

D I V E R S E D I S C I P L I N E S , O N E C O M M U N I T Y

BiomedicalComputation

REVIEW

Ramping Up To Multiscale

Taking biomedical modeling
to the next level

PLUS:

THE FEMALE FACTOR

Is the gender gap in computer
science carrying over to
biomedical computing?



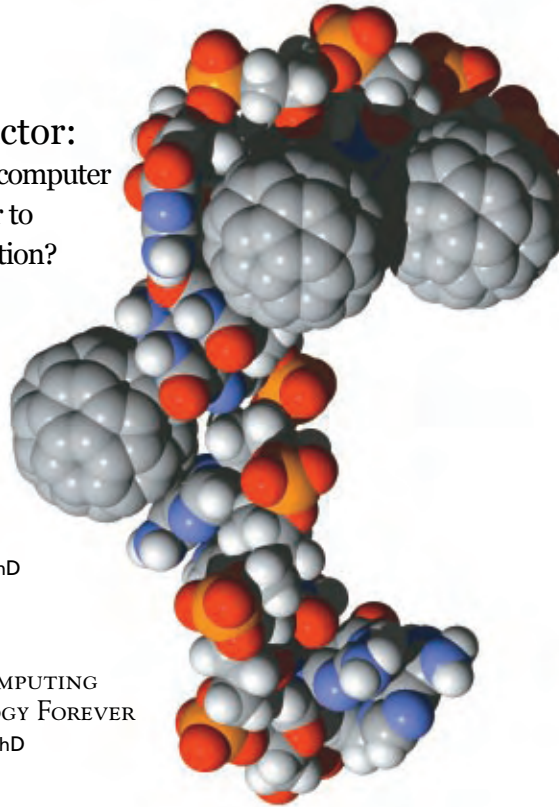
FEATURES

8 The Female Factor: Is the gender gap in computer science carrying over to biomedical computation?

BY KRISTIN COBB, PhD

19 Ramping Up to Multiscale: Taking biomedical modeling to the next level

BY DANA MACKENZIE, PhD



DEPARTMENTS

1 FROM THE EDITOR: COMPUTING HAS CHANGED BIOLOGY FOREVER

BY JOHN C. WOOLEY, PhD

2 NEWS BYTES

- Simulated Buckyballs Bind to DNA BY HANNAH HICKEY
- Where Proteins Go to Work BY LINLEY ERIN HALL
- The Six Faces of E.Coli BY HANNAH HICKEY
- A Powerful Model of Relaxation BY KRISTIN COBB, PhD
- Noisy Genes BY KRISTIN COBB, PhD
- Smoking Addiction Explained BY LINLEY ERIN HALL

27 UNDER THE HOOD: REPRESENTING ROTATIONS WITH QUATERNIONS

BY JAMES COBURN, MS, AND JJ TREY CRISCO, PhD

28 FEATURED LAB: THE INSTITUTE FOR SYSTEMS BIOLOGY

BY KRISTIN COBB, PhD

30 SEEING SCIENCE: DECODING THE DECODER BY KATHARINE MILLER

ON THE COVER:

COVER ART BY RACHEL C. JONES OF AFFILIATED DESIGN

IMAGES:

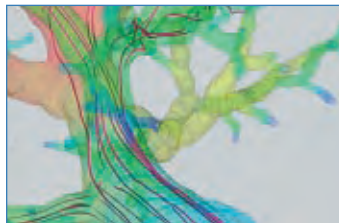
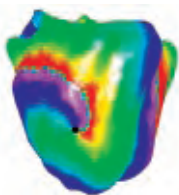
TORSO DIPOLE COURTESY OF A J PULLAN, M L BUIST, AND L K CHENG.

HEART SHOWING FIBER DIRECTIONS, COURTESY OF PETER HUNTER, PhD, BIOENGINEERING INSTITUTE, THE UNIVERSITY OF AUCKLAND, NEW ZEALAND

CONFOCAL IMAGES OF CARDIAC MYOCYTES, COURTESY OF DR. SETH L. ROBIA AND DR. JEFFERY W. WALKER, UNIVERSITY OF WISCONSIN, DEPARTMENT OF PHYSIOLOGY.

CARDIAC MYOCYTES FROM TANIWAKI ET AL., 2006, CELL AND TISSUE RESEARCH, COURTESY OF RENATO MORTARA. WITH KIND PERMISSION OF SPRINGER SCIENCE AND BUSINESS MEDIA

THIN FILAMENT FROM XU ET AL., 1999, BIOPHYSICAL J. 77, 985-992, COURTESY OF WILLIAM LEHMAN, PHD., PROFESSOR OF PHYSIOLOGY AND BIOPHYSICS, BOSTON UNIVERSITY.



Spring 2006

Volume 2, Issue 2

ISSN 1557-3192

Executive Editor

David Paik, PhD

Managing Editor

Katharine Miller

Science Writers

Katharine Miller

Kristin Cobb, PhD

Louisa Dalton

Hannah Hickey

Dana McKenzie, PhD

Linley Erin Hall

Community Contributors

John C. Wooley, PhD

James Coburn, PhD

JJ Trey Crisco, PhD

Layout and Design

Affiliated Design

Printing

Advanced Printing

Editorial Advisory Board

Russ Altman, MD, PhD

Brian Athey, PhD

Andrea Califano, PhD

Valerie Daggett, PhD

Scott Delp, PhD

Eric Jakobsson, PhD

Ron Kikinis, MD

Isaac Kohane, MD, PhD

Mark Musen, MD, PhD

Tamar Schlick, PhD

Jeanette Schmidt, PhD

Michael Sherman

Arthur Toga, PhD

Shoshana Wodak, PhD

John C. Wooley, PhD

**For general inquiries,
subscriptions, or letters to the editor,
visit our website at
www.biomedicalcomputationreview.org**

Office

Biomedical Computation Review
Stanford University
318 Campus Drive
Clark Center Room S170
Stanford, CA 94305-5444

Biomedical Computation Review is published quarterly by Symbios National Center for Biomedical Computing and supported by the National Institutes of Health through the NIH Roadmap for Medical Research Grant U54 GM072970. Information on the National Centers for Biomedical Computing can be obtained from <http://nihroadmap.nih.gov/bioinformatics>. The NIH program and science officers for Symbios are:

Peter Lyster, PhD (NIGMS)

Jennie Larkin, PhD (NHLBI)

Jennifer Couch, PhD (NCI)

Semahat Demir, PhD (NSF)

Charles Friedman, PhD (NLM)

Jerry Li, MD, PhD (NIGMS)

Wen Masters, PhD (NSF)

Richard Morris, PhD (NIAID)

Grace Peng, PhD (NIBIB)

David Thomassen, PhD (DOE)

Ronald J. White, PhD (NASA/USRA)

BY JOHN C. WOOLEY, PhD

Computing Has Changed Biology Forever

In 1991, a prescient editorial in *Nature* by Harvard's **Walter Gilbert, PhD**, ("Towards a paradigm shift in biology") included these observations on the utility and impact of computing:

"The new paradigm now emerging is that all the 'genes' will be known (in the sense of being resident in databases available electronically), and that the starting point of any biological investigator will be theoretical. ... We must hook our individual computers in the worldwide network that gives us access to daily changes in the database. ... The programs that display and analyze the material for us must be improved—and we must learn how to use them more effectively."

Today, Gilbert's paradigm shift at the interface of biology and computation is essentially complete: Genomic data allow researchers to start their investigations through a theoretical approach—an analysis of sequence databases; computer literacy is an essential part of the biologist's toolkit; data resources are available anywhere anytime; and software for biologists is becoming more widely available and useable.

Another emerging idea for biology is third party open access and standards (such as the open development efforts that extended Linux) to establish annotation for the vast amounts of data emerging. No single group will ever be able to annotate the data arising from the ever accelerating pace of genome sequencing, let alone that from metagenomics (Venter et al., 2004). Subsequent automated re-annotation following advances in biological understanding must be a feature of knowledge management. Similarly, the development of sophisticated computational methods for predicting function is needed to refine experiments in functional genomics and make explicit the information flowing from high throughput sequencing (Friedberg et al., 2006).

An exclamation mark for the introduction of computational approaches to biology has recently appeared: the 2005 report on the frontier at the interface, which summarizes a study of the National Research Council (*Catalyzing Inquiry at the Interface of Computing and*

Biology). While the field has grown so much in the past decade that no study could be inclusive, the NRC report is a comprehensive review of the history and current status of the field. The report, pointing also to the future, was conducted independent of the interests of individual agencies and subdisciplines, and should build support for research on the frontier.

Catalyzing Inquiry documents the paradigm shift through reviewing the accomplishments, opportunities and challenges at the interface; provides further sup-

An exclamation mark for the introduction of computational approaches to biology has recently appeared: the 2005 report on the frontier at the interface.

port for the commitment envisioned by Gilbert for advancing software for quantitative biological research, such as the efforts of the National Centers for Biomedical Computation and the computing-based NIH Research Resources; and points to the importance of programs for bioinformatics education for extending the impact of computing.

In 1998, Stanford professor **Michael Levitt, PhD**, reflected that "computing has changed biology forever, even if most biologists don't know it yet." Today, perhaps, they do. □

REFERENCES

Catalyzing Inquiry at the Interface of Computing and Biology, 2005 report of the National Research Council of the National Academies, National Academy Press, Washington, DC. <http://www.doegenomestolife.org/pubs/NRCCComputingandBiology/index.shtml>. Editors: John C. Wooley and Herbert S. Lin.

Friedberg, I., Harder, T., Godzik, A. (2006) "A Protein Function Annotation Metaserver," *Nucleic Acids Research*, in press.

Gilbert, W. (1991) "Towards a paradigm shift in biology." *Nature* 349: 99.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rush, D. (2004) "Environmental Genome Shotgun Sequencing of the Sargasso Sea." *Science* 304: 66-74.

DETAILS

John C. Wooley, PhD, is Associate Vice Chancellor of Research at the University of California, San Diego and Senior Fellow of the San Diego Supercomputer Center.

NewsBytes

Simulated Buckyballs Bind to DNA

Recent research illustrates a nightmare scenario for nanotechnology: simulated particles called buckyballs eagerly glomming onto nearby DNA. The study, published in *Biophysical Journal* in December 2005, has been widely read as a warning against the use of such materials for drug delivery, or any other purpose that could release them into the environment.

“If [buckyballs] can get into the cell, and into the nucleus, then they look like they have a significant impact on the DNA,” says co-author **Peter**

Cummings, PhD, professor of chemical engineering at Vanderbilt University in Nashville, Tennessee, and director of the Nanomaterials Theory Institute at Oak Ridge National Laboratory.

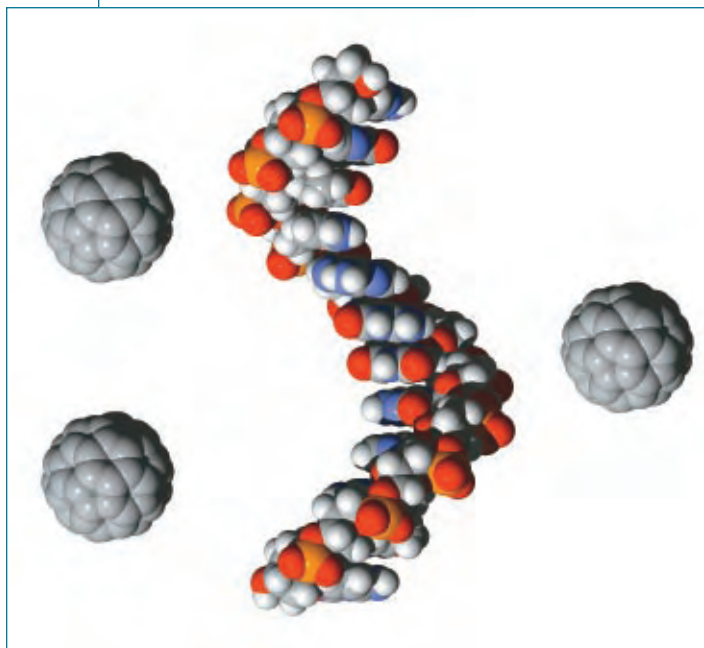
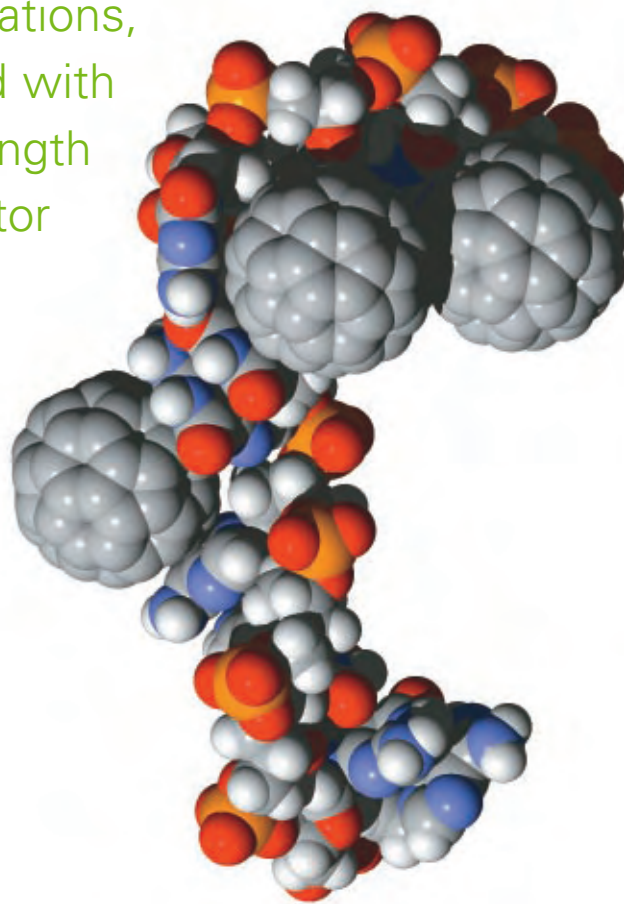
But because buckyballs’ ability to enter the cell nucleus is by no means certain, Cummings emphasizes it’s too early to suggest that buckyballs are unsafe. “This [simulation] is showing a possibility of what buckyballs could do,” Cummings says. “Now it’s worth investigating to find out if they can actually get into the cell and if they can do this kind of damage.”

