BiomedicalComputation
REVIEW

Simbios:

Bringing Biomedical Simulation to Your Fingertips

Fall 2009

contents

ContentsFall 2009

Cover Art: Created by Rachel Jones of Wink Design Studio using
Simbios images and hands image © Glo5 | Dreamstime.com.
Page 3: Supercomputer image is © Heizfrosch | Dreamstime.com, black servers image is © Jbron | Dreamstime.com.
Page 28, 29: Dashboard image is © Leonid Dedukh | Dreamstime.com, engine image is © Mashe | Dreamstime.com.

BY RUSS ALTMAN, MD, PhD

Biomedical Computation Review:
The Simbios
Fifth Anniversary Issue

Dear Reader,

In this eighteenth issue of Biomedical Computation Review (BCR), we bring you a special edition devoted to the work of the magazine's publisher: the Simbios National Center for Physics-Based Simulation of Biological Structures (http://simbios.stanford.edu/).

The magazine itself is one of the innovations Simbios introduced to help build a national community of biomedical investigators with a devotion to computation. But Simbios' national presence is also being built in other ways. The initial Simbios team, formed in 2004, was based at Stanford and included biologists, biomedical computation and informatics scientists, and computer scientists. We therefore made special attempts to grow from the strong core at Stanford to the national and international community. Taking our "national center" mandate seriously, we set out to serve a much larger com-

munity of biomedical scientists who could benefit from software technologies that brought easier and better ways to use simulation in the pursuit of new knowledge.

This issue's feature story, appropriately titled "Simbios: Bringing Biomedical Simulation to Your Fingertips," describes some of the technologies and projects that Simbios created in its first five years and how Simbios tools are being adopted by the research community well beyond Stanford.

The story focuses primarily on three of the center's "driving biological problems" or DBPs—the problems that drove us to develop tools that would matter for bioscience. Each of these DBPs—neuromuscular, protein-folding, and cardiovascular—introduced at least one key open-source computational tool during Simbios' first five years. These include OpenSim, a neuromuscular simulation package for simulating human movement; Open Molecular Mechanics (OpenMM), which permits the use of very fast GPUs for molecular dynamics simulations, including pro-

Map of Simbios collaborators. Red markers indicate institutions that have worked with Simbios during the last five years: some collaborated on a driving biological project; others received a collaborative R01 grant or seed funding from Simbios; others have made significant contributions to our software offerings; and still others were lead or senior authors on a Simbios publication.

tein-folding; and SimVascular, for simulating blood flow velocities and pressures in deformable vessels.

Simbios researchers are using these tools to advance their own research goals, but Simbios' reach extends much farther, as you'll see from the feature's focused stories about 13 people who are Simbios collaborators, tool users and/or alumni. These individuals describe how they are using Simbios tools to further their biomedical research goals. We hope their stories inspire more of you to sample Simbios' wares.

Also within these covers you'll find our regular Point/Counterpoint column where Vijay Pande highlights key issues in the debate between Supercomputers and Superclusters and which is better suited to the needs of molecular simulations.

And in this issue's "Under the Hood" column, our chief software architect, Michael Sherman, provides insight into how the incredibly complex technologies required for fast, accurate simulations can be bundled into software packages that are usable by domain experts.

Some basic principles that have guided much of the Simbios software development include the following:

• Simulations must be fast, in order to yield results of potential interest on a time-scale compatible with scientific progress. Accuracy must be measured and documented, but given a specification of tolerances in accuracy and precision, simulation is all about speed.

• "General purpose" applications are nice for software engineers to create, but are not generally useful to domain scientists who want narrowly defined tools with narrowly defined capabilities, tuned to their domain. This means that an underlying toolkit must be available, so that narrow applications can be created quickly to address this need, while not causing the software team to endlessly reinvent the wheel.

• Putting the above two rules together is the special sauce: the very complex technologies required to make simulations fast and accurate must be totally hidden from the users, who must be presented with application-specific, easy to use software

At the end of the day, the National Centers for Biomedical Computing are charged with catalyzing high-impact research in biomedicine, and with making tools available for the community to use. To that end, Simbios created Simtk.org, a repository for software, data and models related to physics-based simulation. On Simtk.org, we distribute the open source software products we've developed. And any other biomedical investigator may do the same. Simtk.org is intended to gain momentum and have a sustainable life even after the end of the Simbios center. As of this writing, Simtk.org had nearly 350 (348 to be exact) software projects, including eight that are actively being disseminated in our workshop and tutorial program. Over 6500 users have registered on the site to download our software, and that number continues to grow. In fact, for each of the last few years, we have seen a doubling in the membership growth rate and currently have an average of 400 new users signing up each month.

As you can see, I am thrilled to report that in its first five years, Simbios has made great strides in our key missions. Our scientists have made important, high-impact contributions to their domains of research, while also creating software, data, and models that can be used by others in the field. We have hardened key software (such as those described above) in order to guarantee utility to the community. We have engaged in a blitz of workshops and tutorial sessions in order to disseminate these programs, and show the target audience how to use them. Our singular devotion to physics-based simulation has also attracted a network of collaborators (see map) who are co-developing software projects, using them in their research or engaging as serious "alpha" testers to see if the software is performing well.

The task of becoming a National Center is not an easy one. It is not sufficient simply to post a sign and declare national center-dom. Instead, a national center must start with service. By building a core of useful technologies, generating compelling scientific results, and then making the technologies available along with training materials, we can best serve the scientific community. We begin generating the core on the inside, but we make our impact by bringing it outside. □

> Simbios has made great strides in our key missions. Our scientists have made important, high-impact contributions to their domains of research, while also creating software, data, and models that can be used by others in the field.

Supercomputer VERSUS Supercluster

Say you are performing biomolecular investigations that are extremely compute intensive. You have a finite amount of money and time. You could get (1) a supercomputer (fast custom CPUs and high-speed interconnect facilitating parallelization of a single computation) or (2) a big cluster (lots of cheap commodity CPUs with slower commodity networking). What approach will deliver the most science per resources consumed? Professor Vijay Pande overheard the two types of systems arguing the pros and cons and recorded it word-for word here …

Supercomputer: Supercomputers have reached a point where they are very powerful, are readily available via supercomputer centers, and can do it all. In terms of physics-based simulation, models have become sufficiently accurate that we can make useful, quantitative predictions.

Supercluster: Yes, no argument about the utility of simulations and the need for large-scale computer resources, but traditional supercomputers are expensive and the network can cost as much as the processors. For those of us with limited budgets, buying lots of computers to build a "super-cluster" with twice the computing power and a cheaper network seems more cost effective.

Supercomputer: Well, ok, but what would you do with all those processors without a fast network? Without the fast network, the processors are useless because they can't work together. For example, in molecular simulation, one can only simulate about a nanosecond (10^{-9} second), maybe 10 ns, in a day on a single processor. That's not going to get you far. Using a tightly coupled (i.e., fast network, supercomputer)

machine, one can use thousands of processors' cores to get orders of magnitude longer simulations.

Supercluster: Yes, you can, but not very efficiently. Even these fast networks aren't truly instantaneous so there can be a heavy overhead cost associated with processors communicating. For example, if one needs to

> **Supercomputer: "Supercomputers have reached a point where they are very powerful, are readily available via supercomputer centers, and can do it all."**

simulate a 10,000 atom system, a 10,000-core supercomputer is likely not going to be useful, since breaking the problem up that small won't scale. In other words, the processors will spend too much time communicating and not enough time calculating.

Supercomputer: True, but scalability is a classic problem in computer science and the solution is a *faster* network, not a slower one. I don't see your point.

Supercluster: Well, there is another way for many types of simulations interested in kinetic or thermodynamic properties, such as simulations on the molecular scale. Instead of using the processors to work together to

simulate a single long trajectory (i.e., a single simulation of a protein going through its dynamics), one could run many shorter simulations in parallel. Since molecular processes are inherently stochastic, one would be able to use multiple, *independent* simulations to get a performance boost *without* a fast network.

Supercomputer: Yes, I've heard of these tricks. The problem is that they have a very limited regime of applicability. These methods assume single exponential

Supercluster: True, for very, very large systems (which, for example, could not even be run on single processors due to memory or other constraints), scaling works reasonably well on multi-core boxes present in superclusters or on multiple cores in supercomputers, so I suggest a combination of methods: scale as far as one can go well (i.e., linearly, where doubling the number of processors doubles the speed), and then use additional methods on top of that. The MSM approach, for example, just seeks to use trajectories as efficiently as possible,

> Supercluster: "For those of us with limited budgets,
> buying lots of computers to build a "super-cluster"
> with twice the computing power and a cheaper
> network seems more cost effective."

kinetics, i.e., that the probability of an event occurring after t nanoseconds of simulation time looks like $p(t) = k \exp^{-kt}$. This distribution is sharply peaked at short times, so one could see events even at times much shorter than the average time $<t> = 1/k$. The problem here is that most complex systems have multiple states so this simple two-state approximation would break down.

Supercluster: A simple method like that would break down. However, over the last five years or so, several groups have been working on a much more sophisticated method, called Markov State Models (MSMs). Here, one combines many (relatively short) simulation trajectories with Bayesian statistics to build a kinetic model of the process of interest. One does not have to make approximations regarding single exponential kinetics of the overall system, assume reaction coordinates for the system, *a priori* identify initial and final states, etc. Recent advances in adaptive methods where one builds an MSM by gathering some simulation information, then adaptively decides where to run new simulations in order to optimize some property (such as minimizing the uncertainty in some variable of interest) have shown that this approach can be *more* efficient than a few long runs, *even if* there were no cost for traditional parallelization. This is because MSMs are much more efficient at skipping over traps and other places where simulations simply "wait" for some rare stochastic event.

Supercomputer: OK, but what about large systems that would take a lot of computing power to generate even short trajectories? Even if you use cool methods to combine short trajectories, you are bound to be stuck if the trajectories are too short. And this can be a problem for large systems, where generating even short trajectories can be a challenge on a single CPU core or even a multi-core CPU.

but there are other possible synergies as well. The generation of those trajectories still becomes an interesting challenge for the future.

Supercomputer: Isn't there another alternative? I think you're forgetting about GPUs. While traditional processors (i.e., CPUs) have not been getting faster (just with more CPU cores packed on a single chip), GPUs have been getting more and more powerful, in part due to their unique architecture with lots of floating point units for scientific calculations. Indeed, GPUs remind me of supercomputers of the past, like the old Cray vector supercomputers, which could do certain types of heavy floating point calculations quickly, with some coding effort to take advantage of this unique hardware. Indeed, this opens the door to smaller *personal* supercomputers like a GPU-accelerated under-your-desk minicluster. That can pack a lot of cycles into a small space and you don't have to share it with anyone. It's cheap and you can easily replace it with the next generation.

