baldridge *www.sdsc.edu/~kimb*

Kim Baldridge is a professor in residence, Chemistry and Biochemistry, UCSD, as well as a professor of Theoretical Chemistry and Director of Computational Applications of the OCI-Institute at the University of Zurich. Baldridge has been an active computational scientist and key developer of a community code for quantum chemistry, GAMESS. She has a long track record of publications and working with students. She initiated and oversees the annual Marie Goeppert Mayer Interdisciplinary Symposium. She is active in both NBCR and the Center for Theoretical Biological Physics.

Helen Berman *www.rcsb.org*

Helen Berman is Board of Governors Professor of Chemistry and Chemical Biology at Rutgers University, the Director of the Protein Data Bank, and a Visiting Scholar in CRBS of UCSD. Berman received her Ph.D. in 1967 from the University of Pittsburgh. She has been active in the international crystallographic community. Very early on, she recognized the need for community resources such as the PDB and the Nucleic Acid Database, and the corresponding need to provide access to these resources using contemporary information technology. In 1996 she helped develop the Research Collaboratory for Structural Biology consortium, which consists of Rutgers, The State University of New Jersey; UCSD; and Center for Advanced Research in Biotechnology (CARB) of the University of Maryland Biotechnology Institute and the National Institute for Standards and Technology. She has also been an active proponent of creating standards for the description of the crystallographic experiment and is a co-author of the Macromolecular Crystallographic Information File (mmCIF), which formed the basis for the creation of a stable and extensible infrastructure.

Philip E. Bourne *www.sdsc.edu/pb*

Phil Bourne is a professor in the Department of Pharmacology at UCSD, Director of Integrative Biosciences at SDSC, Co-Director of the PDB, an adjunct professor at the Burnham Institute and the Keck Graduate Institute, the Immediate Past President of the International Society for Computational Biology, and an elected fellow of the American Medical Informatics Association. He is an Associate Editor of the journal *Bioinformatics*, a member of the Advisory Board of Biopolymers, and a long-standing member of the National Science Foundation and National Institutes of Health panels responsible for reviewing proposals relating to biological infrastructure. He is a member of the National Committee for Crystallography and past chairman of the International Union of Crystallography Computing Commission IUCrCC and past chairman of the American Crystallography Association Computing Committee.

Bourne's professional interests focus on bioinformatics and structural bioinformatics in particular. This implies algorithms, metalanguages, biological databases, biological query languages, and visualization with special interest in cell signaling and apoptosis.

He received his Ph.D. in 1979 from The Flinders University of South Australia. He is the Principal Investigator on the Systematic Protein Annotation and Modeling grant from NIH.

Edward A. Dennis *www-chem.ucsd.edu/Faculty/bios/dennis.html*

Edward Dennis is a UCSD professor in the departments of Chemistry & Biochemistry and in Pharmacology, the Principal Investigator on the LIPID Metabolites and Pathways Strategy (LIPID MAPS) glue grant, and Editor-in-Chief of the *Journal of Lipid Research*. He received his Ph.D. from Harvard University in 1968, and was appointed to the faculty of UCSD in 1970.

Dennis' laboratory is focused on understanding the regulation of lipid second messengers and signal transduction processes, especially the role of various phospholipases in their generation. Special attention is paid to the cytosolic, secreted, and membrane-bound phospholipase A2s responsible for the control of prostaglandin and leukotriene biosynthesis in macrophage cells. These are produced from arachidonic acid released upon cell stimulation. Their goal is to characterize and elucidate the regulatory mechanisms of various phospholipase A2s both *in vitro* and in the intact cell.

The Dennis laboratory also designs and synthesizes chemical inhibitors of phospholipase A2. Many different inhibitor classes have been developed and are being studied using *in vitro*, *ex vivo*, and *in vivo* animal models as part of the first TransMed grant and a UC Discover grant. In addition, there are studies of synthetic oxidized phospholipids and their role in LDL scavenger receptor uptake, autoantibody formation, and apoptosis, all of which are processes that lead to atherosclerosis. In summary, the laboratory uses organic synthetic approaches, enzyme kinetics, molecular biology, site-specific mutagenesis, cell and tissue culture, and mass spectometric techniques as well as traditional biochemical approaches in addressing phospholipase and membrane problems.