Buckyballs—also known as buckminsterfullerenes—are hollow, soccer-

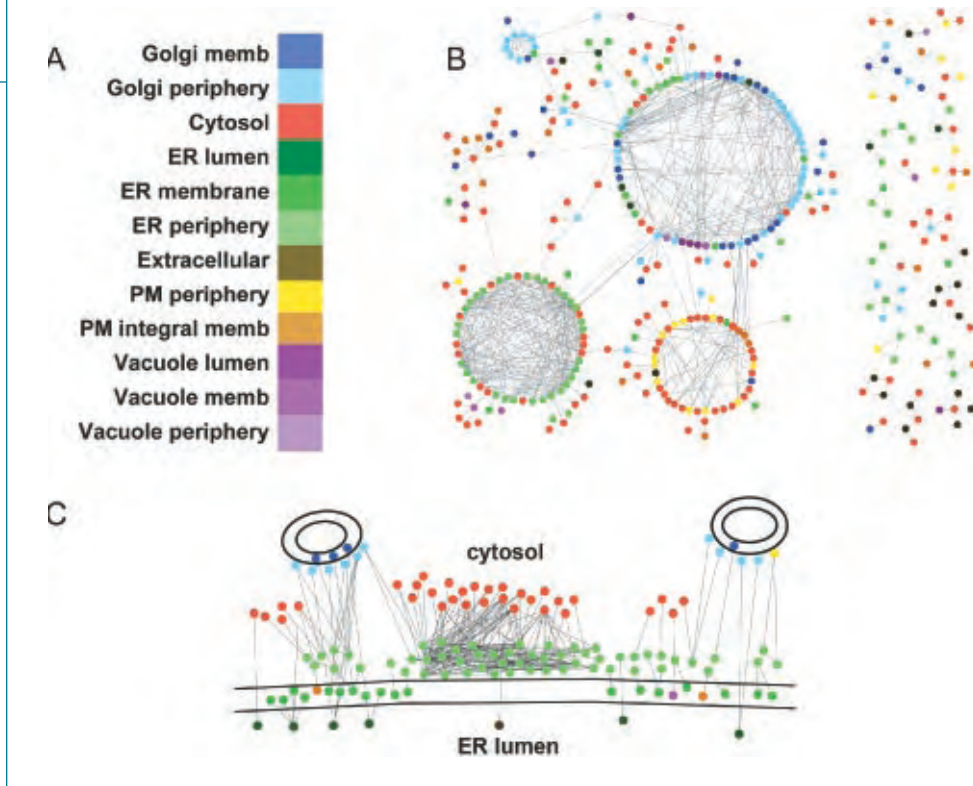
ball-shaped carbon molecules that, researchers believe, have the potential to transmit electricity or deliver drugs to targets inside the human body. Safety concerns were raised by a 2004 experiment that detected buckyballs accumulating in the brain tissue of fish. This finding prompted Cummings’ group to investigate buckyballs’ behavior inside cells.

Cummings and collaborators at Oak Ridge National Laboratory ran computer simulations of buckyballs placed in saline solution near a short strand of DNA. Within two nanoseconds, the buckyballs either stuck to the free end of the DNA or lodged into the minor

According to the model calculations, buckyballs form a strong bond with DNA—comparable to the strength of a drug attaching to a receptor and four times the binding energy of one buckyball to another buckyball.



Simulated buckyballs start near a DNA strand (left) and quickly bind to hydrophobic sections of the genetic material (above). The twisted DNA-buckyball compound remains stable for the rest of the simulation. Courtesy of Xiongce Zhao, Oak Ridge National Laboratory.



Hallett and colleagues used protein-protein interactions to help predict the location of yeast proteins in the cell. Here the proteins of the secretory pathway (B) and endoplasmic reticulum (C) are colored according to their location; lines represent interactions.

groove between the two strands. Single-stranded DNA tended to wrap around the buckyballs, dramatically changing the DNA's shape. On confronting a damaged piece of DNA, the buckyball wedged itself into the gap created by the tear.

These scenarios suggest that nanoscale materials such as buckyballs could interfere with DNA replication, transcription and repair. Such disruptions might cause long-term damage, including heritable mutations and cancer.

Scientists had predicted buckyballs would be harmless because they are hydrophobic, or water-hating. The nanoparticles were expected to bind to one another and "clump out" of solution. But it appears that inside a cell's nucleus, buckyballs tend to latch onto hydrophobic sections of DNA molecules, rather than onto one another.

According to the model calculations, buckyballs form a strong bond with DNA (in the range -27 to -42 kcal/mol). This is comparable to the strength of a drug attaching to a receptor and four times the binding energy of one buckyball to another buckyball.

Cummings and other scientists caution, however, that these results must be verified in experiments

before sounding the alarm.

"This paper adds little to the debate over what might happen in the physiological milieu," comments **Martin Chaplin, PhD**, an expert in water clustering at London South Bank University in the United Kingdom. The initial distance may be small enough that a dehydrating transition drew the buckyballs to the DNA, Chaplin says. He also questions representing a buckyball in solution exhibiting no electric charges.

To address these concerns, Cummings is using x-ray diffraction to study actual buckyballs and DNA. He plans to run another simulation incorporating the same fluid that he uses in the experiment.

—*Hannah Hickey*

Where Proteins Go To Work

Joe works in a factory; Jane works in a hospital; protein X works in the Golgi apparatus. Just as one might guess a worker's job by knowing where he or she is employed, biologists can guess a protein's function by knowing where it does its job—whether in or near the cell membrane, the endoplasmic reticulum or the Golgi apparatus—some of the important job sites inside a cell.

Determining thousands of proteins' correct cellular addresses is a daunting task. But a new yeast model takes a pretty good stab at predicting which proteins will wind up in 18 possible destinations inside this single-celled organism. The model is described in the November 2005 issue of *PLoS Computational Biology*.

"The trafficking and localization of proteins are very fundamental questions in biology," says **Michael Hallett, PhD**, a professor at the McGill Centre for Bioinformatics at McGill University in Montreal. But the places where 30 to 50 percent of all cellular proteins settle down to do their tasks are unknown. To get a better handle on this question, Hallett and his colleagues and graduate students

"People have done a lot of analysis using protein subcellular localization to predict protein-protein interactions. This work turns that around to good effect," says **Mark Gerstein**.

at McGill created the Protein Subcellular Localization Tool 2, or PSLT2.

PSLT2 is composed of three modules that predict where a protein will go. The *motif* module makes predictions based on the presence of particular sequences of amino acids that suggest a protein's function—a good indication of where it belongs in the cell. The *targeting* module relies on sequences that act like a known zip code, indicating where the protein should end up—such

ment, in its membrane, or associated with its surface. The model places the proteins into 18 sub-compartments correctly 83 percent of the time.

The ability to determine sub-compartments is new to this model. “When we use classical techniques for finding the localization of a protein [in, for example, the endoplasmic reticulum (ER)], we can't use them to tell if a protein is in the ER membrane, in the cytosol, or on the periphery,” Hallett

The Six Faces of *E. Coli*

Biologists' favorite bacterium grows almost anywhere—from the human gut to the pounding surf. But *E. coli*'s remarkable adaptability apparently stems from being predictable rather than accommodating. In a recent computer simulation, thousands of environments provoked only a handful of shifts in the microbe's physiology. The work was published in *Proceedings of the National Academy of Sciences* in December 2005.

“A network comprised of thousands of molecules, in response to a myriad of inputs, takes on relatively few overall responses,” says Bernhard Palsson.

as mitochondrial targeting peptides and transmembrane domains. And the *interaction* module concerns itself with the protein's likely comrades—the other proteins it associates with when doing its task. If protein A always interacts with protein B, and B has a known location in the cell, then A must be active in that vicinity as well.

Each module can individually predict the localization of a protein using Bayesian methods. The combination of the three modules improves the prediction when proteins lack motif and interaction data or traffic through multiple compartments.

For the entire yeast genome, the new tool predicts in which of nine compartments a protein is located with at least 72 percent accuracy. These compartments are mostly organelles but also include the cytosol and cell membrane. PSLT2 also predicts proteins' sub-compartmentalization—whether they are inside the compart-

says. “We need a computational method to pin down where the protein is in the organelle.”

The computational model's predictions compared well with databases from two high-throughput laboratory experiments, but they didn't always agree; Hallett and colleagues suggest that the model and two databases should be used in parallel as checks on each other.

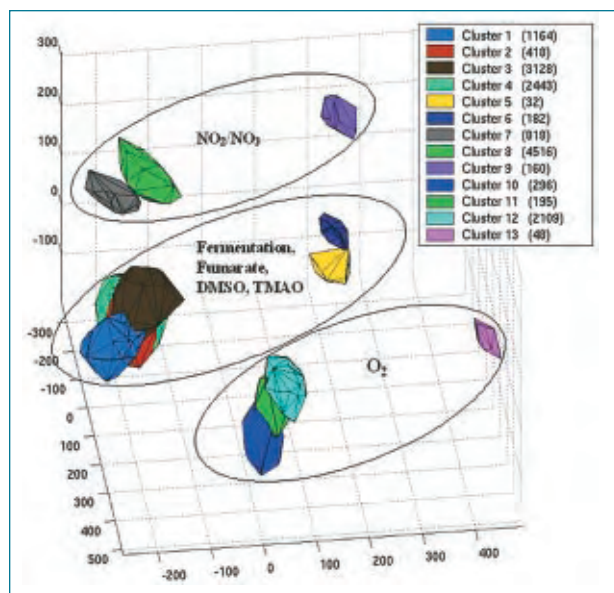
According to **Mark Gerstein, PhD**, an associate professor of biomedical informatics at Yale, the paper goes beyond what has been done before. “In particular,” he says, “people have done a lot of analysis using protein subcellular localization to predict protein-protein interactions. This work turns that around to good effect.”

—Linley Erin Hall

A map of possible states for *E. coli* metabolism. The axes represent the Hamming distance—a mathematical comparison of the output from different simulations. The closer two points, the more similar are those results. The ovals show the terminal electron acceptor for different types of respiration: NO₂/NO₃ for anaerobic; O₂ for aerobic; and fumarate or DMSO or TMAO for anaerobic fermentation. Courtesy of Christian Barrett, UCSD.

“A network comprised of thousands of molecules, in response to a myriad of inputs, takes on relatively few overall responses,” says senior author **Bernhard Palsson, PhD**, professor of bioengineering at the University of California, San Diego. The systems biology study of *E. coli* metabolism might help scientists understand how cells function and adapt to different environments.

To simulate *E. coli*'s environment, Palsson and his colleagues first drew up a list of nutrients that could meet the



microbe's needs—carbon, nitrogen, sulfur, etc. From this, they generated an exhaustive list of media that could support its growth. Then they wrote mathematical algorithms—based on 1,010 genes—for each step in *E. coli*'s well-understood metabolic process.

Combining the different inputs with these mathematical algorithms, they “grew” *E. coli* in 108,728 hypothetical simulated Petri dishes, of which 15,580 nurtured bacteria growth. Each of these *in silico* cultures produced a simulated gene expression profile, which researchers visualized in 3-D using a statistical tool known as principal component analysis.

The 3-D space was mostly empty: physiological outcomes appeared as thirteen clusters organized into six groups. Cells based their metabolic decisions largely on two factors: the availability of glucose as an energy source; and the identity of the terminal electron receptor(s)—the molecules that dictate whether the cell carries out aerobic respiration, anaerobic respiration, or fermentation. These responses are reasonably similar to laboratory experiments, Pálsson says, but he was surprised by the limited scope of all possible responses.

The researchers chose to study *E. coli* because it has the best-characterized DNA on the planet, but the technique could apply to other organisms. For example, ecologists might map microbial communities in an ounce of soil to see how hundreds of microbes' metabolisms interact. And engineers might use the technique to design whole bacterial ecosystems for useful tasks, such as eating toxic waste.

According to **Costas Maranas, PhD**, professor of chemical engineering at Pennsylvania State University, the study will help “to flesh out dominant organizing principles for complex systems.” In addition, he says, “One could look at whether the dominant behaviors that they have elucidated will hold under different kinds of perturbations, [such as] genetic perturbations.”

But the larger question of how all the complexity in the *E. coli* genome

results in only a few metabolic activities, Pálsson says, “is something that we still have to study, and understand.”

—**Hannah Hickey**

A Powerful Model of Relaxation

When a heart beats, millions of muscle cells contract in unison to pump blood to the body; then they relax, allowing the heart to refill. Though scientists have carefully characterized the mechanisms that govern contraction, they are less certain about the dynamics of relaxation. But a new mathematical model of calcium ion concentration in cardiac muscle—published in March 2006 in *Biophysical Journal*—has resolved at least one controversy.

“There’s been a lot of emphasis on contraction, because it’s the first thing you measure experimentally,” says

the cell reaches a critical length or tension), rather than on biochemical factors (a drop in calcium levels).

Smith and his colleagues combed the literature and found decades worth of experimental data (from humans, chickens, rats, mice, ferrets, rabbits, cows and cats) on calcium concentration and binding, as well as cell velocity, length, and tension during a heart beat. They combined these diverse data into a series of mathematical equations that simulate cellular contraction and relaxation. Then they simulated the tension changes in the beat of a heart cell—and found that their predictions closely approximated tension changes measured in the lab (data that had not been used to build the model).

Their simulation also showed that cell relaxation depends predominantly on the drop in calcium levels. “In some

“Often models get published that are very limited in scope, because authors are only interested in fitting their particular dataset. But these authors tried to match a diverse set of data,” says John Jeremy Rice.

Nicolas Smith, PhD, senior lecturer in the Bioengineering Institute and Department of Engineering Science at the University of Auckland in New Zealand. “But it’s just as important that the heart relaxes. We wanted to be very clear that we were characterizing the relaxation properties just as well as the contraction properties in this model.”

Here’s what a heartbeat looks like from within a cell: An electric signal spurs the release of calcium ions, which bind to motor proteins and activate contraction; then, the calcium ions are pumped away, and the cell relaxes. The rise in calcium clearly governs contraction, but scientists still debate the key trigger for relaxation. Some have suggested that relaxation depends more heavily on mechanical factors (when

ways this is less exciting than more esoteric ideas of length dependence and tension dependence, because it’s actually quite simple. But it does clear up a lot of the debate,” Smith says.

Smith and his colleagues are extending their model to study life-threatening biochemical changes that arise during ischemic heart disease (where oxygen is not getting to the heart). In ischemia, heart tissue becomes acidic, which wreaks havoc on calcium signaling. An unchecked overload of calcium will cause the heart to perpetually contract—a deadly deficiency of relaxation.

The model is limited to the cellular level, Smith notes, as modeling at the molecular or atomic scale would take too much computing power. But, he adds, “Because of the way we for-

mulated it, it would be absolutely clear how we would interface with a much more detailed protein model.” The model can also be embedded into tissue-level and whole-heart models of contraction.

“This paper is unique because the authors searched the literature pretty extensively to come up with the estimates for different muscle responses,” comments **John Jeremy Rice, PhD**, a researcher in the Functional Genomics and Systems Biology Group at IBM’s T.J. Watson Research Center in New York. “Often models get published that are very limited in scope, because authors

are only interested in fitting their particular dataset. But these authors tried to match a diverse set of data.”

—**Kristin Cobb**

Noisy Genes

Genetically identical cells or organisms grown in identical environments will differ phenotypically, because—even with a common script—gene expression is inherently variable, or noisy.

Such noise is counter-intuitive to many molecular biologists, who would expect gene regulation—the process that shapes all life—to run as precisely as a Swiss watch, says **Jeff Hasty, PhD**, professor of bioengineering at the University of California, San Diego.

Hasty and his colleagues are trying to expose the biological origins of this variability. In the December 22, 2005 issue of *Nature*, they report their latest finding in yeast cells. Using a combined

chromatin unwinding, nucleoside binding, and mRNA degradation—were stochastic (random) in nature. Extrinsic noise affects multiple genes within one or more cells, for example fluctuations in environmental conditions or in a cell’s global transcription or translation machinery.

“If you could classify the noise into these two different types, it gives you a handle on what might be causing the noise,” Hasty explains.

His team engineered yeast cells with one to five copies of the gene for an easily quantified green fluorescence protein (GFP), and its promoter. As expected, the cells with five copies lit up five times as brightly on average as the cells with one copy. More interestingly, fluctuations in the signals of the different strains were almost completely correlated, whether there were five gene copies or one, suggesting that extrinsic sources of variability dominate—which agrees with findings from other groups, in different experimental systems.