Supercluster: I hate to agree with you, but it's true: Both the supercomputer and megacluster have problems associated with unleveraged acquisition of capital equipment, real estate, cooling, and power. Maintenance, obsolescence, and other "big science" issues could be avoided by equipping each scientist with a smaller, personalized resource. Scientists can afford even small GPU clusters which can be quite powerful. Of course, the next question will be how one wants to use all of those GPUs in parallel! □

News Bytes

Studying Force in 3-D

Mechanical forces drive many processes in the human body, from organ and tissue formation during development, to stem cell differentiation, to wound healing. Until recently, scientists could only study these forces at the single cell level in two-dimensional experimental models. Now, researchers have developed a new tool and computer model to study forces generated by cells in 3-D contexts at a very small scale.

"This is the first time we have been able to measure three-dimensional forces in very small structures or with a small number of cells," says **Christopher Chen, PhD**, bioengineering professor at the University of Pennsylvania and senior author of the work that appeared in a June 2009 issue of *Proceedings of the National Academy of Sciences*.

Chen's group measured forces exerted

> "This is the first time we have been able to measure
> three-dimensional forces in very small structures or
> with a small number of cells," says Christopher Chen.

by human and mouse fibroblasts, a type of cell that is abundant in connective tissue in the body. Fibroblasts secrete proteins to form an extra-cellular matrix that binds cells together into tissues. "All cells aside from blood cells are adhered to this matrix, sort of like a carpeting that they're embedded in, and they're pulling against that matrix," Chen says. "When they feel those forces, there's a fair amount of data to suggest that they change their behavior."

The measurement tool Chen's group constructed contains two tiny cantilevers connected to a sensor with a collagen gel between them. Chen's group then put fibroblasts into the collagen goo. When fibroblasts hit collagen, they contract and reorganize the collagen fibers, Chen says, and this results in a mix of collagen and cells suspended between the cantilever rods, like a hammock. The scientists then measured the force from that contraction. They also varied conditions in the set up, such as the thickness of the collagen and the stiffness of the cantilever springs, and looked at how the cells reacted. The stiffer the springs, the more the fibroblasts contracted. And the more contractile forces the cells encountered, the more extra-cellular matrix they pumped out. In the body, this reaction to force is useful. For example in wound healing, the fibroblasts sense the tension from the wound edges pulling apart and secrete more matrix to form scar tissue.

Chen and his group then constructed a computational model to better understand the distribution of force within the collagen mass. They found that the points in the structure where their model predicted the highest stress correlated with the most production of extra-cellular matrix. The model can also be useful for predicting forces in more complicated geometrical structures, more like those found in the body, Chen says.
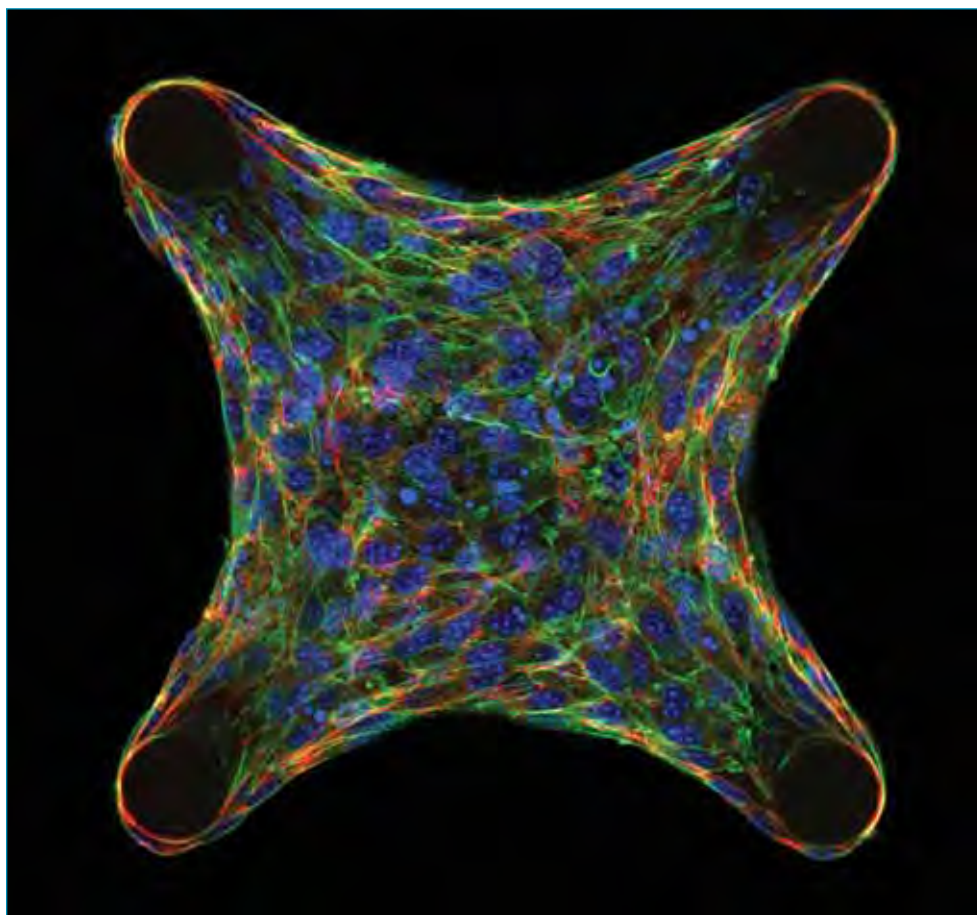
Fluorescent image of fibroblast cells embedded in a collagen matrix suspended between four small rods. Chen's study measured the force exerted by these cells using sensors at the small rods, and used computational models to predict the patterns of force throughout the microtissue. Cell nuclei are shown in blue, the cytoskeleton protein actin in green, and structural matrix proteins in red. Image courtesy of Wesley R. Legant.

Snapshots of a simulation of a dendrimer-DNA complex arriving and docking at a model cell membrane. Reprinted with permission from **The Journal of Chemical Physics 130, 155101, 2009. Copyright 2009, American Institute of Physics.**

"There's a growing appreciation of how important mechanical forces are for many biological processes ranging from directing stem cell differentiation, to tissue formation, to how cells respond to drugs," says **Ali Khademhosseini, PhD**, an assistant professor of medicine at Harvard and MIT. "This work generates a powerful model that can be used for many different applications. There's a lot of scientific follow-up as well as many potential technological and engineering advances that can come out of it."
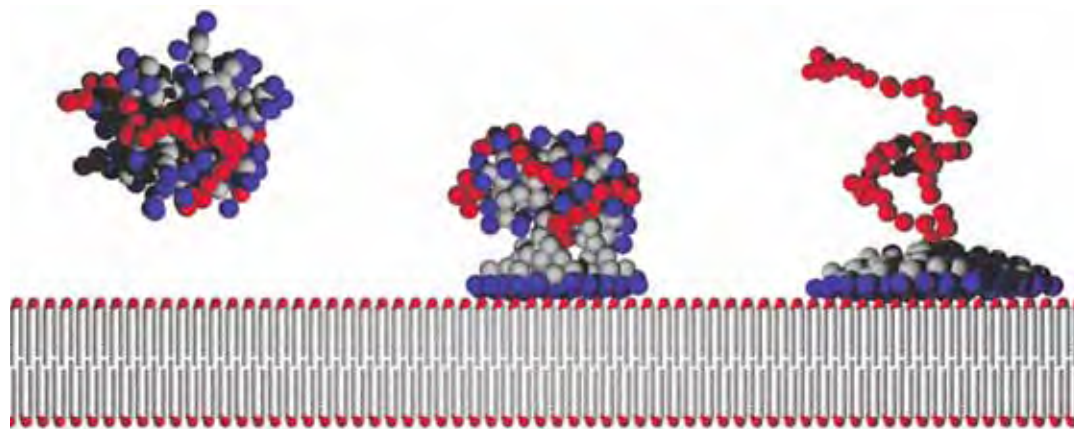
—By Rachel Tompa, PhD

Modeling A Gene Therapy Delivery Vehicle

Gene therapy to correct inherited illnesses hinges on successful delivery of DNA into a person's cells. Most gene therapists work with viruses to ferry their DNA cargo. Yet the body tends to fight even disarmed viruses that should be harmless. As an alternative, researchers have devised dendrimers, branched molecules whose endings can be tailored to package DNA. Now, in the first molecu-

> "With our simple tinker toy model we're going to throw out a lot of information that is certainly important, but it gives us the basic physics," says Paul Welch.

lar-level simulation of a gene therapy vector in action, researchers have simulated a dendrimer docking at a model cell surface and shown how long it can hold on to its DNA cargo.

The simulation rendered a quick but clear sketch of what happens at the cell membrane. "With our simple tinker toy model we're going to throw out a lot of information that is certainly important, but it gives us the basic physics," says **Paul Welch, PhD**, a materials physicist at Los Alamos National Laboratory and lead author on the study published in the April 2009 issue of *The Journal of Chemical Physics*.

In previous work, other researchers have modeled dendrimers interacting with membranes, but no one had simulated them transporting DNA.

Welch and his team created a molecular dynamics model of a dendrimer with an attached DNA strand. In their simulations, they let the dendrimer-DNA complex loose near a simple, planar membrane model to see whether it would bind or wander off. They found that both the propensity to bind and the duration of binding decreased in the presence of a more negatively charged membrane. There is a range of surface charges which allow binding for the optimal length of time— long enough for the complex to transit the membrane but not so long that the dendrimer retains a grip on the DNA after entry. In addition, the researchers found that big burly dendrimers are not necessarily the best delivery vehicles for DNA. In future simulations, Welch's team hopes to use a more realistic model of the membrane's lipid bilayer. Ideally, Welch says, the membrane would undulate, deform and perhaps form a little liposome (bubble) around the complex to pull it in, much as one would expect a membrane to behave in nature.