Mark Ellisman *ncmir.ucsd.edu, ncmir.ucsd.edu/biology.html*

Mark H. Ellisman is professor of Neuroscience and Bioengineering and the Director of the Center for Research in Biological Structure at UCSD, where he has taught since 1977. Ellisman directs the National Center for Microscopy and Imaging Research (NCMIR), an internationally acclaimed technology development center and research resource established by the National Institute of Health. He has received numerous awards, is the interdisciplinary coordinator for the National Partnership for Advanced Computational Infrastructure (NPACI), and leads the Neuroscience thrust for NPACI, which involves integration of brain research and advanced computing and communications technologies.

His scientific contributions include work on basic molecular and cellular mechanisms of the nervous system and development of advanced technologies in microscopy and computational biology. He is a pioneer in the development of three-dimensional light and electron microscopy and combined application of these image-acquisition tools and computational technologies to achieve greater understanding of the structure and function of the nervous system. His group was the first to introduce the idea of telemicroscopy by demonstrating network-enabled remote use and sharing of a high-energy electron microscope in 1992 and then developed practical systems now in use by researchers in the US and abroad.

He spearheaded the development and deployment of the Biomedical Informatics Research Network (BIRN) initiative and is the Principal Investigator of the BIRN Coordinating Center. He is also the Director of an organized research unit called the Center for Research on Biological Systems (CRBS: *www.crbs.ucsd.edu*).

Ellisman received his B.A. in 1970 in Biological Psychology from UC Berkeley, and his master's in Neurophysiology in 1974 and Ph.D. in Molecular, Cellular and Developmental Biology in 1976 from the University of Colorado at Boulder, where he studied under Keith R. Porter.

Eleazar Eskin *www.calit2.net/compbio/hap*

Eleazar Eskin is an assistant professor in residence in Computer Science and Engineering at UCSD. His research areas are Bioinformatics and Computational Biology-specifically, human variation analysis, transcriptional regulation, and protein analysis using machine-learning techniques. Eskin is involved in the HAP project, which provides tools for associating genetic variation to disease phenotypes *(www.calit2.net/compbio/hap)*.

Previously, Eleazar was a postdoctoral student at Hebrew University in the Computer Science Department under the supervision of Yoram Singer and Nir Friedman. He completed his Ph.D. in the Computer Science Department of Columbia University under the supervision of Sal Stolfo.

Trey Ideker *www-bioeng.ucsd.edu/faculty/area/ideker_lab*

Trey Ideker is working to develop large-scale, computer-aided models of biological signaling and regulatory pathways. New types of models, experimental strategies, and statistical frameworks are needed for integrating the enormous amount of data on mRNA expression, protein expression, and protein interactions arising in the wake of the Human Genome Project. These tools will be crucial to the success of Systems Biology, i.e., understanding biological systems as more than merely the sum of their parts.

Ideker received bachelor's and master's degrees from MIT in Electrical Engineering and Computer Science, where he was elected to the HKN Engineering Honor Society and awarded the Northern Telecom/BNR prize for his work in digital circuit design.

Encouraged by developments in the Human Genome Project, Ideker rapidly became interested in applying methods from computer science and engineering to the understanding of biological systems. Toward this goal, he obtained a Ph.D. in Molecular Biotechnology at the University of Washington and at the Institute for Systems Biology under Dr. Leroy Hood. He then moved to the Whitehead Institute for Biomedical Research, in Cambridge, Massachusetts, as the David Baltimore Fellow and Pfizer Fellow of Computational Biology. Ideker is assistant professor in the Department of Bioengineering at UCSD. He serves on the advisory board of Genstruct and the BioCyc Project, has been a Bioinformatics Lecturer for ISTR, Inc., and holds several patents in the fields of microarray analysis and systems biology.

Leslie Lenert *wiisard.org, preferences.ucsd.edu*

Leslie Lenert is Associate Professor of Medicine and Staff Physician at the Veterans Affairs San Diego Healthcare System. His educational background includes an M.D. from UCLA, an M.Sc. degree in Biomedical Informatics from Stanford University, and a postdoctoral fellowship in Clinical Pharmacology under Terry Blaschke and Lewis Sheiner at Stanford.

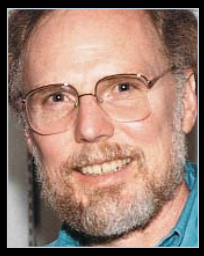
Lenert is a clinically active primary care physician. His research program has two general divisions: the Laboratory for Patient Informatics (LPI) and the Laboratory for Informatics in Terrorism Response (LITR). LPI focuses on the development of computer systems to improve patient and provider decision making and promote health behavior change. Ongoing work includes development of technologies to facilitate patient-physician collaboration in clinical decision making and behavioral change. LPI is also engaged in studying the use of small groups on the Web and integrating telephone counseling and Web technologies.