Hasty’s team then tried to pinpoint the biological sources of this extrinsic vari-

Variability in gene expression is largely due to cells being in slightly different phases of growth and division.

ability with computer simulation. Early models that included fancy terms for common transcription or environmental factors, “didn’t fit quite right,” Hasty says. Then they tried something more obvious. They started with the one source of (extrinsic) variability that has to be there: the oscillation in gene expression that arises naturally during the cell cycle.

They built a completely deterministic (non-random) mathematical model of population dynamics coupled with gene expression. In their model, virtual yeast cells, in slightly different phases, grow at a fixed rate to a particular size, and then bud off smaller daughter cells; cells in different stages of the cell cycle produce differing

ability with computer simulation. Early models that included fancy terms for common transcription or environmental factors, “didn’t fit quite right,” Hasty says. Then they tried something more obvious. They started with the one source of (extrinsic) variability that has to be there: the oscillation in gene expression that arises naturally during the cell cycle.

They built a completely deterministic (non-random) mathematical model of population dynamics coupled with gene expression. In their model, virtual yeast cells, in slightly different phases, grow at a fixed rate to a particular size, and then bud off smaller daughter cells; cells in different stages of the cell cycle produce differing



Identical twins never look exactly alike, despite having identical genotypes. Hasty and his colleagues at UCSD are trying to tease out the biological origins of this kind of “noise.”

amounts of GFP. This model predicted variation in fluorescence production similar to their experimental data. It also predicted a “noise floor”—a lower limit for variation, which was verified in tests of other highly expressed yeast genes.

“It turned out to be a much simpler answer than we were thinking,” Hasty says.

A few added “bells and whistles” improved the model fit—for example, adding random variation in the time to division and the size of the daughter cells—but the core of the model was still mostly deterministic variation.

“Here’s another piece that expands our understanding of the factors that contribute to variability in gene expression,” says **Jim Collins, PhD**, professor of biomedical engineering at Boston University. “This sets us up nicely as a community to begin to understand how the cell actually deals with this variability.”

—**Kristin Cobb**

Smoking Addiction Explained

To non-smokers, the nasty-smelling habit is inexplicable. But now neuro-computational researchers have developed a hypothetical model to explain how nicotine produces addictive behavior. The first computational model of nicotine addiction is described in the

A few things about nicotine addiction were already known. First, nicotine stimulates specific receptors on neurons that produce dopamine, causing them to produce more of this key neurotransmitter for motivation, reward and learning. Second, dopamine modifies learning in the circuits that are responsible for making choices. And third, nicotine results in addictive behavior. A wealth of behavioral data, including experiments in which rats are trained to self-administer nicotine, indicates that once a rat (or human) starts the habit, it will choose to continue. It will even learn to navigate complex mazes or press levers to obtain nicotine.

The model developed by Gutkin and his colleagues examines how a hypothetical animal “behaves” when given two choices: to smoke or not to smoke. “We found that simulation of both the biophysical effects of nicotine at the neuronal level and the subsequent effect on learning reproduced certain aspects of the behavioral data.”

In the biophysical part of the model, the researchers developed a simple kinetic scheme for how a population of neuronal receptors would react to nicotine on three different time scales: short-term stimulation of nicotinic receptors, leading to a burst



researchers asked, essentially, will this smoker choose to smoke again. They found that—especially after the second session—the

model smoker consistently keeps smoking. The researchers suggest that the dopamine message (“nicotine is rewarding”) gets locked in neurally—and becomes hard to unlearn. Smokers continue to seek out nicotine even when the opponent process reduces the amount of dopamine in the system and the pleasure goes out of smoking. Moreover, when nicotine is withdrawn from a smoker, the researchers say, the dopamine-lowering opponent process is still in effect. As a result, dopamine levels drop below normal, and it’s hard to find pleasure in anything.

The researchers were most pleased by the observation that nicotine sensitizes the system to the dopamine response. That is, the response to the second dose of nicotine is bigger than to the first, and has a significant effect on the choices made (to smoke or not smoke), whereas the first dose did not—presumably because of the learning effect on the decision maker. Likewise, a smoker who quits and starts again does not start from zero, but learns to self-administer nicotine faster the second time

Nicotine causes the neural circuits that control behavioral choice to change in a way that locks in smoking-related behaviors—making them difficult to unlearn.

January 24, 2006, issue of the *Proceedings of the National Academy of Sciences*.

“We wanted to tease out different parts of addiction,” says lead author **Boris Gutkin, PhD**, a research scientist at the Pasteur Institute in Paris, France. “The goal is to put behavioral and neural information together and to see how the behavioral effects are produced by the neural effects.”

in the dopamine signal; medium-term production of more receptors, which further increases the dopamine signal; and long-term opponency, a process that attempts to bring the system back into balance by decreasing the amount of dopamine.

Combining these dopamine changes with a module that gave smokers a choice after each smoking session, the

around. “So this was an effect we got for free,” says Gutkin. “We did not set out to *a priori* capture this effect, but found it inherent in the behavior of the model.”

“Gotten-for-free” effects are the stuff that makes us modelers especially happy,” says Gutkin, “they are our intellectual lolly-pops.”

—**Linley Erin Hall** □

THE

Female Factor:

Is the
gender gap
in computer
science
carrying
over to
biomedical
computing?

BY KRISTIN COBB PhD

The field of computer science has seen some wild ups and downs over the last twenty years. And the roller-coaster ride has been reflected in the fluctuating numbers of undergraduate computer science majors nationwide. Yet, throughout that time, the percentage of female computer science majors followed a steady, downhill trend—from 37 percent in 1985 to around 28 percent today (NSF). The drop stands in stark contrast to the rising proportion of women in other science and engineering disciplines. For example, women’s representation in biology and medicine has soared in the past decade, surpassing 50 percent in undergraduate biology in 1996 and in medicine in 2004—despite the male-domination of medicine just a generation ago (NSF; Association of American Medical Colleges).

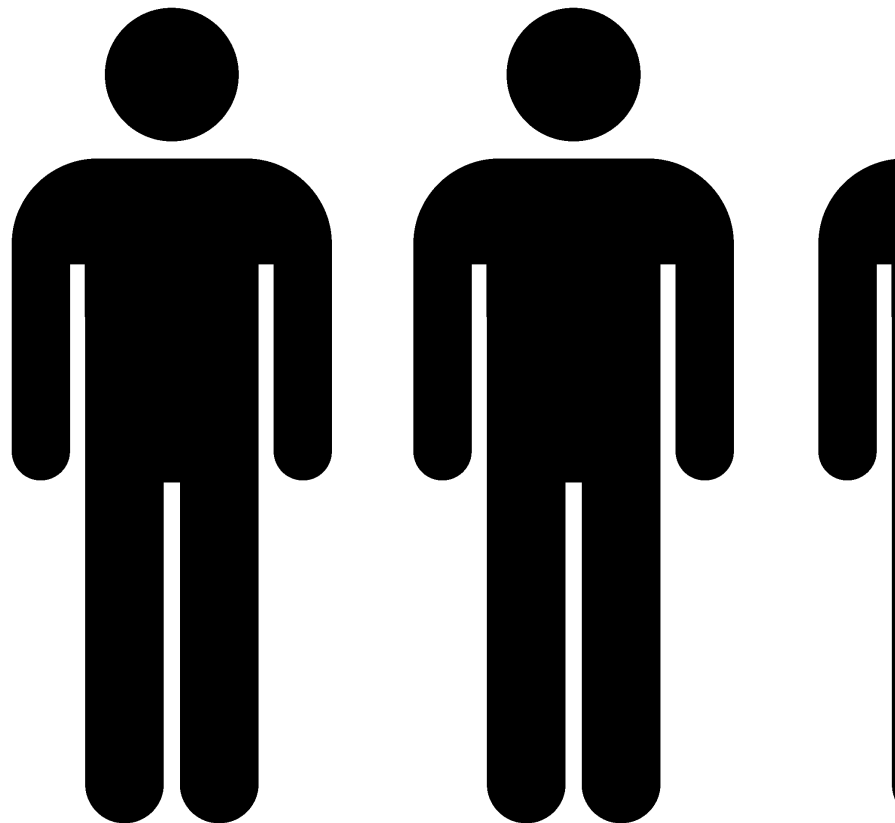


Image designed by Stanford computer science students Chris Chan and Greg Cuellar.

But what
about
biomedical
computation?

So what happens when biomedicine intersects with computer science? Many believe that biomedical computing has the potential to draw women into computing. It offers an antidote to computer science's image problem: whereas pure computing is stereotyped as machine-oriented, solitary, and "geeky," biomedical computing is seen as human-centered, team-oriented and socially relevant.

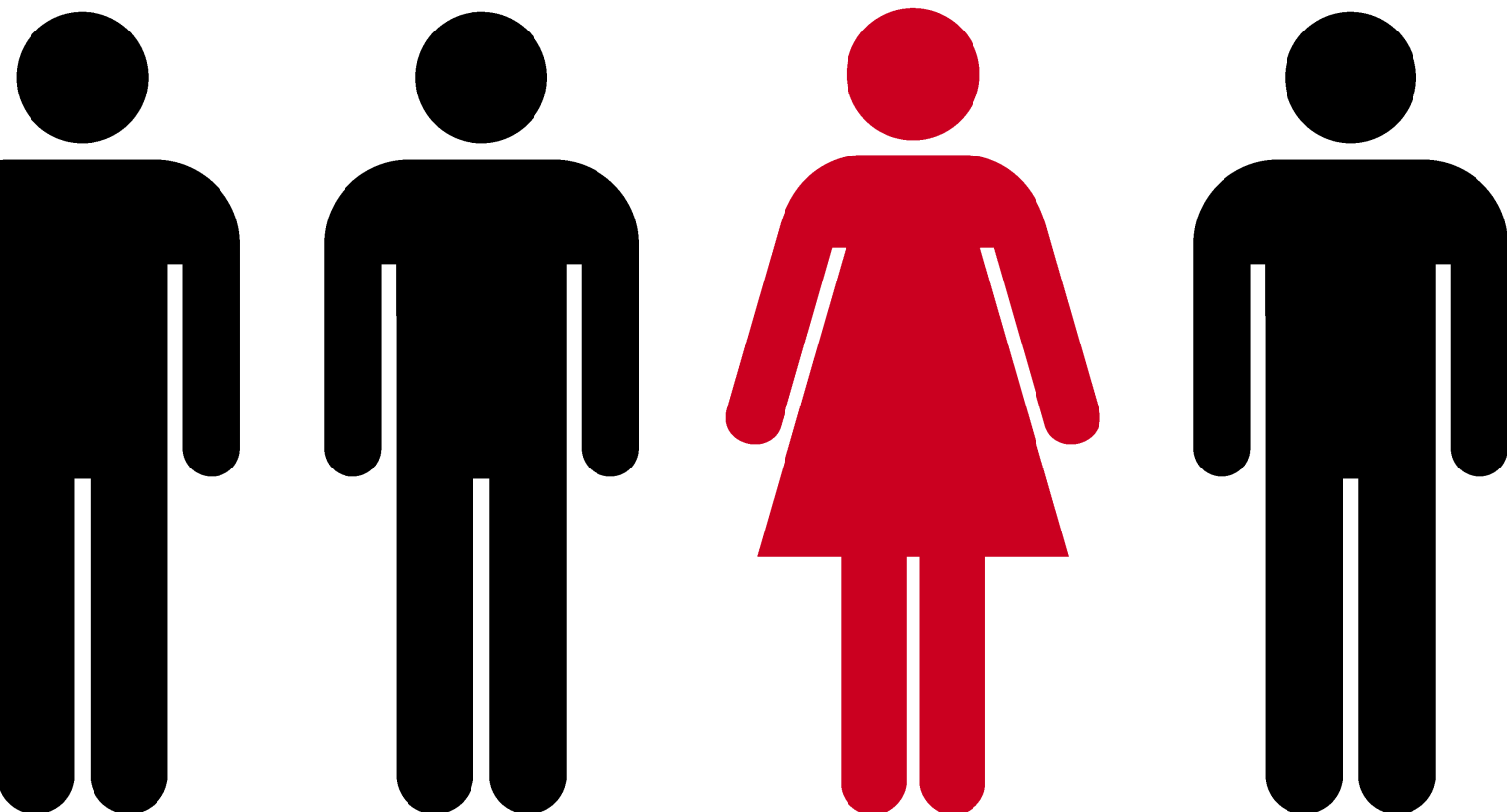
But the early evidence shows a mixed picture. *Biomedical Computation Review* surveyed several universities with programs in biomedical informatics—Columbia, Harvard, the Massachusetts Institute of Technology, Stanford, the University of California, Los Angeles, and the University of Michigan—and found that, so far, just 24 percent of graduate students in these programs have been women. That's closer to women's representation in computer science than in biology.

Because biomedical computing has thus far drawn from students with computational backgrounds, the deficit in females pursuing computer science as undergraduates may create a bottleneck. But a lack of women entering the

field is only half the problem. Across all scientific disciplines, including biology and medicine, the numbers of women shrink as you go up the academic ladder. This "leaky pipeline" phenomenon is worse in the more technical fields. Biomedical computing is no exception: Women make up just 13.8 percent of tenure-track faculty in biomedical informatics and computational biology at the surveyed institutions.

Biomedical Computation Review spoke to women in computing, biomedical computing, sociology, and education about why so few women go into computing; why they leave; and what biomedical computing can do to avoid the gender gap of computer science.

Their message: Biomedical computing holds the promise of drawing more women into computing, but this influx may not happen spontaneously. Moreover, biology and medicine are becoming ever more computational. If women don't have the computer savvy or the influence, they may be left out. More than being an injustice, they say, this would be a huge loss of talent and perspective for both biomedical computing and biomedicine.



PRECOLLEGE: NATURE OR NURTURE?

The gender gap in computer science starts early. Though equal numbers of men and women enter college with prior computer experience (UCLA Higher Education Research Institute), and young women are using computers at an incredible pace, only 14.8 percent of those taking the computer science AP exam in 2005 were women. Compare that to 46.3 percent for cal-

in computer science (in college) varies widely: from 1.79 (in Turkey) to 6.42 (in the Czech Republic). This variation is better explained by societal and cultural norms than genetics. Similarly, the under-representation of minority men in computer science cannot be explained by innate gender differences.

Margolis also found that socialization plays a major role. In her interviews, 40 percent of the male students

children's interests playing out as self-fulfilling prophecies?" Margolis asks in *Unlocking the Clubhouse*.

Suzanna Lewis, MS, of the University of California, Berkeley, recalls the early influence of her four older brothers: "I think by the time I came along, I didn't get treated any differently. Boys were who I played with. So it never occurred to me that being female meant that much." She later designed control systems for steel mills

"We found that while girls were not frightened of the computer or disinterested, there was a difference in how involved they got with computing," says Margolis.

culus (College Board AP Program).

Boys develop a "magnetic attraction" to computers at an early age, observe **Jane Margolis, EdD**, and **Allan Fisher, PhD**, in *Unlocking the Clubhouse: Women in Computer Science* (MIT Press, 2002). In their 1995 to 2000 interviews of 46 male and 51 female computer science majors at Carnegie Mellon University, three-quarters of the men—but only one-quarter of the women—fit the profile of someone who spent much of their youth consumed with computers.

"We found that while girls were not frightened of the computer or disinterested, there was a difference in how involved they got with computing," says Margolis, a researcher at the UCLA Graduate School of Education.

Some people attribute this disparity to innate gender differences: Women are nurturing and human-focused; men are analytical and object-focused; so, computers appeal more to men.