Ron Larson, PhD, a polymer physicist at the University of Michigan in Ann Arbor who models the use of dendrimers to poke holes in membranes to kill bacteria or deliver drugs, wonders whether the model should address possible interactions between the four bases of the DNA and the membrane. And he looks forward to experiments that would test the model. "People make these different particles by the seat of their pants and see how many go in," Larson says. "When things go wrong, they often don't know why. It's really helpful to have a theoretical model."
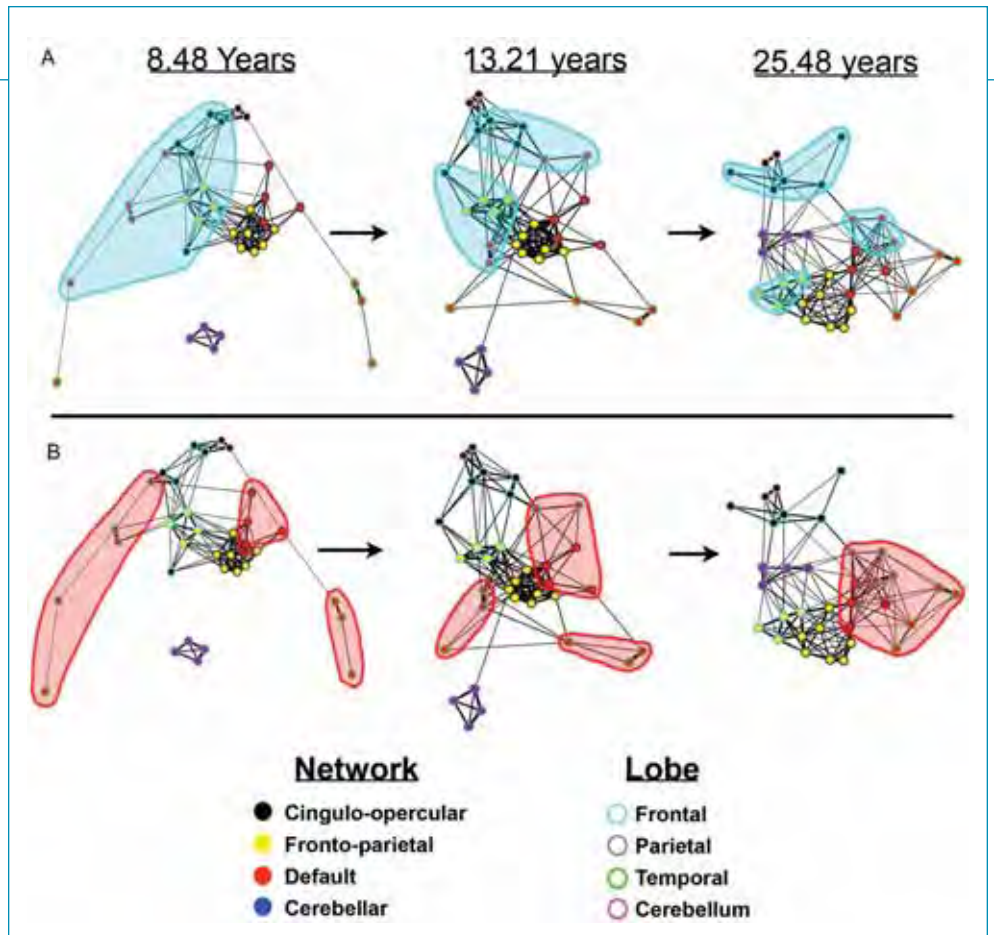
—By Roberta Friedman, PhD

Different But Equal

Kids often claim they are just as smart—if not smarter—than their parents. Childish nonsense? Perhaps not, according to a recent study. It turns out that young children's brains are as efficient in solving information-processing tasks as their adult versions, despite being very differently organized. This finding could improve our understanding of normal brain development as well as of disorders such as autism and Tourette syndrome.

"Whether you are a kid or an adult your brain is organized in a pretty damn efficient way," says **Steven Peterson, PhD**, a neurophysiologist at the Washington University School of Medicine in St. Louis, and senior author of the study which appeared in the May 2009 issue of *PLoS Computational Biology*.

By seven years of age, a typical human brain has already attained 95 percent of its adult size and most of the wiring that connects neurons to each other is already in place. As the brain matures further, two things happen: its wires (axons) get better at transmitting signals and its unused junctions (synapses) are progressively trimmed out. Along with these physical changes, the brain changes the way it configures its various regions into functional networks during resting, reading, singing, walking, or other tasks. This phenomenon is a key to both normal and abnormal brain development, but until now it has been difficult to quantify.

In the new study, Petersen and his team used magnetic resonance imaging to study functional brain connectivity in a sample population of 210 subjects aged 7-31 years. When they cross-correlated the temporal activity of 34 key brain regions in each subject while resting, a clear pattern emerged: brain regions in children interacted mostly with their neighbors while those in adults enjoyed longer-range interactions. While this confirmed prior theories, further quantitative analyses turned up a surprise. Functional brain networks in both adult and child subjects proved to consist of tightly knit communities loosely linked to each other—both possessing a "small world" structure that typifies efficiently connected systems such as social networks and the Internet. Although these communities start off being spatially localized in children and grow more diffuse with maturity, measures of computational efficiency remain high throughout. "All of us were surprised when those numbers came out," says Petersen, who notes that these findings have now been replicated by Stanford University neuroscientist Vinod Menon's research team

in the July 2009 issue of *PLoS Biology*.

Petersen and his colleagues, including pediatric neurologist **Bradley Schlaggar, MD, PhD**, have applied this methodology to study Tourette syndrome, a neurological disorder characterized by physical and vocal tics. Brain connectivity patterns in adolescent Tourette sufferers appear to lag by 2-3 years compared to

normal, says Petersen. "The context we got from studying normal development allowed us to interpret what we observed in Tourette subjects."

"This study is incredibly innovative in providing original and direct evidence about how circuits are formed in the brain," says **Beatriz Luna, PhD**, a developmental psychologist at the

Functional brain networks in both adult and child subjects consist of tightly knit communities loosely linked to each other—both possessing a "small world" structure that typifies efficiently connected systems such as social networks and the Internet.

NewsBytes

University of Pittsburgh Medical Center. Luna suggests that the method could next be applied to study subjects engaged in specific activities. **BJ Casey, PhD**, a neuroscientist at the Weill Medical College of Cornell University in Ithaca, New York, finds the study to be "a very novel characterization of neural system development" ideally suited to study developmental disorders such as autism. "It's going to drive a lot of research," she says.

—*By Chandra Shekhar, PhD*

Chromatin Fiber: Zigzag or Solenoid?

Try packing a two-meter-long stretch of DNA into a cell nucleus just a few millionths of a meter thick—with key coding segments readily accessible. It's a seemingly impossible feat that eukaryot-ic cells routinely pull off by building a highly compact, fibrous mix of DNA and proteins called chromatin. Now a new study uses a combination of novel lab experiments and computer simulations to provide long-sought details about the structure of chromatin fibers.

"Our study appears to resolve a 30-year-old controversy about the structure of chromatin fiber," says **Gaurav Arya, PhD**, assistant professor of nanoengineering at the University of California, San Diego. The findings, published in the August 11 issue of *Proceedings of the National Academy of Sciences*, could improve our understanding of cell growth, differentiation, and cancer.

This much is generally accepted: Chromatin starts off as a series of nucleosomes—protein spindles wrapped with about a turn and a half of DNA—connected by stretches of linker DNA; this "beads on a string" structure then folds itself into stiff, compact fibers. What is debated is the interaction and arrangement of nucleosomes within this fiber.
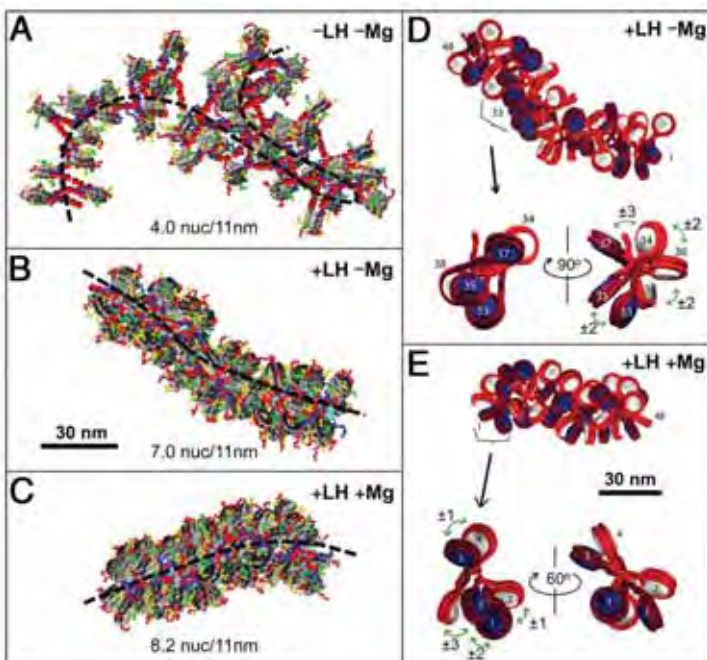
One school of thought favors a spiral arrangement, or solenoid, in which successive nucleosomes interact and are connected with bent DNA linkers. Another school argues that DNA is too stiff to bend easily, and proposes instead a zigzag structure with straight linkers in which alternate nucleosomes interact. Until now, this issue could not be resolved because the available experimental techniques required the chromatin fiber to be unwrapped before it could be studied.

In the new work, researchers first used formaldehyde to create permanent cross-links between interacting nucleosomes. These interactions give rise to loops in the fiber when it is unwrapped under various conditions. Studying these loops under an electron microscope, the researchers found evidence to support the existence of the zigzag structure in the absence of divalent ions such as magnesium; in the presence of such ions, however, a fraction of nucleosomes switch to the solenoid motif.

The researchers then used a computational model developed by New York University researcher **Tamar Schlick, PhD**, to simulate the structure of chromatin fiber. The model confirmed the experimental results and added additional details: Without divalent ions present, the zigzag fiber packs about 7 nucleosomes per 11nm stretch; with divalent ions, about 20 percent of the linkers in the fiber bend, solenoid-style, and this helps the fiber accommodate about 8 nucleosomes per 11nm.

> "Our study appears to resolve a 30-year-old controversy about the structure of chromatin fiber," says Gaurav Arya.

"These results show that both the zigzag and solenoid topologies may be simultaneously present in chromatin fiber," says the study's lead experimentalist, **Sergei Grigoryev, PhD**, associate professor of biochemistry and molecular biology at Pennsylvania State University. "It's very exciting that we could show this using both computational and experimental techniques."

University of Wyoming molecular biologist **Jordanka Zlatanova, PhD**, who has been studying chromatin for more than 30 years, says the paper is an important contribution "because finally we seem to really understand what the

Chromatin packing gets denser with the addition of linker histones (LH) and divalent ions (Mg) in this computational simulation (A-C) . In the close-ups at right, the cores of alternate nucleosomes have different coloring (white or blue) with red linkers for better visualization. The zigzag structure dominates at low ionic concentrations (D) but in the presence of magnesium chloride, several nucleosomes have bent linkers and the nucleosomes interact in more of a solenoid arrangement (E). Reprinted from Grigoryev, S, et al., Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions, Proceedings of the National Academy of Sciences, 106: 32:13317-13322 (2009).

chromatin fiber structure is." It's also a major advance experimentally, she says, because it captures nucleosome interactions under physiological conditions. Further, no other group has been able to come up with a computational model that fits the native structure of chromatin so well, she says.

—By Chandra Shekhar, PhD

Predicting Cancer Treatment Success

No two cancer patients respond identically to treatment. Some will be cured while others will see their cancer return, and physicians are at a loss to explain why. Now, using MRI imaging researchers have developed a mathematical model of tumor growth that identifies two factors that are predictive of cervical cancer treatment success: responsiveness to radiation and the ability to clear dead cells.