The LITR focuses on development and evaluation of technologies that enhance preparedness for response to terrorist attack. The WIISARD project (described above) is part of the LITR portfolio. A second project, Absence Causing Symptom Level Assessment Network (ASLAN), is developing and validating new approaches for syndromic surveillance. ASLAN will use automated telephone questionnaires to monitor symptoms that cause absenteeism in school children, providing early warning of bioterrorist attack and identifying children with undiagnosed health problems.

Lenert has published more than 80 original articles, received numerous awards, and is a member of the American College of Medical Informatics. He is member of the editorial boards of several medical informatics-related journals, including the *Journal of the American Medical Informatics Association* and the *Journal of Biomedical Informatics and Medical Decision Making*. He is also a member of the Agency for Healthcare Quality and Research's Healthcare Technology and Decision Sciences study section.

J. Andrew McCammon

mccammon.ucsd.edu



J. Andrew McCammon is the Joseph E. Mayer Professor, Department of Chemistry and Biochemistry. He is also professor in the departments of Pharmacology. He is also an Investigator in the Howard Hughes Medical Institute.

Most biochemical reactions occur in solution or at interfaces between lipid bilayers and aqueous phases. The McCammon group studies these reactions using principles of statistical mechanics, classical and quantum mechanical models, and homology models of proteins to gain insight into functions of biological macromolecules. His research focuses on rational drug design, molecular simulations, and questions in structural biology. His studies benefit from the excellent computing facilities to which we have access, including superb visualization tools, networks of high-performance personal computers, and parallel supercomputers.

McCammon is a co-investigator in the National Biomedical Computation Resource (NBCR) and is one of the founding members of the La Jolla Interfaces in Science Program for Graduate Students and Postdocs and of the NSF Center for Theoretical Biological Physics.

Andrew McCulloch

cardiome.ucsd.edu

Andrew McCulloch is professor in and Vice-Chair of the Department of Bioengineering. He received is Ph.D. from the University of Auckland in 1986 and is a member of the Whitaker Institute for Biomedical Engineering, and the Institute for Molecular Medicine.

McCulloch's Cardiac Mechanics Research Group studies the mechanics and electrical dynamics of the normal and diseased heart from molecular to organ scales. To understand normal cardiac development and the pathogenesis of heart failure and arrhythmia, the group uses *in silico* models to integrate *in vitro* experimental observations. These multi-scale simulations integrate systems models of intracellular networks into anatomically detailed continuum models of the whole heart, which are validated using *in vivo* measurements. The principal investigators are Andrew McCulloch and Jeffrey Omens, and the research is funded by the National Institutes of Health, the National Science Foundation, the American Heart Association, the National Space Biomedical Research Institute, and the Defense Advanced Research Projects Agency.

McCulloch is a co-investigator on the National Biomedical Computation Resource and is the lead advocate behind the community code Continuity.

Bernhard Ø. Palsson *systemsbiology.ucsd.edu*

Bernhard O. Palsson is a professor of Bioengineering and an adjunct professor of Medicine at UCSD. His research interests include areas of hematopoietic stem cell transplantation, cell culture technology, bioreactor design, gene transfer, cell separations, genome-scale *in silico* model building. and metabolic engineering. He received his Ph.D. from the University of Wisconsin in 1984.

His current research at UCSD focuses on 1) reconstruction of genome-scale biochemical reaction networks, 2) development of mathematical analysis procedures for genome-scale models, and 3) experimental verification of genome-scale models with emphasis on cellular metabolism and transcriptional regulation in E. coli and Yeast.

He founded or co-founded several biotechnology companies. The areas of focus include purging of occult tumor cells in autologous bone marrow transplants (ONCOSIS), instrumentation for high-throughput screening and *in situ* cell sorting and processing (CYNTELLECT), *in silico* biology (GENOMATICA), and tracing the genetic basis for common human diseases in the Icelandic population (Iceland Genomics Corporation).