But socialization is working to a greater degree than we realize, argues **Maria Charles, PhD**, professor of sociology at the University of California, San Diego. "Both men and women believe in these fundamental gender differences. Independent of their truth, these beliefs can be very powerful in affecting people's choices."

In her study of 21 countries, she found that the male-to-female ratio

reported being given a computer early in life, compared with only 17 percent of the female students. "When parents place computers in boys' bedrooms and spend more time nurturing their sons' computing interests than their daughters', are they responding to innate difference in the children's level of interest? Or are their assumptions about the

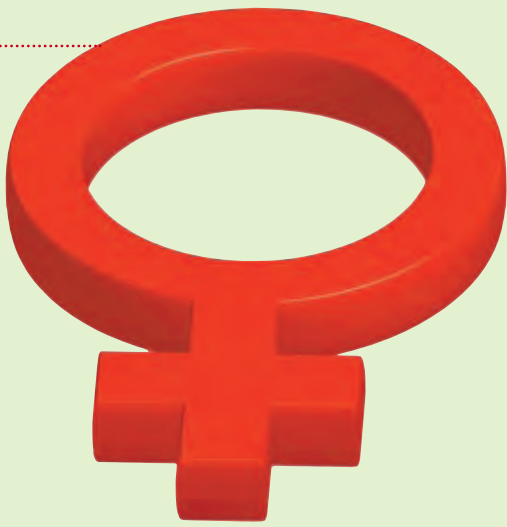
and shipyards for a decade (which she likens to working on "a big tinker toy set"), before moving to bioinformatics.

Grace Peng, PhD, grew up with math exercises over dinner. "I have a Chinese background. And so engineering was not thought of as something very strange for a woman," she says. When she was 10 years old, she attend-

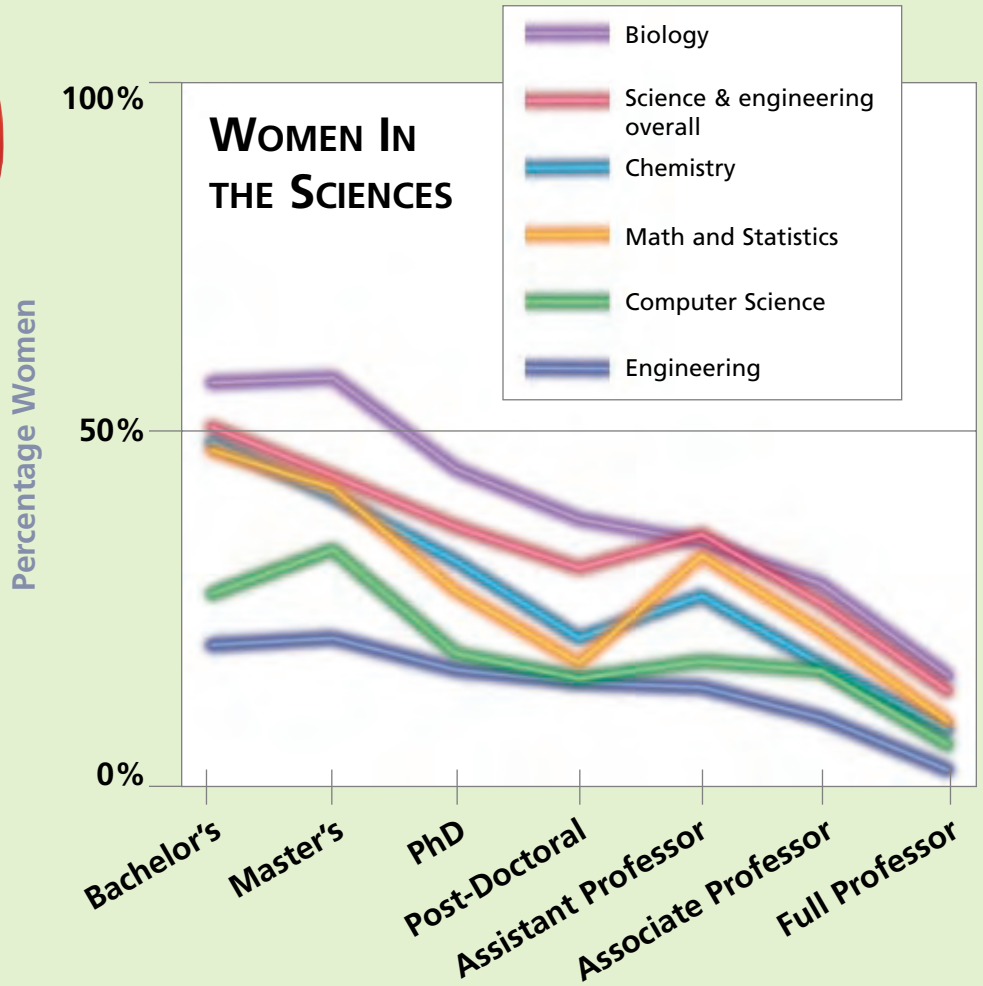
JANE MARGOLIS, EdD:
researcher at the
University of California,
Los Angeles, Graduate
School of Education

Only seven percent of
Carnegie Mellon
University's computer
science students were
women in 1995, when
Margolis and Fisher
began extensive
interviews to find out
why. The research and
ensuing action culminated
in a book (*Unlocking the
Clubhouse: Women in
Computing*. MIT Press,
2002) and left CMU's
entering class in
computer science at
42 percent women
in the fall of 2000.

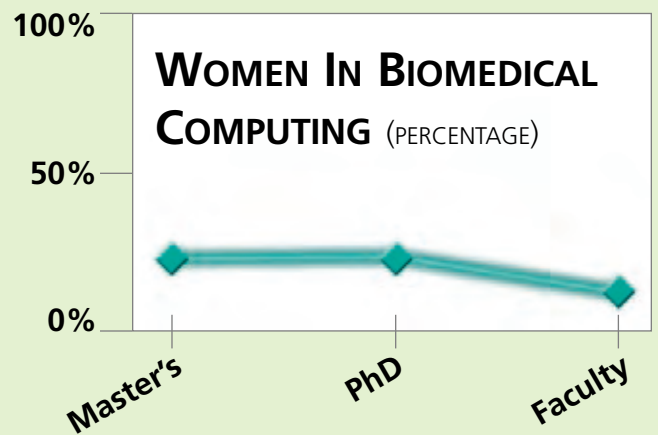
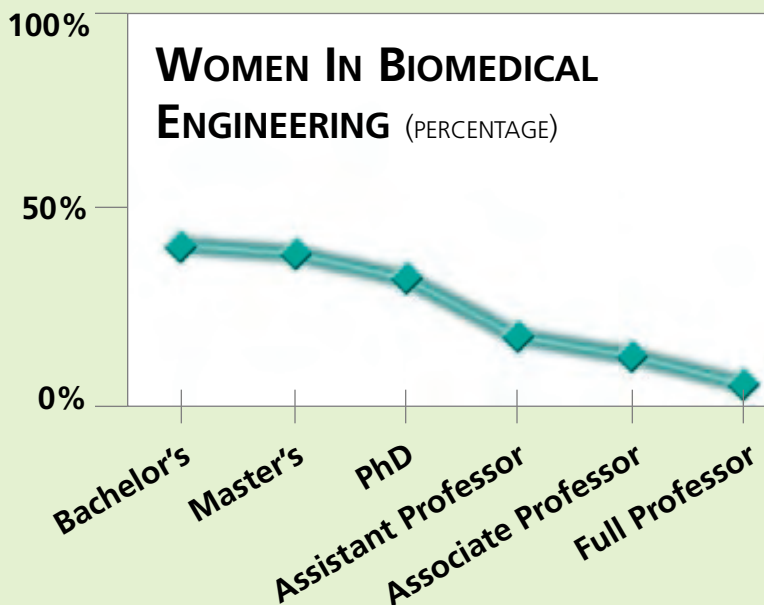




A LOOK AT WOMEN IN SCIENCE, BIOMEDICAL ENGINEERING AND BIOMEDICAL COMPUTING



Statistics from: Women, Minorities, and Persons with Disabilities in Science and Engineering: 2004, National Science Foundation (NSF). Data are from 2001. Notes: the chart includes only professors at 4-year colleges and universities; the category "science and engineering overall" also includes psychology and social science; for professor levels in chemistry, all physical sciences are included.



Left: Statistics From: American Society for Engineering Education (ASSE). Above: Data are from 2003. Data compiled in 2005 from: Stanford, Harvard, MIT, UCLA, Columbia, and the University of Michigan. Student data are limited to graduates and current students of degree programs in biomedical informatics (n=429). Faculty data are limited to tenure-track faculty in programs in biomedical informatics or computational biology (n=239).

ed a program at the University of Illinois, Urbana-Champaign aimed at encouraging girls in computing—which she says made her comfortable with computers from an early age. She is now program director in the Division of Discovery Science and Technology at the National Institute of Biomedical

“I think that it is an inaccurate and a repulsive kind of an image that keeps terrific men and women away. Not just women. Both,” she says.

“I don’t find the computer science world to be very socializing,” agrees **Mia Levy, MD**, a first-year student in the bioinformatics PhD program at

technology world changes so quickly. In response to Margolis’ study, the computer science department at Carnegie Mellon University changed their admissions imperative to finding thinkers and leaders—“and they got fabulous thinkers and leaders,” Sanders says. This admissions change, coupled with curriculum change, increased female representation from 7 to 42 percent in just five years. It also changed the computer science culture at Carnegie

Mellon for the better, Margolis says.

Biomedical computing could draw from a large pool of thinkers and leaders in biology and medicine if they keep the technical barriers to entry low, Sanders says.

COLLEGE: THE CONFIDENCE GAP

Women often begin college with less computing experience than their male counterparts. “Boys have been tinkering and experimenting and working at the computer since they were very, very young; it becomes almost a physical intelligence,” Margolis says. This may intimidate women, and erode

Are parents’ assumptions about children’s interests playing out as self-fulfilling prophecies?

Imaging and Bioengineering managing programs related to computation and engineering systems.

If socialization is playing a major role, then early access and encouragement may be keys to effecting change. For example, foundations that donate computers to schools should require schools to demonstrate that girls are using the computers as much as boys—who tend to be more aggressive about grabbing them, says **Cherri Pancake, PhD**, professor of electrical engineering and computer science at Oregon State University. Charles also urges parents and educators to downplay gender stereotypes and to require more math and computer science classes for everyone.

Finally, whether nurture- or nature-driven, girls tend to be less interested in the computer games that appeal to boys and more interested in how computing can solve real problems. Thus, integrating computing into subjects outside of the computer lab, such as biology, may help draw more women into technology before college.

COLLEGE: “GEEK CULTURE”

Tying computing into compelling problems—such as those in biology and health—may also help discredit the computer science stereotype of “this geeky white guy sitting behind a terminal, getting his suntan from the terminal rays,” says **Lucy Sanders, MS**, CEO of the National Center for Women and Information Technology at the University of Colorado, Boulder. She formerly worked at Bell Labs for 24 years, including as a chief technology officer.

Stanford. She says this stereotype was reinforced by her computer science classes, where the assignments were always to “create a computer game.”

“Having the month before gone from the critical care unit to then suddenly programming Boggle, I was like, ‘My life seems very insignificant right now in comparison to how important it all seemed before,’” she says.

Besides curriculum reform, broadening admissions criteria for computer science can also attract more diverse candidates, Sanders says. Requiring extensive computing know-how for freshmen is crazy, she says, and also irrelevant—since the

MARIA CHARLES, PhD
(sociology); professor of sociology, University of California, San Diego; co-author of *Occupational Ghettos: The Worldwide Segregation of Women and Men* (Stanford Press, 2004)

In an upcoming paper, Maria Charles argues that abolishing gender discrimination does not increase women’s representation in technology because widely held beliefs about gender differences (“women are more nurturing; men are more analytical”) reinforce the technology gender gap.



Integrating computing into subjects outside of the computer lab, such as biology, may help draw more women into technology before college.

SUZANNA LEWIS, MS (engineering and biology); director of bioinformatics, Berkeley Drosophila Genome Project; co-principal investigator of the National Center for Biomedical Ontology, University of California, Berkeley

As a student at the University of Michigan, Suzanna Lewis foresaw the overlap of biology and computer science long before bioinformatics was even a term—so she pursued a dual degree in engineering and biology. For a decade afterwards, she focused only on the engineering side (designing control systems for heavy industry), until the Human Genome Project lured her back into biology.



GRACE PENG, PhD (biomedical engineering); program director at NIH/NIBIB
Grace Peng says her decision to move from electrical engineering (which she studied as an undergraduate at the University of Illinois, Urbana-Champaign) to biomedical engineering (which she studied in graduate school at Northwestern University) may unconsciously have been driven by her desire for more human interaction and impact. After several years on the faculty at Catholic University of America and John Hopkins University, she took a position at the NIH, where she can “help people on a daily basis, in a different way than being in academia.”

CHERRI PANCAKE, PhD (computer engineering); professor of electrical engineering and computer science at Oregon State University
Cherri Pancake worked as an ethnographer studying the Mayan Indians in Latin America for more than a decade. She later went back for her PhD in computer engineering (because “museum work required a second source of income”), and was, in 1982, the first woman admitted to any graduate engineering program at Auburn University. She now applies her anthropology expertise to improve computer usability for practicing clinicians, scientists, and engineers.



their confidence at a time when—studies show—women generally experience a dip in confidence.

Indeed, Pancake found that women were leaving the computer science major at Oregon State University because they believed they couldn't keep up—but, in fact, they were doing just as well as the men.

“What we believe from observing in

the classroom is just that men tend to act more self-assured, not that they're any better prepared,” Pancake says.

“It's the imposter syndrome,” echoes **Francine Berman, PhD**, director of the San Diego Supercomputer Center. “People of both genders feel like secretly they don't have what it takes, and they're just trying to make sure that no one knows that. Women tend to take it



LUCY SANDERS, MS (computer science); CEO and co-founder of National Center for Women and Information Technology, University of Colorado, Boulder
After earning her MS in Computer Science at the University of Colorado, Boulder, Lucy Sanders worked at Bell Labs for 24 years, where she was a Bell Labs Fellow and a Chief Technology Officer. She says she never noticed the gender ratio until she climbed the technical ladder. “Up in the leadership ranks on the technical side, there are very few women that are really making the choices about invention.”

MIA LEVY, MD, PhD
candidate in biomedical informatics, Stanford University

Mia Levy studied bioengineering as an undergraduate at the University of Pennsylvania, where, she notes, “Fifty percent of my class was women, which was certainly not the case for the engineering school at large.” She programmed for a start-up during the dot-com boom (“when they were hiring anybody who had ever taken a computer science course”), before entering medical school at Rush University, in time to beat the bust. She is now a fellow in medical oncology at Stanford University and also a first-year PhD student in biomedical informatics.



more seriously and personally. They are encouraged not to stick with it, because they're afraid it's true. But it's not true."

Interestingly, Margolis found that international women seem to better resist this confidence trap; though they often had no computing experience to start, they stuck out the rigorous program at Carnegie Mellon University. International students expressed the belief that if you work hard, you will make it, but American women too often succumbed to the belief that intelligence is fixed—not malleable—and either you have that

certain intelligence or you don't.

"There have been studies showing that if you believe that intelligence is fixed and innate, your confidence is much lower than if you believe that, with work, it's malleable," she says.