"This work gives us strategies to find out early on if the tumor does not respond to cancer therapy … and to adjust treatment to increase the chance of cure," says, **Nina Mayr, MD**, radiation oncologist at Ohio State University and principal investigator of the study. The work was presented at the annual meeting of the American Association of Physicists in Medicine.

Currently, "little is known about the underlying biological mechanisms that govern the tumor response to radiation therapy," says **Zhibin Huang, PhD**, a postdoctoral researcher at Ohio State University and lead author of the study. "We wanted to see if this imaging technology could find some early indications of the outcome."

The research group, headed by **Jian Z. Wang, PhD**, medical physicist and the director of the Radiation Response Modeling Program at the Ohio State University, followed 80 women with various stages of cervical cancer—with tumors ranging from the size of a cherry to the size of a grapefruit. All of the patients received MRI scans before, during and after radiation therapy—the standard treatment for cervical cancer. With these scans, the researchers could measure the change in tumor volume over the course of the cancer therapy.

The team developed a mathematical model to fit the tumor volume data from the MRI scans and, using this model, identified two factors that correlated with the likelihood of a

> "This work gives us strategies to find out early on if the tumor does not respond to cancer therapy … and to adjust treatment to increase the chance of cure," says Nina Mayr.

patient's cancer returning. The first is the patient's radiation sensitivity—essentially, the percentage of the cells that survived the radiation dose. The higher this number, the worse the outcome. More specifically,

Ohio State University researchers used magnetic resonance imaging and a mathematical model to predict cancer recurrence. These images show decreasing tumor volume over a 5 week radiation course in a patient who was alive and cancer free 9 years later. Photo Credit: Dr. William Yuh and Dr. Nina Mayr.

Only a very small number of transdermal drugs have made it to market.
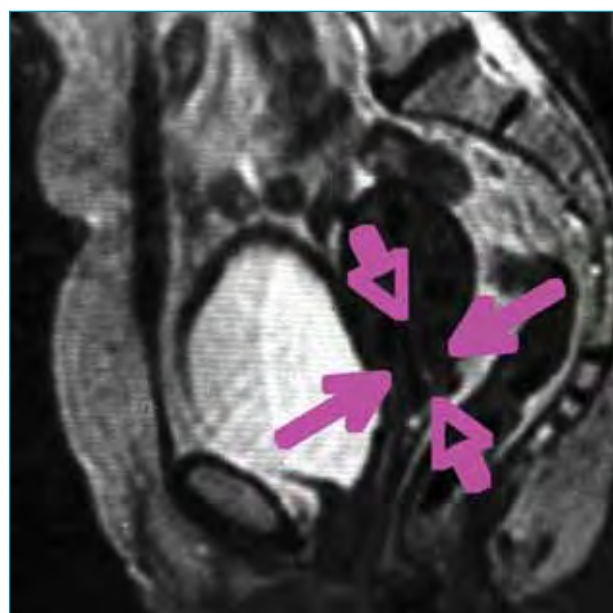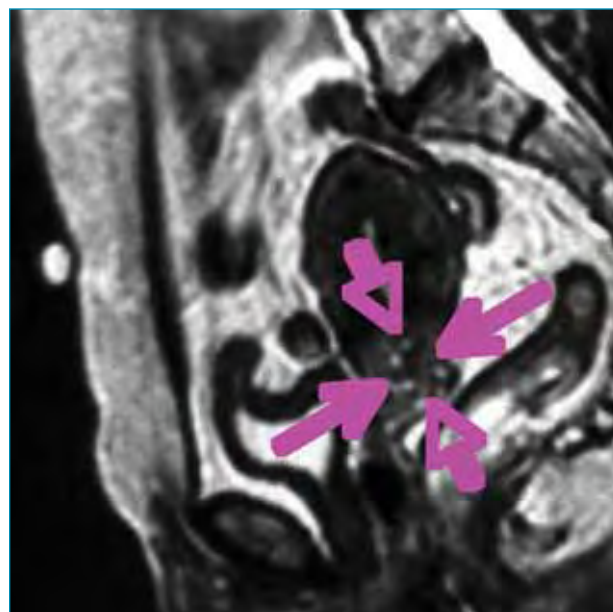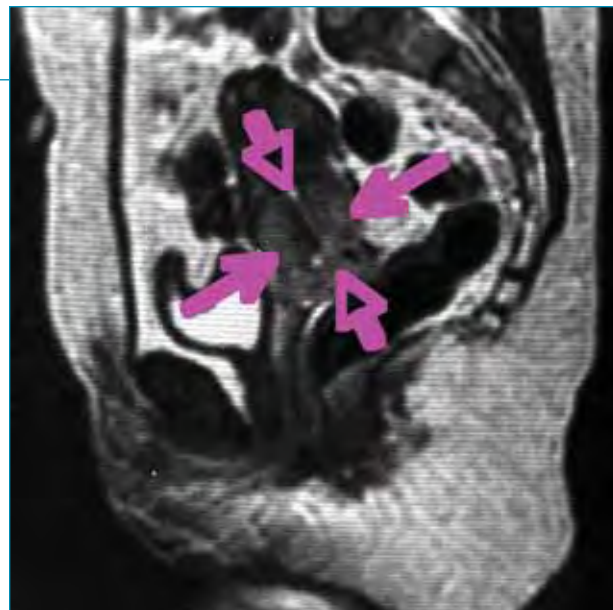
if radiation killed 30 percent or more of a woman's tumor cells during each day of treatment, then she is 33 percent more likely to be cancer-free than a woman whose tumor is more resistant to the radiation. The second factor is how quickly the dead cells are removed from the tumor area. For example, if it takes more than 22 days to clear the dead tumor cells after treatment, then that woman is nearly four times as likely to have her cancer return later on compared to a woman whose body clears the dead cells more quickly. When these two factors indicate that the cancer is likely to return, alternative treatments may be suggested for the patient. "Maybe we can use a more aggressive intervention instead," Wang says.

This is a very active area of research, says **William Small, Jr**, **MD**, professor of radiation oncology at Northwestern University Medical School. This kind of modeling could potentially be applied to other types of cancers treated with radiation. "It is very important to try to identify outcomes with surrogate markers," he says. "Doing so could allow us to finish clinical trials much quicker and dramatically improve our ability to test new therapies."

— By Liz Savage

A Multi-scale Model of Drug Delivery Through the Skin

Medicinal patches applied to the skin are an attractive route for drug delivery since they can release medicine slowly into the bloodstream and avoid being metabolized by the digestive system. Yet only a handful of medications have ever made it to market in patch form, largely because the stratum corneum, the skin's top layer, acts as a barrier. A new multi-scale computational model describing how chemicals move through that layer could help change that, opening the door to development of patches for a wider variety of drugs. The work was published in the June 2009 *Annals of Biomedical Engineering,*

"We believe that by separating the contributions from different levels of scale, the model can provide better insight about the barrier properties of the skin," says **Jee Rim**, **PhD**, a postdoctoral researcher at the University of California, Los Angeles, who did the research in collaboration with pharmaceutical company Alza while he was at Stanford.

Rim's model began on the micro level, with a molecular dynamics simulation of how a drug molecule diffuses along the middle of lipid bilayers like those in the stratum corneum. Since the lipid bilayers actually weave around impermeable cells called corneocytes, the next step was to consider the path the drug must take to avoid these cells and modify the diffusion coefficient (determined by the molecular dynamics simulations) to account for the behavior. The final step zoomed out to an even larger scale, modeling the effects of boundaries between the patch, the stratum corneum, and the rest of the epidermis.

"One of the insights gained from the study is that the stratum corneum acts as a regulating layer," says Rim. The slow journey of molecules through the lipid labyrinth seems to be the main reason drugs delivered via patch keep such a stable concentration in the blood.

Experiments with cadaver skin showed similar diffusion parameters to the ones calculated by the model. Drugs moved somewhat slower through the skin samples than the model, which the authors think may be due to water content. "[A future goal] is to consider the effect of water more carefully," says Rim. Other areas to pursue include studying in more detail how penetration enhancers, like oleic acid, enhance diffusion.

"This is a good example of how you can take calculations from a molecular level up into a more macro level and apply them to a practical drug delivery problem," says **Gerald Kasting**, **PhD**, of the University of Cincinnati. He cautions that the model uses assumptions about the stratum corneum that have been challenged—including the idea that the corneocytes are impermeable to drugs, and that the lipid bilayers contain enough cross-connections that they can be treated as isotropic. Despite those criticisms, he says, "I think this is a wave of the future. As we remove some of the limiting assumptions made in the analysis, some very useful results are going to emerge."

—By Beth Skwarecki □

> "We believe that by separating the contributions from different levels of scale, the model can provide better insight about the barrier properties of the skin," says Jee Rim.

Simbios:

Bringing Biomedical Simulation to Your Fingertips

By Kristin Sainani, PhD, and Katharine Miller

Simbios began with a simple idea: that physics-based simulation of biological structures at all scales could benefit from a unified tool-building effort. >

At the same time, the thinking went, the task of assembling universal tools had to be driven by real research needs—the so-called Driving Biological Projects (DBPs). The center's initial DBPs therefore involved multiple scales including molecular dynamics (protein-folding, RNA structure prediction, myosin dynamics), neuromuscular simulation, and cardiovascular simulation.

Five years ago, Simbios started from scratch trying to meet its ambitious goals. And it has built remarkable momentum in that time. In its first five years, this National Center for Biomedical Computing has assembled a toolkit (SimTK), pieces of which lie within a variety of Simbios applications that address real biomedical research problems at a range of scales.

"Ultimately, the hope is that Simbios' toolkit [SimTK] will provide the underlying—and mostly invisible—computational foundation for a whole range of simulation tools developed both here at Simbios and elsewhere," says **Russ Altman**, **MD**, **PhD**, and co-PI of Simbios.

Getting Simbios' ideas and tools accepted and used in the wider community is, of course, a major challenge. But by producing high quality, useful tools, making them free and available and providing training and assistance, the center's tools are finding a welcoming audience.

In this feature, you'll read about how three of the center's DBPs—neuromuscular, protein-folding, and cardiovascular—have introduced state-of-the-art software tools that are already directly contributing to high-impact biomedical research. You'll also learn about the wider impact that Simbios is having, based on interviews of thirteen people who are not part of Simbios but who are actively using the center's tools to advance their own research goals.

Scott Delp, PhD, professor of bioengineering at Stanford University, is co-PI of Simbios and leads the OpenSim development team.