Pavel Pevzner *www.cs.ucsd.edu/users/ppevzner*

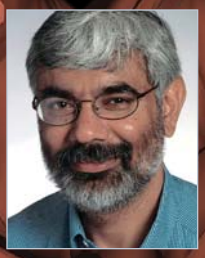
Pavel Pevzner is the Ronald R. Taylor Chair Professor of Computer Science in the Department of Computer Science and Engineering and an adjunct professor in the Department of Mathematics. He is Executive Editor of the *Journal of Computational Biology* and on the editorial boards of many bioinformatics, computer science, and mathematical journals. He is the co-founder and Chair of the Steering Committee of the Annual International Computational Molecular Biology Conference (RECOMB). He has published over 100 articles, patents, and software manuals. He received his Ph.D. from Moscow Institute of Physics and Technology in 1988 in mathematics and physics and subsequently moved into the emerging area of computational biology.

Pevzner has interest in many areas of research including fragment assembly in DNA sequencing; pattern discovery and regulatory genomics; genome rearrangements; optimization of DNA array manufacturing, and computational mass-spectrometry. His work has focused on algorithm development, and he has applied these new algorithms to addressing problems in molecular biology.

Nicholas Schork *psychiatry.ucsd.edu/faculty/nschork.html*

Nicholas Schork is a professor of psychiatry at UCSD. Prior to his appointment at UCSD, Schork was an associate professor of epidemiology and biostatistics at Case Western Reserve University and an associate professor of biostatistics at Harvard University. He was also formerly the Associate Director of the Program for Population Genetics at the Harvard School of Public Health and an adjunct associate staff scientist at The Jackson Laboratory. Between 1999 and 2000, Schork took a leave of absence to conduct research as the vice president of statistical genomics at the French Biotechnology company, Genset, where he helped guide efforts to construct the first high-density map of the human genome. Schork is internationally renowned for his work on theoretical and applied aspects of the genetic basis of multifactorial traits and conditions, has been selected as a member of a number of scientific journal editorial boards, is a frequent participant in U.S. National Institutes of Health-related steering committees and review boards, and has also been on the advisory board of five different companies. Schork has published over 150 scientific articles and book chapters on the analysis of complex, multifactorial traits and diseases. Schork earned a B.A. and an M.A. in Philosophy, an M.A. in Statistics, and a Ph.D. in Epidemiology all from the University of Michigan.

Shankar Subramaniam *genome.ucsd.edu/People*



Shankar Subramaniam is a professor of Bioengineering, Chemistry & Biochemistry, and Biology and Director of the Bioinformatics Graduate Program at UCSD. He also has adjunct professorships at the Salk Institute for Biological Studies and SDSC. He received his Ph.D. from the Indian Institute of Technology. His interests are in computational approaches for analysis of experimental data to understand the integrated functions of thousands of genes. Understanding how genomes work requires sophisticated computer-based information handling tools-bioinformatics and new high-throughput technologies for understanding the function of genes on a genome-wide scale (functional genomics) and cellular signaling networks. Projects that he is involved in include the Alliance for Cellular Signaling, the Biology Workbench, and LIPID MAPS.

Subramaniam has played a key role in raising national awareness for training and research in bioinformatics. He served as a member of the NIH Director's Advisory Committee on Bioinformatics, which resulted in the Biomedical Information Science and Technology Initiative report. The report recognized the dire need for trained professionals in bioinformatics and recommended launching a strong NIH funding initiative.

Opportunities for Students:

In addition to the Interdisciplinary Bioinformatics Program, there are other opportunities for the support of students.

Cal-(IT)^2 supports undergraduates for summer scholarships to do research. See *www.calit2.net/students/ugrad/index.html*

In addition, Sixth College together with PRAGMA have begun an international summer research program. Details can be found at *prime.ucsd.edu*.

In addition, Cal-(IT)^2 supports graduate students via fellowships. See *www.calit2.net/students/grad/index.html*.

There is a UCSD Bioinformatics Research Projects Website at *www.cs.ucsd.edu/~eeskin/projects*, established by Eleazar Eskin, that links faculty, projects, and undergraduate students together, to provide the students with a research experience in bioinformatics.

Final Word:

Our intention of highlighting these activities and people under the DeGeM banner is to promote a broad discussion with researchers and students, with industry, with private foundations, and with government agencies about the exciting future for digitally enabled genomic medicine.

We openly invite your comments, your constructive energies to enhance our activities and our accomplishments, and your involvement.

If you find the activities presented in this brochure of interest, we encourage you to contact the layer leader, John Wooley *(jwooley@ucsd.edu)*, the deputy layer leader, Peter Arzberger *(parzberg@ucsd.edu)*, the investigators or the people within DeGeM.

www.calit2.net

published September 2004