Confidence and competence are not always related, Sanders advises. Rather than accepting confidence at face value, she developed a trick: "Somebody would say something like they just knew it was true—which is common in my discipline—and I would say: 'oh, why?' I'd go through two or three rounds of questions before I would decide whether they knew what they were talking about or they didn't."

Tinkering with computers doesn't make you a good computer scientist, adds Samantha Chui, a Stanford University computer science graduate who is now a first year master's student in bio-medical informatics: "I feel like there are two kinds of people: the people who know everything about the hardware and can't program, and the people who know how to program but don't know how to fix the computer. So I just categorize myself as the latter, and move on."

GRADUATE SCHOOL: LEAKY PIPELINE

Women who enter graduate programs in science are less likely than their male counterparts to complete a PhD and to move on to tenure-track faculty positions, particularly at top research institutions. A year ago, Harvard University's outgoing president, **Larry Summer, PhD**, ignited a firestorm with his provocative suggestion that innate differences in ability play a role. It's an idea uniformly rejected by our panelists. They say women are opting out because they perceive the academic environment to be unappealing (more specifically: unwelcom-



FRANCINE BERMAN, PhD (mathematics); director, San Diego Supercomputer Center
Francine Berman's pioneering research in grid computing led her to become professor of computer science at the University of California, San Diego, and then Director of the San Diego Supercomputer Center. She was named one of the top women in technology by Business Week magazine in 2004. She says: "It's been really interesting because there are not a lot of women at my level in the community. There are some differences in the way that you're perceived, but like any job, people get used to you and what you have to offer, and then they go beyond gender."

Carnegie Mellon University "got fabulous thinkers and leaders" when they changed their admissions criteria, Sanders says.



SAMANTHA CHUI, BS
(computer science); MS
candidate in biomedical
informatics at Stanford
University

Samantha Chui says the stereotype of computer science majors “spending every night alone hacking and talking in chat rooms and gaming” applied to only a minority of men at Stanford, and that, in fact, there was plenty of camaraderie and teamwork: “It could be social if you let it be.” As a first year master’s student in biomedical informatics, she’s noticed a greater diversity of students in her classes, including more women.

ing, harsh, individualistic, combative, competitive, political, secretive, and focused on empire building) and incompatible with having a life or raising a family.

These impressions are formed early in graduate school and they’re largely untrue, says **Semahat Demir, PhD**, associate professor of biomedical engineering at the University of Memphis and the University of Tennessee. “When graduate students see their faculty as role models, they don’t get a good feeling for how much the faculty enjoy their jobs,” she says. “They see the faculty working so many plus hours. They see deadlines and frustrations.” But, in fact, she says, academia is very fulfilling and offers a lot of flexibility for balancing life and work.

“Women faculty have to talk to these women early and show that you can get a balance,” she says. She advocates mentoring to facilitate communication between faculty and students, and to help women students develop soft skills that are otherwise inadequately addressed in graduate school, such as negotiation.

“I know for myself, if I’m in a kill ‘em and eat ‘em environment, that doesn’t encourage me to do my best work,” says Berman, who has promoted a more collaborative environment at the San Diego Supercomputer Center. Though she doesn’t make a conscious effort to recruit women, they make up half of her senior management team.

Biomedical computing is team-based and collaborative by nature, which may help to attract and retain women, our panelists say.

FACULTY: SECRET LIVES

After earning a science doctorate, women are more likely to leave academia or to take a

SEMAHAT DEMIR, PhD
(computer engineering);
associate professor of
biomedical engineering at
University of Memphis and
University of Tennessee;
program director,
National Science Foundation,
Biomedical Engineering and
Research to Aid Persons
with Disabilities

Semahat Demir studied electrical engineering as an undergraduate at Istanbul Technical University, and worked as a clinical engineer in Turkey and Germany—where, at the time, she was often “the only female engineer in the workforce.” She now advocates peer mentoring for women in technical fields to give them a forum to discuss their experiences, build their confidence, and develop soft skills.



non-tenure track position—choices often attributed to a conflict between the biological clock and the tenure clock. Among science and engineering doctorates who are not working, 35 percent of women cite family responsibilities as the reason, compared with only 2 percent of men. Among employed doctorate holders, 33 percent of women versus just 17 percent of men have never been married (NSF).

“Professional life in general—and high stakes professional life in particular—is not family friendly, for either women or men. You’re incentivized for working all the time and for choosing your career over your family,” Berman says. “Professional success encourages you not to have a life. But everyone does have a life, so they have secret lives,” she says.

Berman now makes a point of talking about her family, to show young women that it is possible to have both a family and a successful career in computer science. Academia actually has a lot of flexibility for parents, she says. “Earlier in my career,” she says, “if I had to go home to watch my daughter play a fork in the school play, I’d just say I had an off campus commitment. My husband did the same.”

FACULTY: SUBTLE EXCLUSION

Even when women decide to pursue a tenure-track academic position, they may face unconscious biases in hiring, compensation, and evaluation. For example, studies show that both men and women will rate an identical resume or journal article lower (on average) if they are told that the subject or author is a woman. A 1999 report on the status of women in science at MIT revealed that,

even when controlling for other factors, women faculty members received less pay and less lab space than men—findings that were replicated at other elite institutions.

Jeanette Schmidt, PhD, executive director of Simbios at Stanford and former vice president of research and development at InCyte Corporation, says she never felt her capabilities doubted as a graduate student in Israel. But as an assistant professor at Polytechnic University in New York, she initially felt a subtle exclusion that she attributed to being a woman.

The manifestations were often well-meaning. For example, she was assigned to teach a morning undergraduate course rather than a more desirable evening graduate course, because—she was told—that way she could “be home with her kids at night.”

Biases creep into academia more than industry, Lewis speculates. In industry, you are judged by how you affect the bottom line, she says. In academia, you are judged on something less tangible—your ability to debate, persuade, and make yourself known. These

Biomedical computing is team-based and collaborative by nature, which may help to attract and retain women, our panelists say.



TAMAR SCHLICK, PhD (mathematics); professor of chemistry, mathematics, and computer science at New York University; director of the new PhD program in computational biology at NYU

Tamar Schlick says she didn't dwell on the fact that she was one of the few women in mathematics as an undergraduate (at Wayne State University in Michigan) or as a graduate student (at NYU), but she did feel that mathematics was a lonely discipline. She says computational biology has a more team-oriented focus that appeals to women.

“...it’s not a women’s problem, it’s everybody’s problem,” Berman says.

kind of subjective criteria introduce a soft bias that favors the dominant group, a “deafness” as she describes it: “How many times have you [as a woman] been standing in a group and you say something and nobody recognizes it? And then someone else who happens to be male says the same thing, and everybody says ‘Ah, what a great idea!’”

Counteracting unconscious biases requires a pro-active approach. By actively recruiting women, MIT increased from 22 to 34 women faculty in the School of Science from 1994 to 1999 (previously, the faculty had stuck fast at 8 percent women for more than a decade). At Oregon State University, Pancake says that sending targeted recruitment letters to eligible women and minorities has increased their proportion of women engineering faculty to the fourth highest in the country.

Biomedical computing needs to be similarly pro-active in recruiting and promoting women, our panelists stress. And this impetus can’t just come from women in the field.

“Typically you see a lot of the programs developed and executed by women for women. But it’s not a women’s problem, it’s everybody’s prob-

lem,” Berman says. “If it was important enough for everyone to address this problem, it would cease to be a problem.”

TWO RIVERS MERGING

Molecular biology and computing were born in the same era and have evolved side by side, like two rivers winding down a mountain in parallel, Lewis reflects. The two rivers are now merging, as computing becomes more and more integral to biology, she says. And before too long, it’s not going to be easy to distinguish between the two fields.

Computing is becoming integral to many other disciplines as well, which is precisely what makes the gender gap in computer science so worrisome. At a time when too few American students are going into technology to keep up with the demand, we can’t afford to lose half our talent pool.

We also can’t afford to lose women’s perspectives, Sanders says. “Technology can’t be designed and invented by a homogenous group of people because it will be less than it should be.”

Biomedical computing is poised to open up computing and technology to a larger audience, in particular women. The field’s collaborative nature and its

connection to medicine (which has been so successful in attracting women) both portend a bright future for women drawn to this kind of work. The women highlighted here are just a few of the many who are impacting the field by acting as role models and mentors as well as by influencing university admissions and educational policies.

In 2003, 24 percent of the biomedical computing PhD students at the universities we surveyed (see chart) were women, compared with about 18 percent in computer science programs nationwide that year. These numbers suggest a positive trend that, if it continues, will impact the future of several disciplines: not only biomedical computing, but also biology, medicine, and computer science. □

JEANETTE SCHMIDT, PhD (computer science); executive director, Simbios, Stanford University
Jeanette Schmidt had an early interest in combining math and biology, but “It was impossible because all the courses conflicted. It was like nobody had ever thought of it.” So, she majored in math, followed by a PhD in computer science (Weizmann Institute of Science in Israel). Later, as an associate professor of computer science at Polytechnic University in New York, she rekindled her interest in biology, designing algorithms for biologists. She eventually joined InCyte Corporation, where she worked for seven years, including as vice president of bioinformatics research and development.



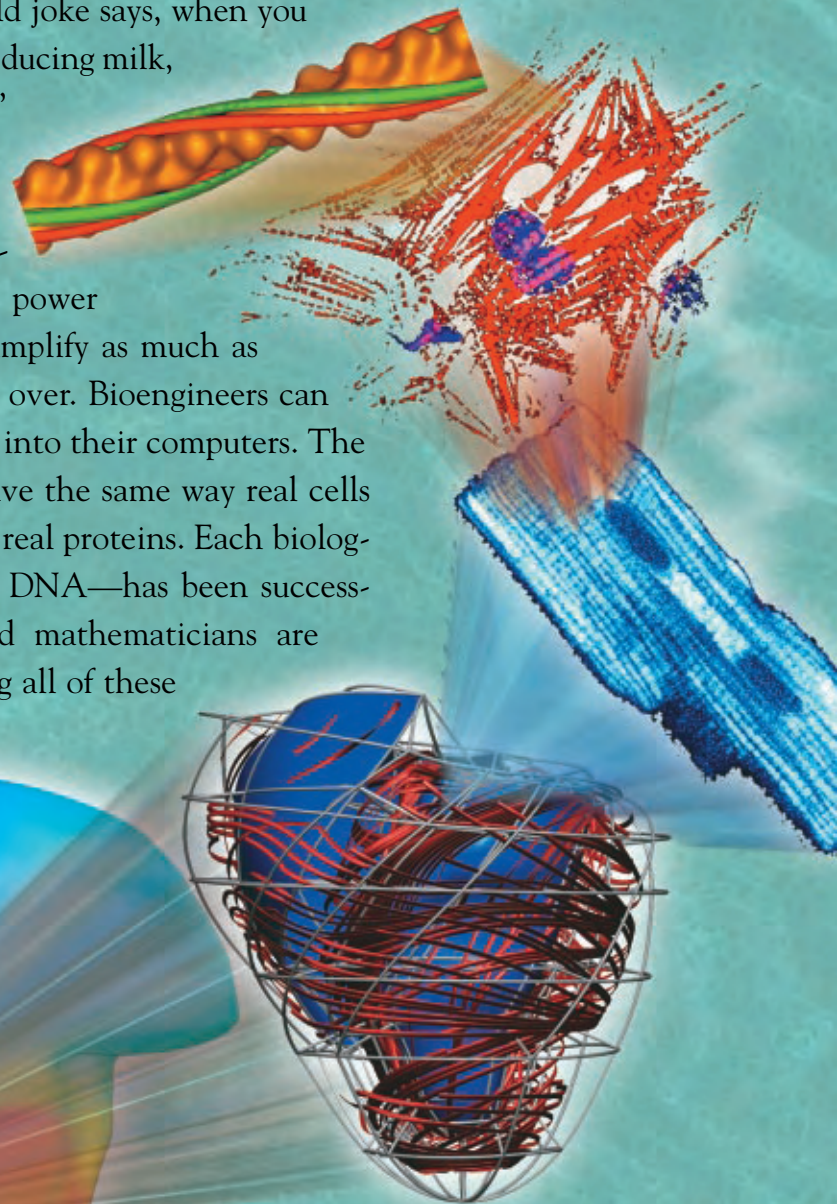
RAMPING UP TO Multiscale

TAKING BIOMEDICAL MODELING TO A NEW LEVEL

BY DANA MACKENZIE

For centuries, mathematics has been an indispensable ally of the physical sciences and engineering. Planes fly and telephones work because engineers know how to simplify physical systems into convenient mathematical models. But biologists and mathematicians have had a harder time communicating. As the old joke says, when you ask a mathematician to explain why a cow isn't producing milk, he'll probably begin, "Consider a spherical cow..."

However, attitudes are changing in both disciplines. With the advent of computational biology, some biologists are shifting toward more quantitative models. And today's vast computing power means that mathematicians no longer have to simplify as much as they used to. The days of the "spherical cow" are over. Bioengineers can program an anatomically correct cow (or human) into their computers. The organs can be made out of virtual cells that behave the same way real cells do, and contain virtual proteins that interact like real proteins. Each biological scale—organism, organ, tissue, cell, protein, DNA—has been successfully modeled in isolation. Now, biologists and mathematicians are beginning to grapple with the problem of unifying all of these layers into a single multi-scale model.



Among the most mature types of multi-scale models are simulations of the human heart. Accurate equations that describe individual heart cells have existed since the early 1960s. They have greatly clarified how the flow of ions through channels in cell membranes causes heart cells to transmit electric signals at precisely timed intervals. Now the models are reaching down to the molecular level, to explain how gene expression or drugs cause changes in the ion channels. At the same time, they are reaching up to the organ level, placing the cell models in the context of macroscopic physiology.