Simulating human movement is a challenging and complex task, and yet it is essential for understanding how the nervous system coordinates normal movement and how to improve treatments for movement disorders, like cerebral palsy. When Simbios was created, there was no common, open-source platform available for the neuromuscular simulation community to use. In the last two years, Simbios has introduced the OpenSim environment for this purpose. OpenSim is built on the Simbios core software toolkit, called SimTK, including the Simbody multibody dynamics code that Simbios created and released. The OpenSim development team has provided workshops for new users at international conferences. Users provide feedback, develop new features, and perform a wide variety of scientific studies.
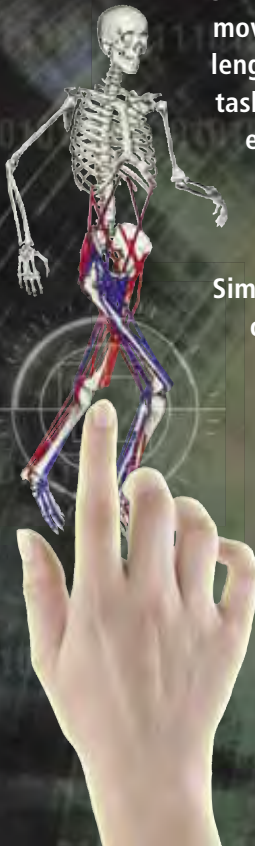
OpenSim:
Neuromuscular Simulation At Your Fingertips

When muscles of the arms or legs contract, they tug on tendons and move the underlying bones in ways that can be predicted using the laws of physics. At Simbios, researchers seek to understand the

In the two years since its release, OpenSim has been widely adopted, with more than 3000 users around the world, Delp says. "The field of biomechanics is not that large, and so to have that large number of users is a great success."

biomechanics of these movements and how modifying them might help treat people with movement disorders such as cerebral palsy, athletic injury or post-stroke limitations.

To achieve these goals, Simbios' co-PI **Scott Delp, PhD**, professor of bio-engineering at Stanford University, led the effort to create OpenSim, a freely available tool for creating detailed three-dimensional dynamic simula-tions of human movement.

THE TOOL: OPENSIM

OpenSim[1] was released to the research community in 2007 (see: *Biomedical Computation Review*, Fall 2007, p.32), and is now widely used among biomechanics researchers worldwide. OpenSim speeds up the time it takes to do neuromuscular sim-ulations by two to three orders of mag-nitude, making it possible to run sim-ulations of many subjects performing a task and to do statistical analyses of these simulations. With the upcoming release of OpenSim 2.0 (in October 2009), OpenSim promises to become even more beneficial to researchers in the field, Delp says, offering the capac-ity to handle constraints (needed for upper-body simulations), simulate
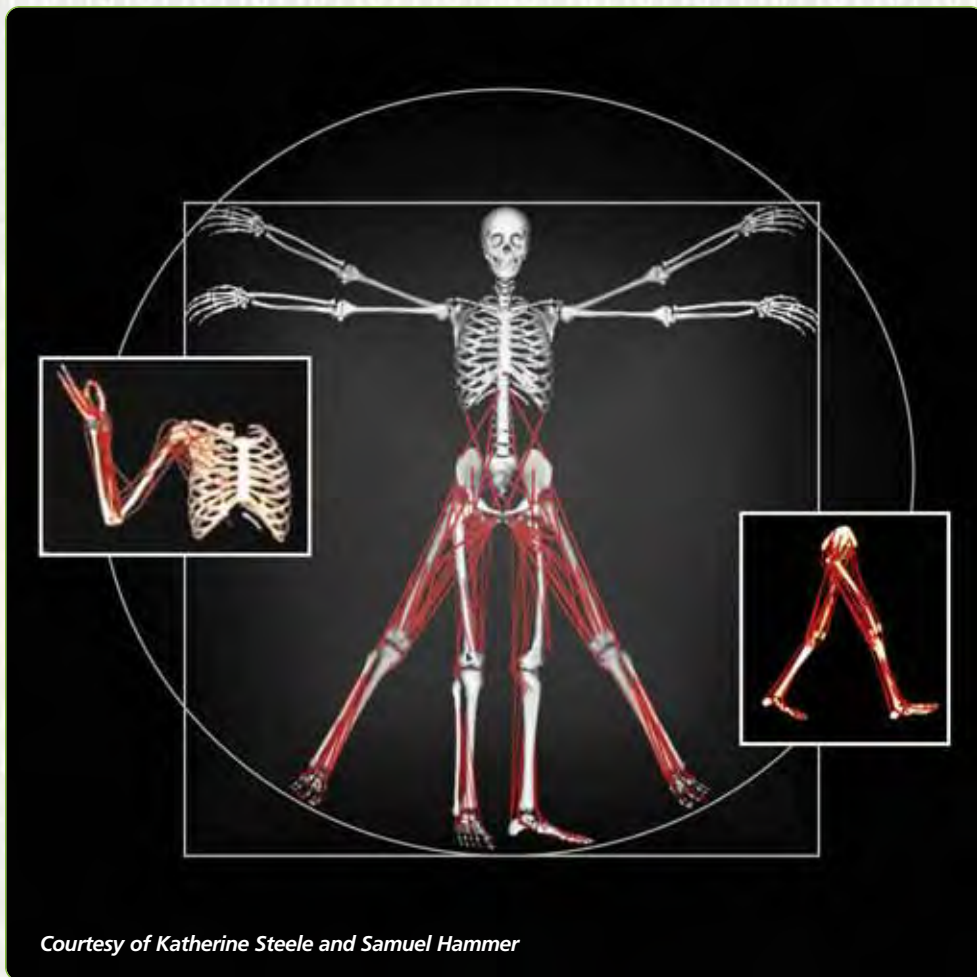
Courtesy of Katherine Steele and Samuel Hammer

Jill Higginson at the University of Delaware uses OpenSim to study stroke.

Jill Higginson, PhD, assistant professor of mechanical engineering at the University of Delaware, had one of the original subcontracts to the Simbios grant—a pilot project to explore simulating post-stroke hemiparetic (i.e., partially paralyzed) gait using OpenSim. The project showed that OpenSim could be used to study pathological gait and that it shortens the process of creating such simulations to an hour or so, as opposed to days or weeks. "That was huge," Higginson says.

In 2008, Higginson received a five-year collaborating R01 to take her work on simulating post-stroke hemiparetic gait further. The project adds a static MRI component to look at muscle atrophy on the para-lyzed side; an activation component to assess the force generating capacity of a muscle; and electromyo-graphic (EMG) data to estimate the forces produced by a muscle. At this point, just about two years into the grant, Higginson says they've collected data from subjects for the EMG, MRI and activation compo-nents and can incorporate each piece into the simulations. "Now we're working to get these parts to function together and get it streamlined into OpenSim."

Higginson is "anxiously awaiting OpenSim 2.0," which will have certain features she desires such as a ground contact model and a cus-tomized cost function. These features would allow her to iterate through different hypothetical treatment protocols for individual patients— such as what would happen if you strengthen a particular muscle or change a muscle pattern. "Would that let them better bend their knee or take a longer step?" Higginson asks. Right now, that kind of iteration isn't possible with OpenSim. "We put in experimental contact forces, but it doesn't allow you to predict what would happen to those forces if you change the experimental conditions. It would be really cool to be able to do that."

Even while awaiting this added flexibility, Higginson says she uses OpenSim every day. "OpenSim is a huge benefit. I couldn't do half of the things that I do without it," she says. "It lets me concentrate on the science rather than the algorithms."

contact forces between two bones (useful for the study of osteoarthritis, for example) or with the ground (essential to see the effect of modifying gait). Users will also be able to set their own optimization parameters (giving researchers more flexibility).

The foundation beneath OpenSim has also been strengthened since the initial release. In 2007, OpenSim used only bits and pieces of Simbios' simulation toolkit SimTK (which includes Simbody, Lapack, Simmatrix, and Simmath). Now that the toolkit is fully developed, OpenSim is built entirely on SimTK, which "provides a very fast

dynamics engine to simulate contact between bodies and to do control of multibody dynamics simulations that we haven't been able to do in the past," Delp says. "Now, anytime we add a new capability to SimTK, we automatically get it in OpenSim."

The SimTK capabilities can be accessed via OpenSim's intuitive graphical user interface, allowing users to easily create, animate, and analyze simulations.

In the two years since its release, OpenSim has been widely adopted, with more than 3000 users around the world, Delp says. "The field of biome-

Simbios broadened University of Virginia's Silvia Blemker's horizons; and OpenSim is helping her understand hamstring injuries in sprinters.

Silvia Blemker, **PhD**, has deep roots in Simbios. As a Stanford graduate student, she watched 15 faculty members put together the grant application for this National Center for Biomedical Computing ("a great experience for me at the time," she says). And as a Stanford post-doc, "Simbios broadened my horizons," she says, by articulating the commonalities between simulating proteins, cells, and whole body systems. Through Simbios, she realized that she didn't need to limit herself to studying the tissue or whole body scales, but could dive into a deeper level of detail—myosin dynamics, for example. "I now think a lot about how proteins interact and give rise to the contractile behavior of muscle, which is definitely affecting the way I'm going forward in my research," she says.

Now an assistant professor of biomechanical engineering at the University of Virginia, Blemker's connections to Simbios remain strong. She and **Darryl Thelan**, **PhD**, at the University of Wisconsin, are co-PIs on a collaborative R01 grant with Simbios which seeks to understand and eventually help prevent hamstring injuries that occur during running or sprinting. The researchers are attempting to couple two different computational strategies—3-D models of muscle (done using a program called FEBio) and dynamic simulations of running (done using OpenSim). "That achievement would have many applications beyond muscle injury," Blemker says, including for the Simbios neuromuscular DBP's work on muscle contractions in cerebral palsy and Duchenne muscular dystrophy.

"OpenSim is pretty pervasive now in the biomechanics field," Blemker says. "And Simtk.org is as well. It's allowing us to share our work and making it easier for others to know about our work."

With OpenSim, it's possible to generate many simulations in a few weeks. So, rather than making inferences from just one example, investigators can now do statistics on a sample of simulations.

Katherine Holzbaur of Wake Forest University Medical School simulates the biomechanics of the upper limb.

Katherine Holzbaur, **PhD**, assistant professor of biomedical engineering at Wake Forest University Medical School, combines computation and experiments to understand how therapeutic interventions can help people with limited upper limb movements, such as the elderly or those with specific neuromuscular disorders. For example, in her postdoctoral work with **Wendy Murray**, **PhD**, now at Northwestern University, they looked at the effect of tendon transfer surgery on hand function in people with spinal cord injuries.

For the computational side of her work, Holzbaur now uses OpenSim exclusively. Initially, the program did not include key functionality for upper limb simulation. It couldn't handle the kinds of complex constraints one finds in the shoulder, for example, where the clavicle and scapula (bones connected to the shoulder's ball and socket joint) slide past one another, constraining each other's movements. So Holzbaur worked closely with the Simbios staff when this new functionality was added to OpenSim, testing it in an already validated model of the upper limb. This will be to be useful to other groups as well, she says. "It's a really important function. Researchers of lower limb function may choose to take advantage of it to have more flexibility in how they describe joint kinematics."