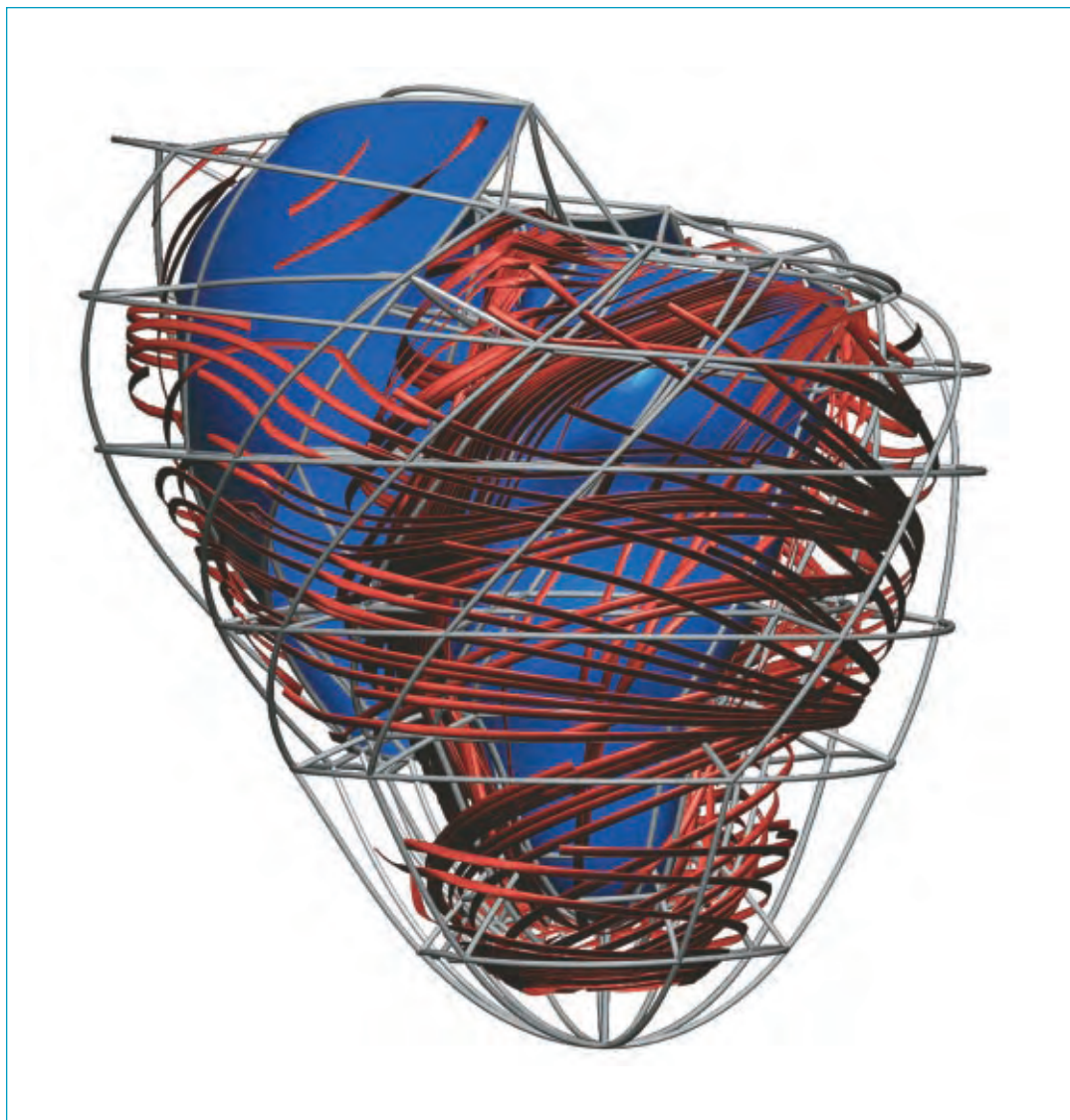
In ischemia, for example, a local event—the blocking of blood flow in a coronary artery—creates organ-wide consequences, as a whole region of heart muscle is deprived of oxygen. This in turn affects the heart tissue at a cellular level, by altering the chemistry inside the cells. The intracellular changes create an arrhythmia, which propagates back up to the whole-organ level. This interplay between the different physical laws at different levels is what multi-scale modeling is all about.

Even so, cardiac models are not necessarily a blueprint for other parts of biology. “We’re a long way from generating the principles by which multilevel work should be done,” says **Denis Noble, PhD**, professor of cardiovascular physiology at Oxford University, one of the pioneers of cellular modeling of the heart. Indeed, multi-scale modeling is now at what might be called its gestational stage. Everybody knows it’s important, but no one quite knows how to do it.

Nevertheless, money is flowing. Last year, an interagency NIH/NSF/ NASA/DOE program funded 24 investigators, to the tune of \$20 million, to work on various projects in multi-scale modeling.

The days
of the
“spherical cow”
are over.

A journal, *Multi-scale Modeling and Simulation*, launched in 2002 and published its first articles in 2003. In almost every part of biology—from bacteria to humans, from the heart to the brain—scientists want to uncover the rules that organize nature’s complexity. “You have to hope there are underlying principles,” says **James Glazier, PhD**, the director of the Biocomplexity Institute at Indiana University and organizer of eight biocomplexity conferences. “If not, you’re out of luck.”



Computer models of the heart incorporate detailed experimental information, both at the level of individual cells and at the level of anatomy. Here, a model developed by Peter Hunter's team at the University of Auckland portrays the changing orientation of the heart's muscle fibers from the outside to the inside of the heart wall. The spiraling of the fibers is believed to affect the flow of electric signals through the heart. Courtesy of Peter Hunter, PhD, Bioengineering Institute, The University of Auckland, New Zealand.

ENGINEERING THE CELL

In 2002, **Yuri Lazebnik, PhD**, of Cold Spring Harbor Laboratory wrote a much-discussed satirical article for *Cancer Cell* called “Can a biologist fix a radio?” Lazebnik’s answer was no. He argued that the usual research method of biologists—knock out one component at a time, and see which ones stop the cell from working—would not enable them to figure out how a transistor radio works. Why, then, should we expect to understand the workings of a cell in this way?

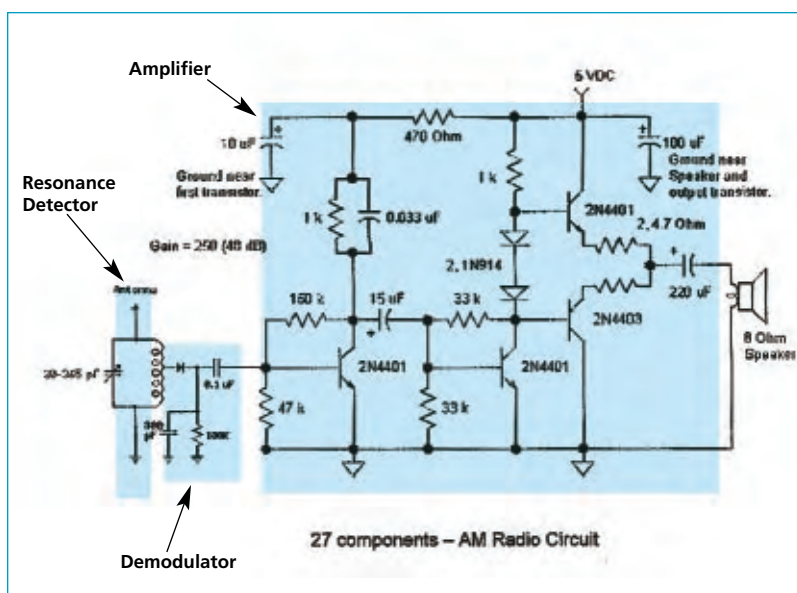
Last year, at the Biocomplexity 7 conference, **Herb Sauro, PhD**, turned the question around. The assistant professor of biochemical control systems at Keck Graduate Institute asked: “Can an engineer fix a cell?” His answer was a qualified yes. “Engineers deal with complex systems day in and day out,” Sauro says. “Today’s computer systems have hundreds of millions of components, a level of complexity that is rapidly approaching that found in biological systems.” But, he says, engineers have a secret that not all biologists have learned: “Engineers modularize.”

It is still far from clear whether nature modularizes. If so, it does so in a very different way from human engineers, because natural systems are not rationally designed; they arise through natural selection. Nevertheless, the final outcome may be the same. A particular network may offer a powerful selective advantage precisely because it performs some function in an optimal manner.

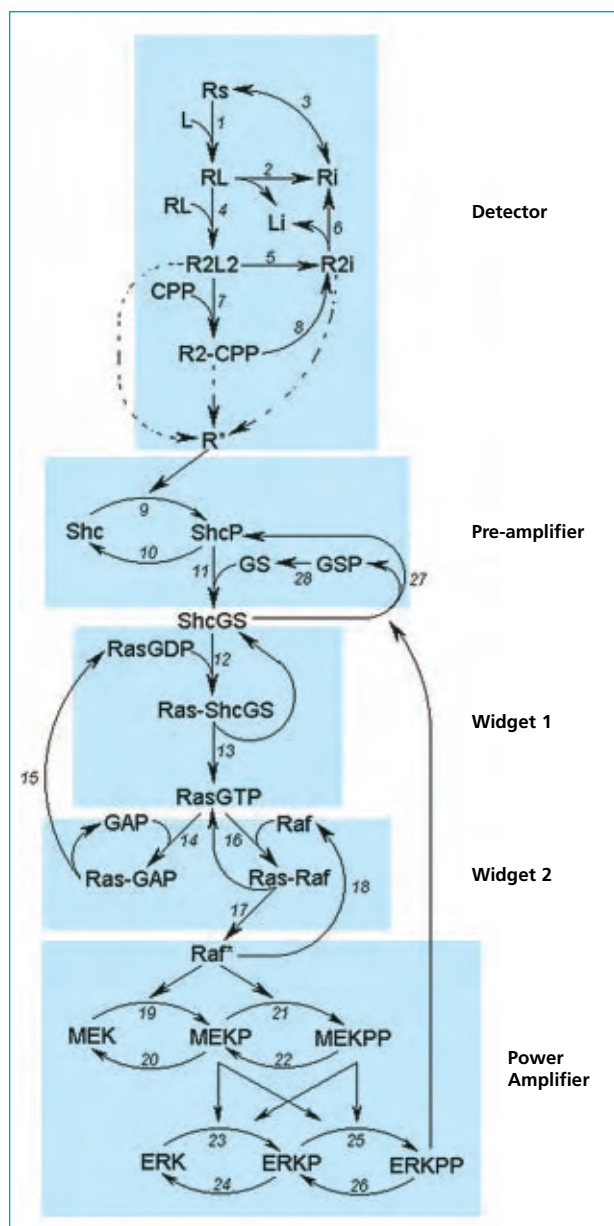
To the layman, the circuit diagram of an AM radio looks incomprehensible. But the system becomes easier to understand once you realize it has three modules: a resonance detector, a demodulator, and an amplifier. From there, an engineer can break the circuit diagram down into smaller modules, each with a specific function. In this way, possibly passing through many layers, the engineer can tell how any electronic device works.

A cell, like a radio or a computer chip, contains many components that interact with each other in a dizzyingly complex network. Most biologists, Sauro contends, are

“We’re a long way from generating the principles by which multilevel work should be done,” says Denis Noble.



The circuit diagram of an AM transistor radio (above) looks forbiddingly complex until it’s overlain with functional modules. According to Herbert Sauro, the same can be true of protein interactions. The MAP (in blue) kinase cascade at the bottom of the protein interaction network (right) looks like a negative-feedback amplifier; however, some of the other “widgets” in the network have functions that are still unknown. Courtesy of Herbert Sauro.



satisfied simply to list the components (the proteome) and identify which ones interact with each other (the “interactome”). He says they should also ask: What are the modules and what do they do?

As an example, Sauro cites the mitogen-activated protein kinase (MAPK) cascade, a complicated series of protein-protein interactions that senses conditions outside the cell and initiates cell division. When the MAPK cascade goes haywire, one possible result is cancer—which explains why many biologists are interested in it. It has a very distinctive and well-

all eukaryotes. Unlike electronic components, Arkin says, biological modules have the ability to evolve and adapt. One particular switch, called the *sin* operon, is ubiquitous in bacteria but plays flexible roles. Arkin has showed that it can function as a graded switch, like a light dimmer; a bistable switch, like a normal wall switch; or a single pulse generator, like the switch of a flashbulb.

If it is true that nature modularizes, it raises the possibility that humans can actually design bacteria to perform certain functions. For larger organisms, such as humans, modularity is impor-

the way a heart muscle cell reacts to ionic currents, with equations from physics that describe how electric currents propagate through conductors. The anatomy of the heart plays an important role, because heart muscle does not conduct electricity equally in every direction: the current flows preferentially along muscle fibers.

The models show that fibrillation starts with tachycardia. This may feel like a “rapid heartbeat,” but it is not really a heartbeat at all. A normal heartbeat is a wave of electrical excitation that progresses from the heart’s pacemaker (the sinoatrial node) and

“Engineers have a secret that not all biologists have learned,” says Herb Sauro, “Engineers modularize.”

understood structure: three staircase steps with a feedback from the third back to the first.

When Sauro showed the “circuit diagram” of the MAPK cascade to engineers, they immediately told him what the circuit does. It’s a negative-feedback amplifier, a type of circuit invented in the 1920s to transmit transcontinental telephone calls. The purpose of the feedback is to cancel out distortion, amplifying only the true signal. Sauro admits that it is “still just a hypothesis” that it performs the same function in a cell. However, if this is the optimal way to amplify a signal without distortion, it’s possible that, during the course of evolution, nature may have stumbled onto the same solution that human engineers did.

Adam Arkin, PhD, an assistant professor of bioengineering and chemistry at the University of California, Berkeley, is one researcher who is taking an engineering approach to the study of cells. He has already compiled a library of protein interaction pathways, organized by their possible functions: switches, oscillators, amplifiers, noise filters (such as the MAPK cascade), and so on. Some of these are very widespread. As far as biologists know, the MAPK cascade is found in

tant because it simplifies multi-scale modeling. “If you’ve identified a module with a crisp function, then you can substitute that whole network with a single equation,” Sauro says. This kind of substitution is what will make multi-scale modeling possible. And such models will generate hypotheses that can be tested experimentally—one of the most important ways that computational biology can contribute to biological discovery.

THE HEART OF THE MATTER

Can an engineer repair a heart? The answer, again, is a qualified yes. Every day, defibrillation—a massive external shock applied to the heart—saves the lives of many people who would otherwise die within minutes. When done correctly and promptly, defibrillation has a success rate well over 90 percent. Ironically, though, scientists are not quite certain why it works. It is certainly a more violent and painful treatment than necessary—although, as Noble says, “In a condition where you otherwise die, you put up with that.”

Multi-scale models have enabled heart researchers to “see” much more clearly into the fibrillating heart. The models work on at least two scales. They couple cellular properties, such as

sweeps over the whole heart. Ventricular tachycardia, on the other hand, is a self-organizing spiral of electrical activity that rotates around a center, like a dog chasing its tail. Opinions differ as to whether the center is an anatomical defect, such as a piece of scar tissue, or whether the “rotor” can form anywhere. Either way, the heart muscle cannot sustain it, and the single spiral wave disintegrates into many. That is the onset of fibrillation.

Defibrillation is a mystery. If the heart were a uniform electrical conductor, the shock from the defibrillator would have no way of penetrating the interior of the muscle, and so the gadget would never work. Evidently the heart is not homogenous, but a debate still rages over where to look for the inhomogeneities. Some heart physiologists believe that the relevant features are large-scale (the muscle fibers). Others claim that the shock sets up a voltage gradient across the gaps between layers of cells (or “interlaminal clefts”). Either explanation, if it could be proved by experiment, would be a triumph for computational biology’s ability to turn qualitative hypotheses into quantitative, testable predictions. The second hypothesis, which proceeds from cells up to the organ

level, is perhaps more in the spirit of multi-scale reasoning, but in fact both of them require multi-scale modeling to work in a quantitative fashion.

Meanwhile, heart models are contributing to scientists' understanding of other heart diseases as well. For example, long QT syndrome is an irregular heartbeat that can be caused either by drugs or by genetic mutations that affect the potassium channel. Often its first symptom is sudden death of an apparently healthy young person. Many drugs affect potassium levels, and it makes much more sense to test their side effects first on a computer model than on a live human.

Simulations can also help identify drugs with positive effects. Noble has used them to study an anti-anginal drug called ranolazine, which affects two channels at once, the potassium and sodium channels. So-called "multiple action drugs," like ranolazine, have a poor reputation, says Noble, precisely because "our minds can't wrap themselves around them." Doctors prefer drugs with a single clear effect. But in the case of ranolazine,

either action by itself would cause arrhythmia. The combination avoids arrhythmia as well as the undesirable side effects of other anti-anginal drugs, such as low blood pressure. In January 2006, the FDA approved ranolazine for general use, making it the first new anti-anginal drug in two decades. While it is unclear to what extent the computer models affected the FDA decision, Noble says that such models "can help a new drug application, since understanding what is going on is an important part of the regulatory process. People feel happier with a new compound as a possible drug the more we understand why it acts the way it does."

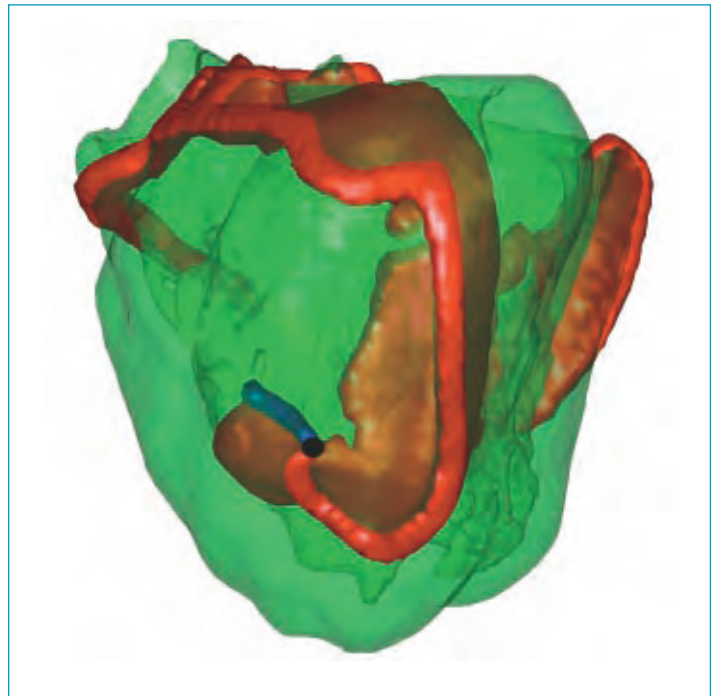
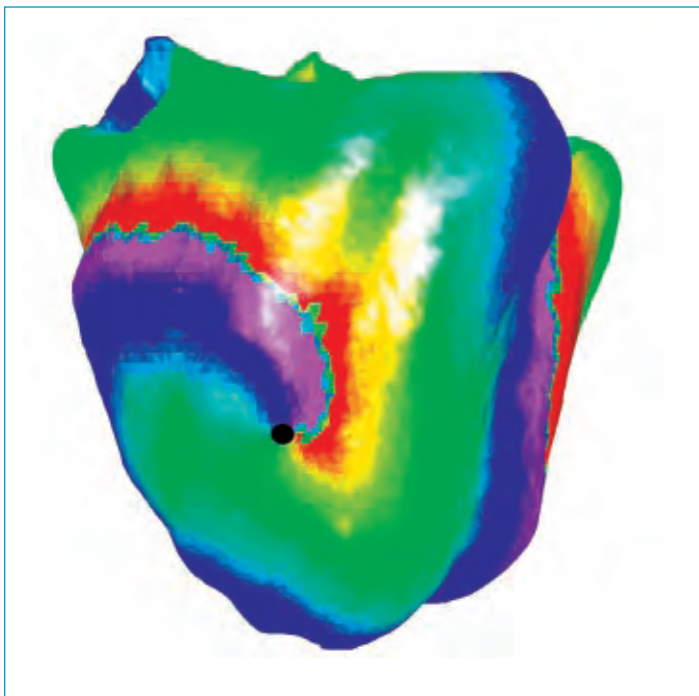
A PANOPLY OF PROJECTS

Last year, the Interagency Modeling and Analysis Group (IMAG), a combined effort of several government agencies coordinated by Grace Peng of the National Institute of Biomedical Imaging and Bioengineering (NIBIB), awarded 24 grants for multi-scale modeling projects in biology. The grants were funded by the individual agencies

(twelve by NIH, ten by NSF, and one each by NASA and the Department of Energy.) Although many of the projects are just beginning, they illustrate the wide diversity of applications envisioned for multi-scale models. Here are a few examples:

■ **James Glazier, PhD**, of Indiana University, will study the processes of limb formation and tissue regeneration. He believes that people in the field count too much on the amazing abilities of stem cells. "The genomic determinants think you'll plunk a stem cell down in the body, and it will spontaneously regrow the tissue that should be there," Glazier says. "Maybe you'll be lucky and it will work that way. But I think that you will have to give complex spatiotemporal signals to those cells." He plans to develop a model of the feedback between the molecular scale—the instructions encoded by DNA—and the large-scale forces that act on cells as a limb grows and takes shape.

■ **George Karniadakis, PhD**, professor of applied mathematics at Brown



Ventricular fibrillation is a complex three-dimensional phenomenon, but experimental methods can probe only the two dimensions of the heart's surface. Using computer simulations researchers can observe (left) a tornado-like "scroll wave" of electric activity spiraling around a filament that passes through the heart muscle (the colors reflect the wave's time of arrival at the heart surface). At right, the wave front inside this semi-transparent rendering of the heart is red and the filament is blue. Lab experiments can only image the places where the filament reaches the heart surface (black dots). Courtesy of KHWJ ten Tusscher and AV Panfilov, Utrecht University, The Netherlands.

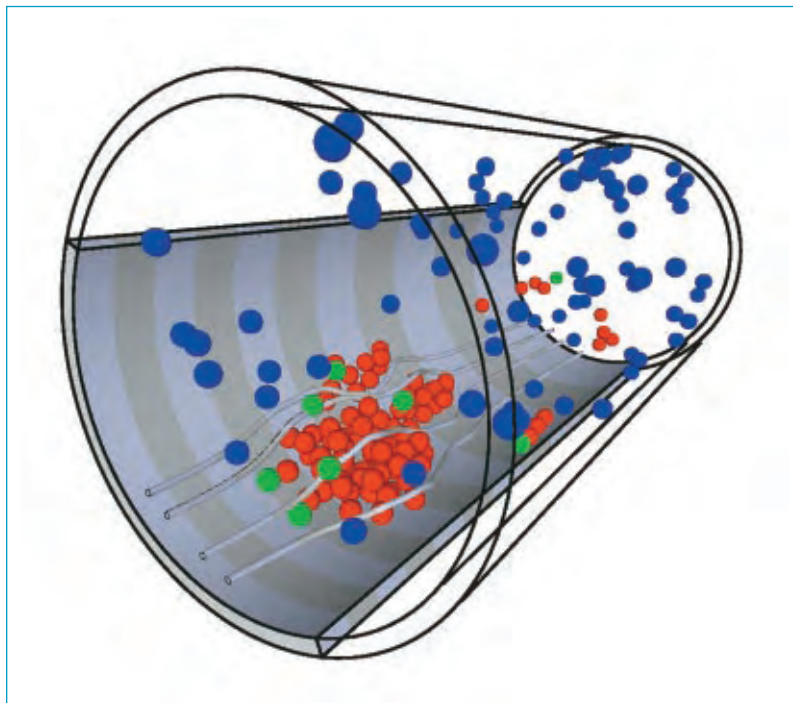
University will model the flow of platelets and the formation of blood clots. Platelets ordinarily look like smooth disks. But when they sense a defect in the arterial wall, they pump

called “dissipative particle dynamics” or DPD, which has been developed by polymer physicists in Europe. DPD is a typical “mesoscale” or inter-

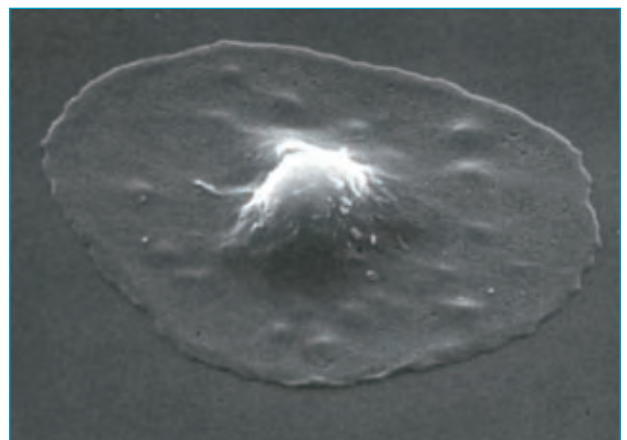
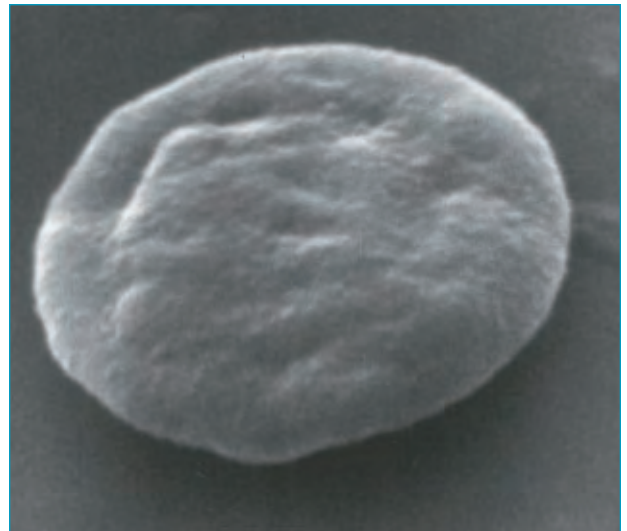
Last year, the Interagency Modeling and Analysis Group awarded 24 grants for multi-scale modeling projects in biology.

themselves up into sticky, spiny spheres. “This kind of phenomenon has never been modeled from first principles, because it’s computationally very complex,” says Karniadakis. Mathematicians and engineers are not used to working with flowing particles that suddenly change their shape and adhesiveness. However, Karniadakis is planning to borrow a new method

mediate-scale mathematical technique, which uses probabilities rather than deterministic equations, as classical physical models do. Ultimately, Karniadakis would like to plug this intermediate-scale model into a large-scale model of the body’s arterial tree. Last year, he and a group of colleagues



A multi-scale model by George Karniadakis and Igor Pivkin aims to be the first to predict clotting time from physical principles. A key ingredient in the model is “dissipative particle dynamics,” a stochastic method designed to model the flow of polymers through a fluid. In the simulation shown here, blue platelets are inactive, green platelets are “triggered” and red platelets are activated. Note that some blood continues to flow through the growing clot. Courtesy of Igor Pivkin and George Karniadakis, Brown University.

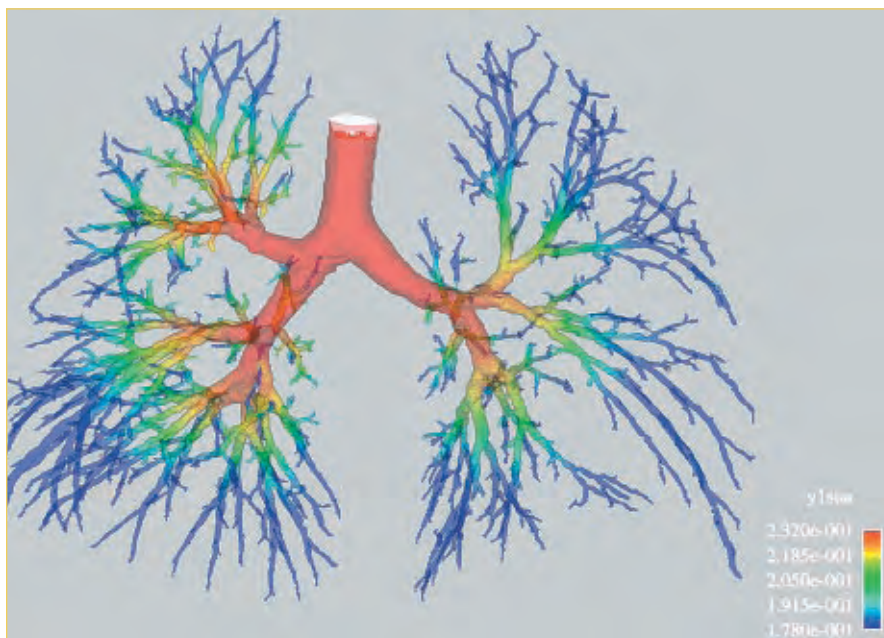
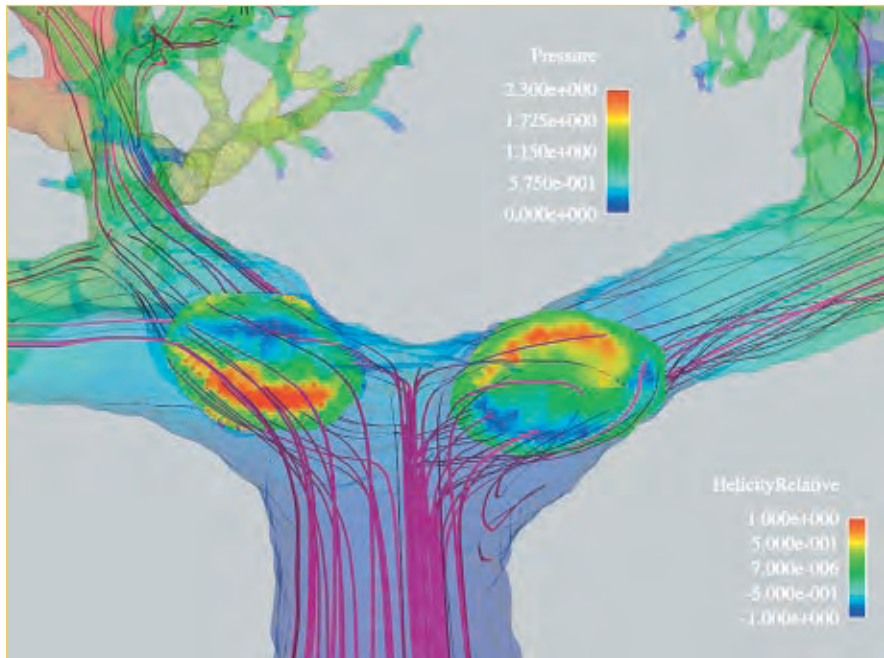


Platelets ordinarily travel through the bloodstream in a disk-shaped, “inactive” form (a). Upon sensing a lesion in the artery wall, they become “activated” and send out sticky pseudopods (b). After adhering to the side of the artery wall, platelets undergo one more change in morphology (c), flattening and spreading out so that the pseudopods no longer protrude as far. Courtesy of James White, University of Minnesota.

used a grid of four supercomputers (based in San Diego, Urbana, Pittsburgh, and Argonne) to prove the basic proposition that you can simulate blood flow in such a complicated set of vessels as the human arterial tree.

■ **Robert Kunz, PhD**, a physicist at Pennsylvania State University, also plans to apply modeling techniques from outside biology. He is developing a simulation of airflow in the human lung inspired, in part, by software used

in the nuclear reactor industry. The flow of coolant in a nuclear reactor is too complicated to model in three dimensions, so computer programs represent the flow with a simplified, one-dimensional model. But if an accident occurs, such as a loss of coolant, the programs immediately switch over to a three-dimensional model of the affected region, and integrate the results seamlessly with the one-dimensional model of the whole reactor. Similarly, Kunz's large-scale lung model will use 3-dimensional



A three-dimensional model of air flow through the lung enables Robert Kunz to predict oxygen concentration (a) and the vorticity of air flow through the bronchi (b). However, only five percent of the lung's volume is contained in its largest bronchi, shown here; 95 percent is contained in its spongelike outer layer. This layer contains billions of bronchi, far too many to model in complete anatomical detail. Kunz is working on a way to integrate the three-dimensional model of the larger bronchi with a smaller-scale, one-dimensional model that describes the terminal bronchi. Courtesy of Robert Kunz, Pennsylvania State University.