Using OpenSim, Holzbaur's team recently generated simulations of upper limb reaching and other movements. The work was presented at the American Society of Biomechanics meeting in August 2009. "We're still in the development phase but getting to where we can really reliably generate simulations and validate them," says Holzbaur.

Because models developed in OpenSim can be made freely available, she says, "If we develop our model and publish it, other people can use and access it in a straightforward way without a lot of duplication of effort. People can build off the previous work of other researchers. That makes a big difference."

University of Florida's B.J. Fregly hopes to use OpenSim to simulate the knee.

B.J. Fregly, **PhD**, associate professor of mechanical and aerospace engineering and of biomedical engineering at the University of Florida, uses computational biomechanics to simulate treatments for knee osteoarthritis and wear performance of knee replacements. He is looking forward to the release of OpenSim 2.0 (expected out in October), which he says "will be the first OpenSim release to meet the specific research needs I have."

OpenSim 2.0 will let users, for the first time, create user-defined loads in musculoskeletal models. For Fregly, this means he can define contact loads between bones such as the femur and the tibia, which interact with each other within the knee. Contact forces play a critical role in osteoarthritis and wear, Fregly says. "Contact is essential to what we do." OpenSim 2.0 will also be more flexible, allowing users to define their own optimization performance criteria, "which will expand the utility of OpenSim," he says.

Fregly has attended several OpenSim training sessions, and in October, he plans to bring his contact models to Stanford and tie them into an OpenSim musculoskeletal model of the knee. "OpenSim is a big timesaver if you want to make a complicated musculoskeletal model," he says. "So I'm looking forward to using it for that purpose." He also plans to use OpenSim for teaching his undergraduate biomechanics students. "They can put it on their PCs, build realistic musculoskeletal models, and perform realistic movement simulations quickly and easily. I think OpenSim will be a great tool to help them learn." And, as these students go on to graduate school or industry and take this experience with them, he predicts that "OpenSim will become a standard for musculoskeletal modeling."

"Simbios is making commercial grade software as part of an NIH project, and I think that's awesome," Fregly concludes. "People are going to use it. It's just a question of time before it gets them over the bar for a particular need. Hopefully for me, version 2.0 will get me over the bar."

chanics is not that large, and so to have that large number of users is a great success."

THE BIOLOGY

Using OpenSim, Delp's team is studying normal and pathological movement, and designing individual treatment plans for those with movement disorders.

During walking, muscles must hold the body up against gravity as well as propel the body forward. To quantify the contributions of different muscles to these two tasks at a range of walking speeds, **May Liu**, **PhD**, a former student in Delp's lab created and analyzed 32 walking simulations (eight healthy subjects at four walking speeds). The results were published in the *Journal of Biomechanics* in November 2008.

Among other insights, the work revealed that when people walk slowly (as many with impairments do), they walk with their knees straight, similar to passive dynamic robots. The aligned skeleton holds the body's weight up, so the muscles don't have to do this work. But, at faster speeds, people flex their knees to give more shock absorption, and the muscles have to work harder to provide support.

The breakthrough in this paper was the number of different simulations run, Delp says. "In the past, generating just a single three-dimensional dynamic simulation of gait was a heroic effort," he says. "Only one would be developed over a five-year period and then people would analyze that simulation of that one subject." But with OpenSim, it's possible to generate many simulations in a few weeks. So, rather than making inferences from just one example, investigators can now do statistics on a sample of simulations. The 32 simulations from this paper are publicly available (https://simtk.org/home/mspeedwalksims), so other investigators can easily download and analyze them.

Using OpenSim, Delp's team has also run extensive simulations to study how cerebral palsy affects movement. Patients with cerebral palsy often walk in a crouched position, with an exaggerated bending of the knee. Over time, this condition progressively worsens, leading to joint degeneration and walking difficulties. To study the effect of crouched gait on the muscles

of the knees and hip, Delp's team built simulations based on data from 316 subjects. The work, published last year in the *Journal of Biomechanics*, found that when people walk with a crouched gait, their muscles have to work harder to fight gravity, because the bent skeleton does not provide this support. At the same time, the muscles lose their ability to straighten the knee and hip. "So it's kind of a double whammy: the effect of gravity is enhanced and the capacity of your

> "Students in biomechanics labs around the world are using OpenSim as part of their courses," Delp says.

muscles to counteract that is diminished," Delp says. These effects lead to a worsening of crouch, which leads to more loading on the muscles and further reductions in the muscles' abilities—and this explains why crouched gait is a downward spiral.

Fortunately, it is possible to correct crouch gait—by strengthening specific muscles or through surgery—and reverse it before it progresses. Delp's team is using simulations of individual patients to predict which corrections will fix crouch in particular patients, and the simulation results are already providing improved understanding of

crouch gait and guidelines for treatment. In one project, Delp's group used OpenSim to simulate dynamics of 127 individual subjects. "That was stunning," Delp says, "and would not have been possible without advanced simulation software."

THE COMMUNITY

Traditionally, the biomechanics simulation community has been fragmented due to the lack of common tools. OpenSim provides a shared platform that can bring the community together. With OpenSim, labs can reproduce, check, and build on each others' work, Delp says. Users are encouraged to post the simulations they create with OpenSim on the Simtk.org website for anybody to access, and neuromuscular models to use within OpenSim are also shared through Simtk.org. Researchers can also improve OpenSim itself by contributing plug-ins that add functionality to the software.

To bring OpenSim to the community, Simbios has held eight workshops and tutorials on OpenSim at Stanford and at conferences around the world; and a ninth workshop is scheduled for October, at which time, OpenSim 2.0—with much greater functionality— will be released. "We have trained over 100 biomechanics researchers during intensive, multi-day workshops at Stanford and have introduced at least three times this many to the software at various conferences," Delp says.

OpenSim has already been downloaded over 5500 times. Researchers are currently using OpenSim to study a wide variety of biomedical problems including joint forces in individuals who are susceptible to osteoarthritis; movement dynamics in individuals with stroke; the movements of the upper limbs in individuals who have a spinal cord injury; and athletic performance and injuries. Besides its use in research, it's also being widely used in education. "Students in biomechanics labs around the world are using OpenSim as part of their courses," Delp says.

Delp concludes: "What's become apparent is that because we've built OpenSim to be a general biomechanical simulation package, it's not just serving the specific DBP at Stanford but is now enabling a much broader research community."

Vijay Pande, PhD, associate professor of chemistry at Stanford University, is lead researcher on the Simbios protein-folding DBP.
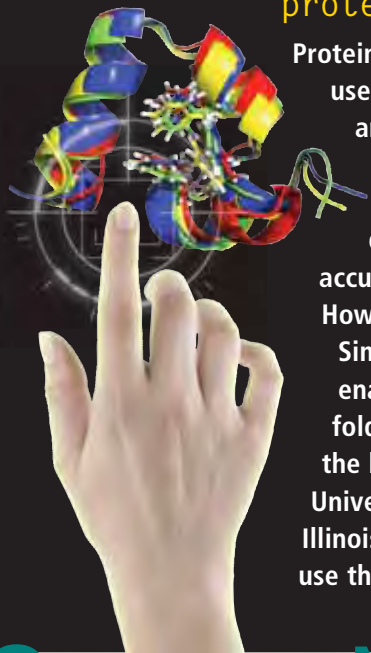
Protein-folding involves molecular dynamics which uses the rules of physics to simulate the motion and dynamics of proteins. The physics of these molecules requires that small time steps (on the order of femtoseconds, 10^{-15} seconds) be taken during a simulation in order to guarantee the accuracy of the computed forces and accelerations. However, interesting biology occurs on the timescale of seconds to hours. Simbios has therefore implemented Open Molecular Modeling (OpenMM) to enable the use of graphical processing units (GPUs) to provide 100- to 1000-fold speedups. Simbios is now aggressively disseminating this technology to the biomedical research community, with collaborators at Notre Dame, the University of Pittsburgh, California Institute of Technology, the University of Illinois, the University of Stockholm, in Sweden and elsewhere to help test and use this technology.

OpenMM:
GPU Acceleration At Your Fingertips

The great limiting factor in simulating molecular dynamics is the speed of computation, says **Vijay Pande, PhD**, associate professor of chemistry at Stanford University and lead researcher on the Simbios protein-folding DBP. Since the launch of Simbios, Pande's team has made major advances in the speed of molecular dynamics simulations and is bringing these speed-ups to the community. The team is also at the forefront of a paradigm shift in molecular dynamics, moving away from simulating single long trajectories of molecular events to instead simulating thousands of shorter trajectories. "It's really about going from something that's anecdotal—kind of cool, but anecdotal—to something that's real, statistical, quantitative, and meaningful," Pande says.

THE TOOLS

Last year Simbios released OpenMM[2] (Open Molecular Mechanics), an open-source, extensible library that brings GPU capabilities to molecular dynamics software (see: *Biomedical Computation Review*, Summer 2008, p.28). GPUs, or graphical processing units, are up to two orders of magnitude faster than traditional CPUs (central processing units), but more difficult to program. OpenMM developers (including developers from the GPU manufacturers) did the dirty work—programming core molecular dynamics algorithms across multiple GPU platforms—and have made these algorithms available to everybody. In early tests, OpenMM achieved speedups of as much as 700-fold compared with traditional CPU implementations.

Since GPUs are available on desktops and laptops, this means that scientists can now do fast molecular dynamics using a computer they have under their desk, Pande says. The preview version of OpenMM was released in September

In early tests, OpenMM achieved speedups of as much as 700-fold compared with traditional CPU implementations.

2008 and the software has already been downloaded 2000 times.

Since the initial release, "we've been rapidly adding functionality," Pande says. One of the most exciting additions, released in May, is the ability to support explicit solvent models (where water is modeled as individual atoms

rather than as a continuous fluid). "For many people molecular simulation means explicit solvent," Pande says. So this advance gives the software "much, much broader applicability," he says. "It's turning out to be the fastest single GPU code to do explicit solvent molecular dynamics," he adds.

Pande's team is also working on new, faster algorithms for molecular simulations that could "be a game-changer," he says. With collaborator **Jesús Izaguirre, PhD**, associate professor of computer science and engineering at Notre Dame University, they have developed a new method (normal mode multiple time stepping Langevin dynamics, or NML) that may speed up molecular dynamics simulations in implicit solvent an additional 10- to 50-fold, Pande says. Then it may be possible to do millisecond simulations in just a few months, he says. Considering that current approaches achieve microsecond simulations in that time frame, "the millisecond timescale is something that most people don't even talk about. But now it would be something where just about anybody could do it on a desktop. So that gets particularly exciting," he says.