Modelers have to learn to walk before they can run.

fluid dynamics to track the flow of air through the wider bronchial passages. However, in the sponge-like outer layer of the lung, where the flow becomes too complicated, his code will switch over to a one-dimensional approximation. In other words, it won't track the twists and turns of every single air molecule, but it will track the progress of an entire breath of air toward its final destination, the alveoli. The model could be used to calculate the uptake of drugs such as inhaled insulin (another drug newly approved by the FDA), or to study how lungs decrease in efficiency with age. One of the other IMAG projects, led by Ching-Hong Lin of the University of Iowa, will also focus on the human lung.

CHALLENGES AND PITFALLS

At present, the number of realistic multi-scale models in biology is very small. "In reality, it has been achieved in only one organ system, the heart," says **Peter Hunter, PhD**, professor of bioengineering at the University of Auckland. "The lungs are getting close. They have all the anatomy of the airways, pulmonary vessels, and gas exchange at the alveolar level, and they are starting to look at the smooth muscle."

Hunter is the co-chair of the Physiome Project of the International Union of Physical Sciences

(www.physiome.org.nz), which runs a website that archives mathematical models of physiology. Currently the site lists around 300 models, essentially all of which work at a single scale. That's no big surprise, because modelers have to learn to walk before they can run. The site is also, at this point, only descriptive: visitors can see computer code for the models but not actually run them. However, **James Bassingthwaighte, PhD**, a bioengineer at the University of Washington,

Glazier feels stymied not only by the lack of data, but the lack of desire to acquire the right kind of data. Every mathematical model incorporates measured parameters. These are like the labels on the radio's circuit diagram that indicate the properties of a resistor or transistor. Glazier itemizes a few that are relevant to biology: "association and disassociation constants, diffusion constants, decay rates, cellular production rates, ..." But biologists aren't convinced that it is worth the effort to measure

them. "Biology is still a 90 percent qualitative discipline," Glazier says. "There's a basic bootstrapping problem. Until experimenters take modeling seriously, you won't have people making measurements to pin down the parameters. And without the right parameters, the record of predictions is not very great."

Another great challenge of multi-scale modeling is that the models at different scales may involve different physical principles and different

assumptions. It will help to put the models on the same computing platform, as Bassingthwaighte is doing, but other fundamental questions need to be addressed. For one thing, "We have no sense of how error propagates from one level to the next," Arkin says. For another, he asks, "Where are the boundaries between fast and slow reactions, or between deterministic and stochastic models?" Physical scientists have developed a very good sense of where the boundaries should be, and which details can be left out when going from one scale to the next. At present, biologists make these decisions in an ad hoc fashion, Arkin says. But perhaps Sauro's modular approach, or switch-on-the-fly software like Kunz's, can make the decisions more rationally based.

In some cases, the mathematical

tools to fit the different scales together may be unfamiliar to biologists. The IMAG program is having a demonstrable effect by attracting researchers like Karniadakis and Kunz, who are bringing in new techniques from physics. Stochastic differential equations, for example, are unlikely to come up in a biologist's mathematical training, but they are a natural fit for biological multi-scale models, because they address the elusive mesoscale. This is the level where there are too many components (such as cells or molecules) to simulate individually, but too few to trust in the law of averages. At the mesoscale, deviations from the average matter. "How elastic are arterial walls?" Karniadakis says. "The answer varies day by day, and across genders and ethnic groups. Even if you know the properties precisely, you need to know how they vary." But including variation in a model is harder than it sounds. It means abandoning the comfortable deterministic models of classical mathematics and using probabilities. The elasticity of an arterial wall is no longer a number but a distribution, a miniature bell-shaped curve of possible values.

Even cardiac models, which have performed well with deterministic equations, may need a dose of randomness. "There's a growing understanding that in some cases one has to do stochastic differential equations," says **James Keener, PhD**, a mathematician at the University of Utah. One such place is the modeling of calcium flow, which he says is "highly inhomogeneous" within the cell. The classical models, which treat the interior of the cell as a uniform fluid, may be getting the right results for the wrong reasons.

Keener's comment suggests a final word of caution about all mathematical models. Even the best-validated model is not guaranteed to last forever. It is always subject to correction, as experimenters discover new phenomena that weren't included in the original assumptions. "Models are never right," says Bassingthwaighte, "they're just not wrong yet." □

For one thing, "We have no sense of how error propagates from one level to the next," Arkin says. For another, he asks, "Where are the boundaries between fast and slow reactions, or between deterministic and stochastic models?"

is taking the next step by putting working versions of the Physiome Project models online. In theory, this will make it much easier to mix and match models at different scales.

However, there is more to multi-scale modeling than picking from a menu of single-scale models. Another thing you need is a lot of data. "Complex models have not caught on in biology the way they have caught on in, say, weather forecasting, because weather forecasters have sensors everywhere," says Arkin. By contrast, much biomedical research has to make do with few sensors and intermittent data. Some fields, on the other hand, are swimming in data—genomics and proteomics, for example—but do not have enough models that can handle that level of complexity.

BY JAMES COBURN, MS, AND JJ TREY CRISCO, PhD

Representing Rotations with Quaternions

Many tasks in biomedical data analysis, such as kinematic data collection, involve three-dimensional motion analysis which requires precise representation of an object's position and orientation. Mathematical operations such as interpolation, averaging and curve fitting are straightforward when applied to translation, but are troublesome when applied to rotation. Another descriptor is necessary to carry out all these operations on rotations.

Various approaches exist such as rotation matrices, Euler angles, helical (or screw) axis and quaternions, but some have significant limitations. The

Quaternion mathematics expands the possibilities of how we can represent and manipulate rotations.

first three all have restrictions limiting their usefulness for interpolating, averaging and curve fitting. Rotation matrices can drift numerically when repetitively multiplying matrices, resulting in undesired scaling or shearing. Euler angles can be defined in 12 different ways and each will give a different answer. They are also subject to gimbal lock, which is when two of the rotational axes align and you lose the ability to continue rotating freely. Both of the above are subject to tumbling during interpolation, where the object over-rotates while getting to the final orientation. Helical axis descriptions are handy for user interaction but

do not provide a unique way to combine rotations into a single desired rotation. Quaternions overcome these difficulties

Quaternions were discovered in 1843 by Sir William Hamilton after years of searching for a natural algebra of 3-D space. A quaternion is a complex four-component unit vector containing a single real component and three imaginary-like parts $q = \langle w, i j k \rangle$. As unit vectors, quaternions always lie on the surface of a hypersphere. This hypersphere represents all possible rotations and all the possible paths between rotations. The shortest distance on the sphere between two points (just

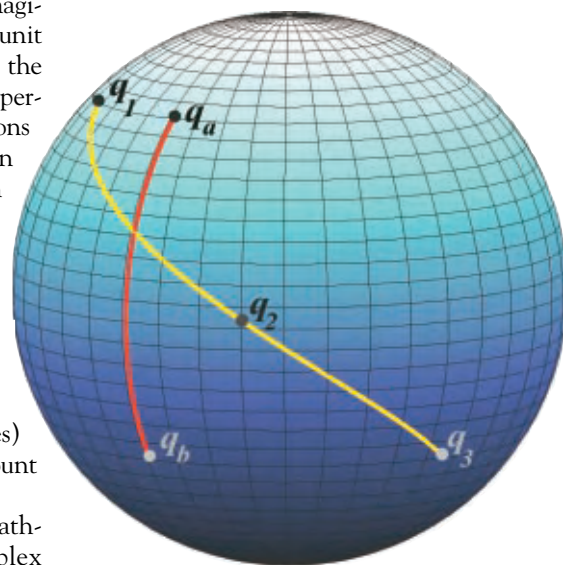
like the great circle airline routes) rotates the object the minimum amount to get to the final orientation.

Quaternions follow the same mathematical rules as standard complex numbers but in vector form. This means operations such as multiplication, division, powers, exponentials, or logarithms can be performed on quaternions. Not only that, interpolating and averaging rotational data is more realistic and more accurate.

Two rotations can be combined simply by multiplying two quaternions together. Averaging and curve fitting rotations become simple vector operations. Most importantly, because



quaternions lie on a surface, interpolations of 3-D rotations can be uniquely defined. Rotation between two orientations can be performed with spherical



linear interpolation on the surface of this sphere (red line). Cubic interpolation on the sphere surface can represent smooth rotation through multiple orientations (yellow line).

Quaternions provide notational convenience and also provide a deeper mathematical foundation for 3-D rotations. Quaternion mathematics expands the possibilities of how we can represent and manipulate rotations. In the lab, 3-D kinematic data analysis and computer motion simulation each require accurate and straightforward methods for calculating rotational data. Quaternion mathematics supply both. When analyzing 3-D kinematics, quaternions provide simple and accurate answers.

For further reading on quaternions, see: Shoemaker, K. "Animating Rotation with Quaternion Curves." *Computer Graphics* 19, 245-254, 1985. □

DETAILS

James Coburn, PhD, received his masters degree in mechanical engineering with the Orthopedic Biomechanics Laboratory at Brown University, and is close to completing his PhD. The research focus of the lab is musculoskeletal injury and repair ranging from studying muscle injury and ligament mechanics, to *in vivo* joint kinematics. Trey Crisco is an associate professor in orthopaedics and adjunct professor in engineering. He directs the Bioengineering Laboratory at Brown University.

BY KRISTIN COBB, PhD

The Institute for Systems Biology

The Institute for Systems Biology (ISB) was founded in Seattle, Washington, in 2000 by **Leroy Hood, MD, PhD, Alan Aderem, PhD, and Reudi Aebersold, PhD**. Five years later, they are pursuing the frontiers of systems biology in an interdisciplinary, non-academic environment with 170 staff members and a 65,000 square foot building.

Biomedical Computation Review spoke to Leroy Hood, president of ISB.

Q. How do you define systems biology?

A. Systems biology is the window by which biological circuitry can be deciphered. If you broke a radio into its parts, this would not give you insight into how the radio works. Rather, you would have to develop techniques for seeing how the parts connect together into circuits and how these circuits interact to convert radio waves into sound. Living organisms also operate by virtue of networks that function together. You need systems approaches to define the networks and understand how the networks interact to carry out biological functions. You have to look at the whole system and see what it's doing as you tweak it in various ways. The old fashioned approach of studying one gene and one protein at a time is inadequate.

Q. What was the motivation for setting up the Institute for Systems Biology?

A. The institute was set up to learn how to do systems approaches to biology. It was set up in a cross-disciplinary environment because we are con-



Leroy Hood, MD, PhD, president of ISB. Courtesy: ISB

vinced that biology needs to drive technology and computation. None of the tools we have today are adequate to the task. It's critical that you not just practice the biology but you invent the future with the new technology.

Q. What was the rationale for moving outside of academia?

A. The bureaucracy of classic academia was getting in the way. We needed to create the cross-disciplinary environment, high-throughput facilities, and computational infrastructure. We needed the flexibility to negotiate strategic partnerships quickly and effectively. We've been responsible in one way or another for the spin-off of nine companies in five years, and we've started a big K-12 science program. Those are things we could not have done at a university.

Q. Can you give an example of systems biology at work in your lab?

A. The halobacterium project, led by **Nitin Baliga**, is very interesting. Halobacterium is a simple organism with 2400 genes and only about 125 transcription factors. We perturb some of its most interesting circuits and see what it does: it's resistant to radiation, so we give it radiation; it deals with metals very well, so we give it large doses of metals; it shifts how it makes its ATP when oxygen is low, so we put it in a low-oxygen environment. From these experiments,



A lab meeting in one of the Institute's many shared spaces that encourages cross-disciplinary interactions. Courtesy: ISB

we've gotten this first glimpse of the interconnected networks that exist. And we've been forced to develop new computational approaches to integrate these data. These tools will all be extended up to the higher organisms. The simpler organisms drive—if not an understanding of the higher organisms—the technology that will enable us to understand them.



The Institute, near Lake Union. Courtesy: ISB

Q. What is some of the work that you've done in prostate cancer?

A. Up until recently, we've done very nice work to show unequivocally that prostate cancer is a genetically heterogeneous disease. Now we are taking tumors and their normal counterpart tissue and looking at DNA arrays to see how the patterns of gene expression have changed. We've compared prostate cancer tissue at an early stage, an intermediate stage,

you actually can cure them. So we'll see if we can do early detection and save people's lives.

Q. What advice do you have for young scientists?

A. I think anybody who doesn't want to be left behind is going to be forced to move toward systems biology. And I think a lot of this will come from younger people, who recognize what the future is. So my advice to them is, first, it's really important for

It's critical that you not just practice the biology but you invent the future with the new technology.

a late stage, and a metastatic stage with normal tissue. Really recently we've begun to delineate the prostate-specific secreted molecules that we believe will constitute molecular fingerprints for telling us the exact state of the prostate: Is it normal? Is it inflamed? Is it hypertrophied? If it has cancer, which of the four or five types of cancer does it have? The idea is that the blood is going to be a window into health and disease.

Q. How far along are we in being able to use molecular fingerprints in the blood to diagnose disease?

A. We are at the very beginning. That's the vision. The vision is that we can computationally make predictions about the molecular fingerprints that will exist for each of the different organs. We've now set up a company (Homestead Clinical Corporation) that is beginning to search for these molecular fingerprints. I think within a year we'll have three to five markers for prostate cancer and three to five for ovarian cancer. We're also starting to do blood molecular fingerprinting analysis in brain cancers. Glioblastomas are an irreversible death sentence, but that's only because they're detected so late. If you can detect them early,

young scientists to learn to think in a more global way. We have many lab meetings where I will come in and say: "Here's an idea we're going to talk about." And everybody will have to say what they think about that idea. It's important to have your focus and do your thing but it's also important to think in a big way. Second, it's really important to enjoy what you are doing. Passion is what makes science fun. Finally, I think this is the most exciting time in science in the 40 years of my career. In some ways I wish I were young again so I could start all over. On the other hand, I'm happy where I am now. Maybe I wouldn't do as well the second time around. □

ISB QUICK FACTS

What: Non-profit research institute
Where: Seattle, Washington
Faculty: 11 members
Staff: 170 members
Facility: 65,000 square feet
Funding sources: 80% federal grants & contracts;
20% grants from foundations, industry, individuals
Annual symposium: every spring

Biomedical Computation Review

Stanford University

318 Campus Drive

Clark Center Room S170

Stanford, CA 94305-5444

seeing science

SeeingScience

BY KATHARINE MILLER, MANAGING EDITOR

Decoding the Decoder

Inside a cell, the ribosome deciphers genetic codes to produce proteins at unfathomable speeds. Now, researchers at Los Alamos National Laboratory (LANL) have simulated this complex nano-machine in action. With 2.6 million atoms moving at once, it's the largest molecular simulation ever attempted by a factor of six. The work by **Kevin Sanbonmatsu, PhD**, a computational structural biologist, and his colleagues at LANL revealed some new details about the essential translating molecule, transfer RNA: it must be flexible in two places for decoding to occur. The simulation also identified a new structural gate, which may act as a control mechanism for selecting the proper transfer RNA.

The ribosome (large subunit in white and small subunit in cyan) uses the transfer RNA molecules (incoming in red; outgoing in yellow) to read the genetic information from the messenger RNA (green) to produce protein. For visualization purposes, the top portion of the ribosome is cut away. Courtesy of LANL.

See a Quicktime movie showing transfer RNA at: <http://www.lanl.gov/news/images/sanbonmatsu.mov>