Developers have also added a graphical user interface, called OpenMM Zephyr (released in January 2009 and already downloaded 500 times), which allows experimentalists to take advantage of OpenMM, even if they've never done molecular simulations before, Pande says. It is designed to be user-friendly and fail-safe, constraining the naïve user to follow good standard practices for molecular dynamics. "They know that whatever they're doing with this code,

Getting Together. Alzheimer's disease develops when amyloid beta aggregates into neurotoxic clumps (oligomers) and then into plaques. Pande's team simulated oligomer formation at the all-atom level and analyzed the results using MSMBuilder. This figure shows three different ways that four-chain oligomers (tetramers) can form: (A) as a dimer bound to two monomers; (B) as a dimer plus a dimer; and (C) as a trimer plus a monomer. The simulations predicted that the trimer is the most stable species for aggregates of up to four chains. Reprinted with permission from Kelley, N., et al., Simulating oligomerization at experimental concentrations and long timescales: A Markov state model approach, J Chem Phys 129:214707 (2008), American Institute of Physics.
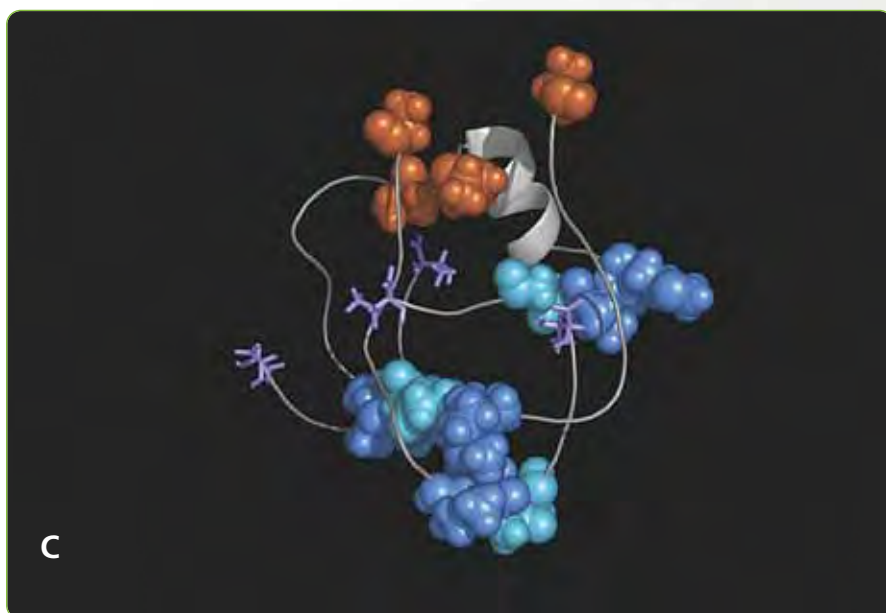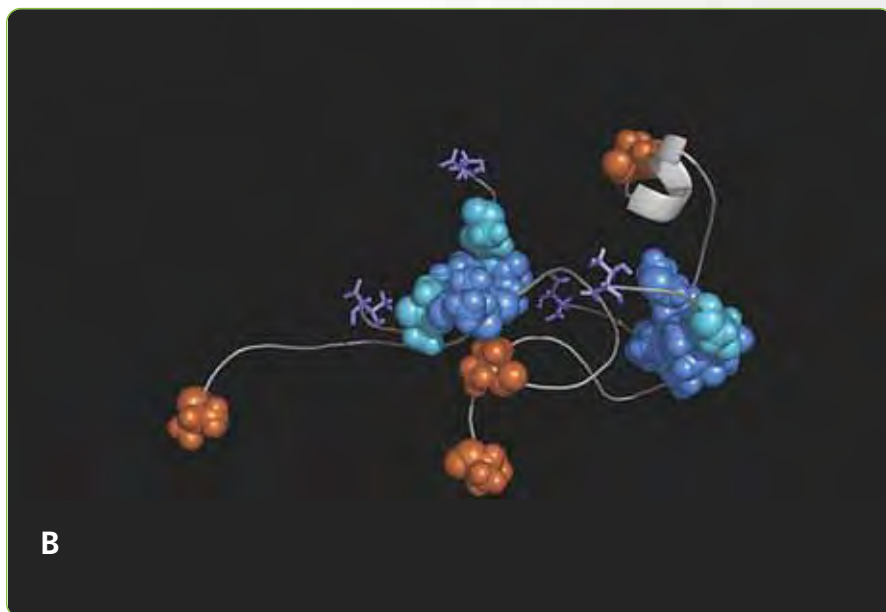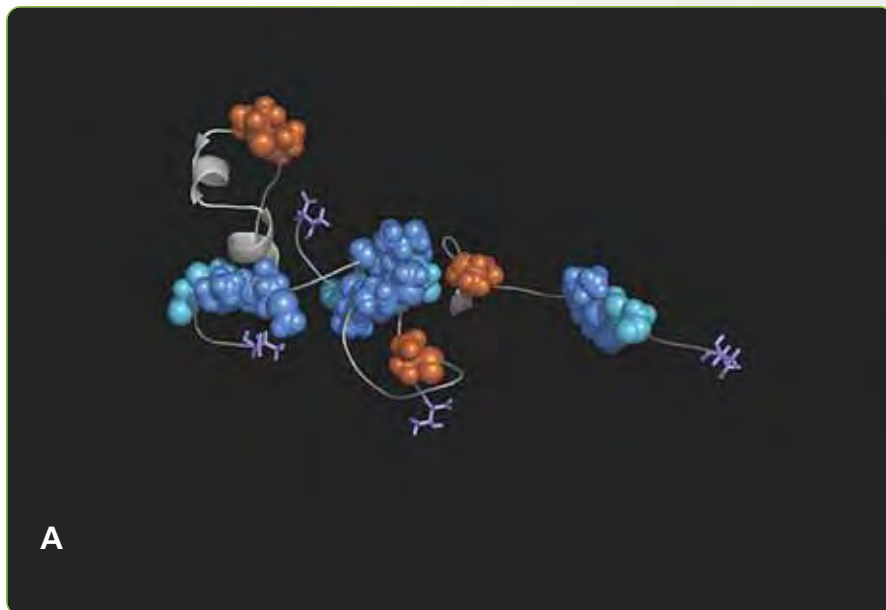
it's been vetted by us," Pande says.

Finally, MSMBuilder[3], a tool that's complementary to OpenMM, was released in April 2009. OpenMM can simulate thousands of trajectories of a molecular event, such as protein folding, sampled across many time points. MSMBuilder then takes these trajectories and analyzes them statistically to build a Markov state model, which consists of a series of states and the transitions between them (see the "Seeing Science" column on the back cover of this issue for more details). Protein folding is not just about the final structure, but about the intermediate structures it assumes along the way. "Characterizing these intermediates is very much what we're interested in," Pande says. MSMBuilder builds a step-by-step model of the folding process—identifying the intermediate states and calculating the transition probabilities between them. The software relies on a novel algorithm, developed with collaborators at Stanford, that clusters intermediate structures based both on geometric similarity as well as kinetic proximity—that is, how energetically easy it is to interconvert between two structures.

> OpenMM can bring GPU capabilities to existing molecular dynamics packages.
>
> "I think it could speed up everybody's work pretty dramatically," Pande says.

THE BIOLOGY

Though Simbios' molecular dynamics tools are relatively new, they are already being used to do groundbreaking science at Stanford and beyond. "The protein-folding and RNA-folding DBPs both have been taking advantage of OpenMM, and I think that has been able to accelerate them pretty dramatically," Pande says.

Using OpenMM, Pande's team simulated the folding of a protein fragment called WW domain, in the most extensive simulation of this folding event to date. Their results, published in the April issue of *Biophysical Journal*, reveal that protein folding is a surprisingly heterogeneous process.

WW domains are the smallest natural beta-sheet structures (35 to 40 amino acids), and have been extensively studied as a model for beta-sheet folding. Previous simulations focused on generating a single long-folding trajectory—including a landmark 10 microsecond trajectory. But Pande's group instead ran thousands of simulations of shorter trajectories (some as long as several microseconds) totaling more than 2.73 milliseconds (2730

Erik Lindahl of Stockholm University uses OpenMM to speed up molecular simulations of membrane proteins and takes inspiration from Simbios' professional approach to software development as he continues developing and maintaining GROMACS.

Simbios provides two key benefits to **Erik Lindahl, PhD**, associate professor of computational structural biology at Stockholm University. First, he says, he values the speed provided by OpenMM. Lindahl's research involves simulating membrane proteins and, he says, "having support for GPUs—that will be critical to speed up simulation throughput 10-fold." He predicts that OpenMM will be widely used as a plug-in when it can do everything that people do in modern molecular simulation. "We are 95% there, but we still have a short way to go before the GPU is a drop-in replacement for the CPU."

Second, says Lindahl, who develops and maintains GROMACS, a widely used molecular dynamics software package, "Simbios' influence is making us take a more professional approach to software development." Lindahl has watched as the Simbios team efficiently developed code with dozens of people involved while still keeping the code maintainable. "What I didn't really appreciate before was the importance of having professional software developers," he says. Simbios' approach to developing modern software libraries on a large scale, getting them to interact with one another, documenting the code thoroughly, and getting everything working consistently—"That's very different from a typical scientific project where the scientist develops the code they need for the next paper," Lindahl says.

Simbios has also been pushing the field by releasing its tools under completely open licenses, Lindahl says. "So they are not just making things available 'in theory and if you ask really hard,' but they are putting it out there with thorough documentation and saying 'Please use this.'"

microseconds). Surprisingly, they discovered that folding proceeded through many disparate pathways and resulted in two distinct end products—the expected 3-stranded beta structure as well as a "misthreaded" version.

"This heterogeneity could not have been revealed by any single molecular dynamics trajectory," Pande says. Capturing this heterogeneity is critical for making quantitative comparisons in experiments as well as for "answering the question that people have been asking me for three decades: how do proteins fold?" Pande says. Citing a single trajectory is like answering the question of "how do people fly from San Francisco to New York?" by citing a single United flight, he says. "This is one anecdotal answer, but the truth is much more complicated and much more diverse."

MSMBuilder is also having substantial scientific impact, for instance, in understanding and potentially treating Alzheimer's disease. Alzheimer's disease develops when a certain protein (amyloid beta) misfolds, enabling it to aggregate into neurotoxic clumps (oligomers) and then into plaques. Oligomer aggregation takes place on the seconds timescale, so all-atom simulations of this process were previously impossible. But Pande's team accomplished such simulations using MSMBuilder; the results were published online in the *Journal of Chemical Physics* in December 2008.

From simulations of more than 6000 short trajectories, totaling 100 microseconds, Pande's team detailed the structures of 14 intermediate states that may occur during aggregation, as well as the transition probabilities between them. They also predicted that a particular mutation (a glycine-to-proline substitution at position 37) would arrest amyloid beta in an early, non-toxic state;

they are currently testing this prediction experimentally. "Such mutants could be useful therapeutically," Pande says. "People have come up with different schemes for inhibiting amyloid beta toxicity, but this would be a very novel one and we hope that it could go pretty far."

THE COMMUNITY

A major goal of OpenMM (and related software) is to unite the molecular dynamics community, Pande says. OpenMM can bring GPU capabilities to existing molecular dynamics packages. It has already been hooked into GROMACS and ProtoMol, and Pande hopes to extend this to others, such as CHARMM, Amber, and NAMD. "I think it could speed up everybody's work pretty dramatically," he says.

OpenMM provides a shared interface that can foster collaboration. Because it is completely open, developers can con-

---

Notre Dame's Jesus Izaguirre collaborates with Simbios to increase the time scales of protein folding simulations with OpenMM.

Why team up with Simbios? Because "they are working on exciting problems and have good people," he says.

Jesus Izaguirre, **PhD**, associate professor of computer science and engineering at the University of Notre Dame in Indiana, says the best thing about Simbios is the people. From mid-2006 to early 2008, Izaguirre spent a sabbatical at Stanford working largely with Vijay Pande. He has since returned to Stanford to attend several workshops, give Simbios seminars, and discuss further collaborations with Simbios PI Russ Altman and Simbios post-doc Sam Flores. "I want to work with them because they are working on exciting problems and have good people who are both competent and nice." He also likes the fact that, at Simbios, science drives technological developments. "That is key," he says.

In 2008-09 Simbios provided seed funding for a collaboration between Izaguirre and Pande. Their goal: to speed up protein folding simulations from the microsecond to the millisecond time scale, where the most interesting motions for biology happen—for example, protein folding and conformational changes. To do this, Izaguirre's and Pande's groups have developed a methodology called NML (Normal Mode Langevin). It is now being imported into OpenMM to reap the benefits of GPU acceleration that OpenMM provides.

Izaguirre is particularly interested in how mutations affect protein folding. For example, certain mutations in the WW domain—a protein domain implicated in Alzheimer's and cancer—speed up folding 1000-fold. To figure out why, Izaguirre's team is using OpenMM and NML, as well as Pande's Folding@Home program—which allows researchers to get statistics about the probabilities of different folding outcomes by generating thousands of simulations. "For WW mutants, we may be talking about 30,000 folding simulations," he says. "Only Folding@Home lets us do that. So it's a good combination: using Folding@Home plus OpenMM plus our methodology that gives the additional speedups."

Izaguirre says Simbios is a terrific initiative. "And I think it's unique because even though I'm working mostly on the molecular level, Simbios includes problems that are of biomedical relevance at a range of scales from molecular to tissue and organ level. And I think it's a good thing to have that kind of cross-fertilization. So I do hope it's around for a very, very long time!"

---

tribute back to OpenMM; and the entire community then benefits from the added functionalities. "If enough people did that, it would make everybody more than the sum of the parts, which would be pretty exciting," says Pande.

To help bring the tools to the community, Pande's group held free workshops over several days in February and June of this year, attended by 60 scientists and developers from 23 institutions from both academia and industry. The sessions included hands-on tutorials on the use of OpenMM, MSMBuilder, and OpenMM Zephyr, as well as time for developers to work with the OpenMM team on integrating OpenMM into their existing molecular dynamics codes. Additional workshops will be held next year.

"There are a lot of interesting things to think about for the next five years," Pande concludes. "But even now, we've gotten to a point where we've got very powerful tools. And I'd love to see what people will do with them."

Pande's lab simulated the folding trajectories of the WW domain. Here, each of four folding trajectories proceeds by a distinct mechanism. Reprinted from Biophysical Journal 96(8), Ensign DL, and Pande VS, The Fip35 WW Domain Folds with Structural and Mechanistic Heterogeneity in Molecular Dynamics Simulations, pp L53–L55 (2009) *with permission from Elsevier.*
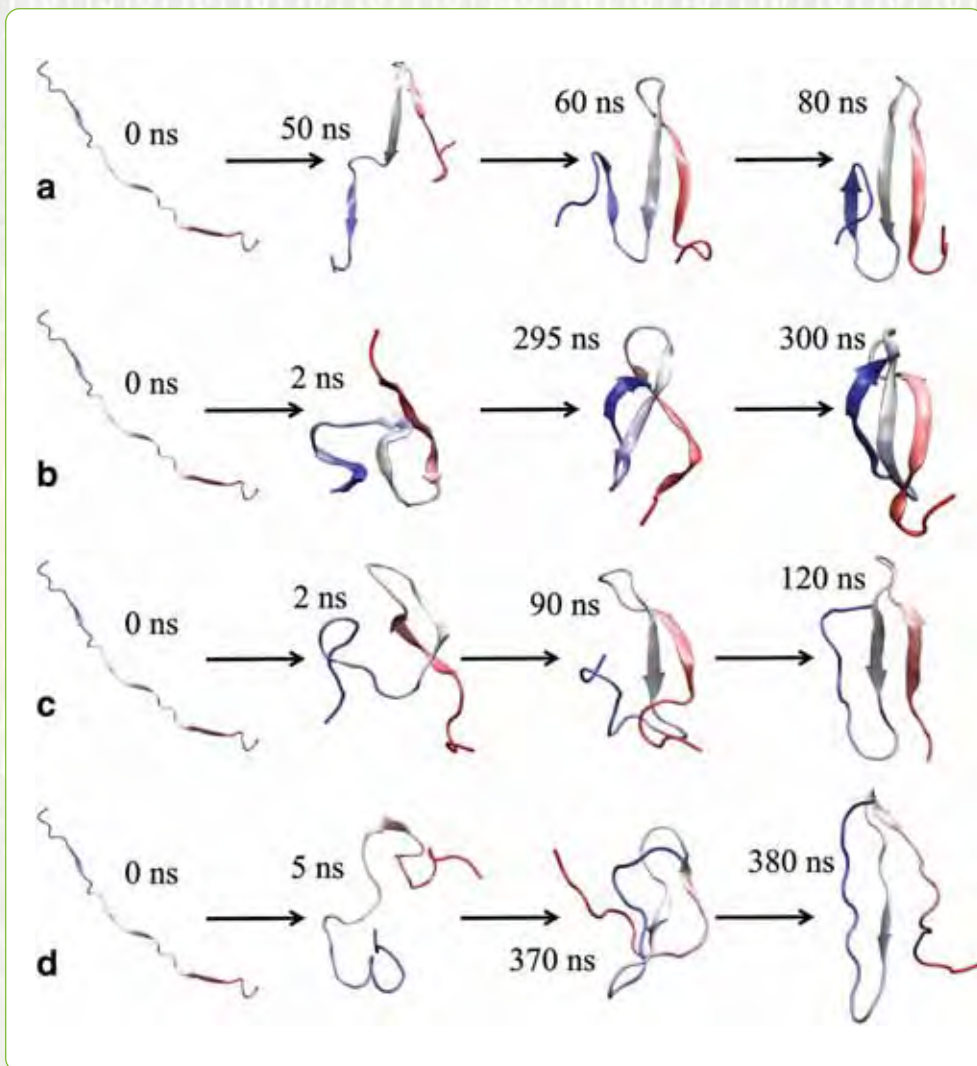
Kim Branson of Vertex Pharmaceuticals uses OpenMM as the GPU accelerator for Yank, a program for quickly estimating molecular binding affinities that he's building with collaborators from Pande's lab.

Kim Branson, **PhD**, a research scientist in the modeling and simulation group at Vertex Pharmaceuticals, together with **John Chodera**, **PhD**, a former Pande group member now at the University of California, Berkeley, are developing a program called Yank along with several other alumni of Pande's lab. "It's become a focal point for everyone to contribute," Branson says.

Yank, which is built on OpenMM, estimates the energy involved when a chemical entity binds to a protein target, a critical calculation in drug design. "OpenMM is absolutely critical for Yank," he says, because it speeds up the calculations. Industrial chemists don't want to wait two weeks for a computational answer that will help design the next experiment, Branson says. "So we need to run these calculations in a day, which OpenMM allows us to do."

Branson wants to make Yank user-friendly, so that chemists, including those at Vertex, can use it. And, like OpenMM, it will be open source and available on Simtk.org, which Branson uses as a centralized repository for all sorts of tools. "It's actually in everyone's best interest to have an open set of tools," he says. The more people use a tool, the more it evolves as people contribute, provide feedback, and gain experience with what works and doesn't work, he says. "Everyone wins."

Charles A. Taylor, PhD, associate professor of bioengineering at Stanford University, is PI for the cardiovascular dynamics project within Simbios.
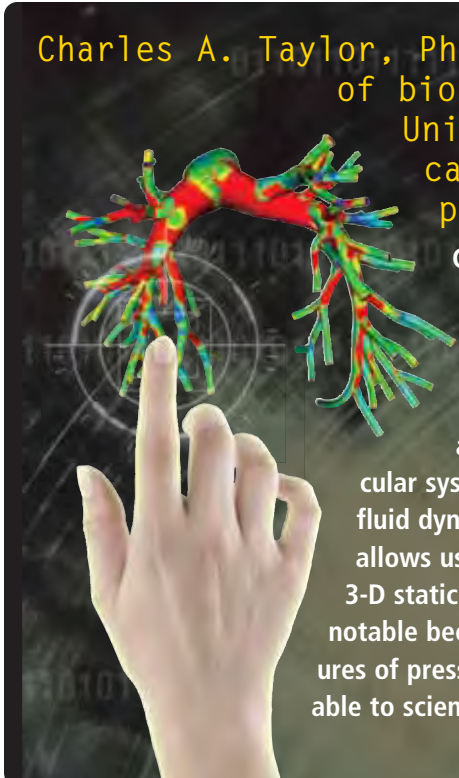
Cardiovascular disease is a primary source of morbidity and mortality in the United States and the world. Fundamental to understanding the mechanisms of this disease, and for formulating strategies for treatment, is the ability to simulate both the normal cardiovascular system as well as the system in disease. SimVascular is an environment for fluid dynamic simulations that provides advanced modeling capabilities and allows users to move from clinical images (CT and MRI) of the vascular system to 3-D static models and then on to dynamic models of circulation. These models are notable because they are among the first to produce quantitatively accurate measures of pressure and flow for the human vascular system. SimVascular is now available to scientists and clinicians around the world.

SimVascular:
Cardiovascular Simulation At Your Fingertips

When the heart or vascular system becomes diseased, medical imaging provides a window into the problem—but the information is limited. Simbios' cardiovascular simulation program SimVascular picks up where imaging leaves off. From imaging data, SimVascular reconstructs an accurate three-dimensional model of blood flow through the arteries of individual patients; this model can be used to predict outcomes and virtually test interventions.

SimVascular is giving doctors new insights into arteriosclerosis, aneurysms, heart disease, and congenital heart defects; guiding the development of better medical devices; and being used to plan the treatments and surgeries of individual patients, says **Charles Taylor, PhD**, associate professor of bioengineering at Stanford University and PI for the cardiovascular dynamics project within Simbios.

THE TOOL: SIMVASCULAR

SimVascular is a simulation program that combines several state-of-the-art commercial components with open-source code. It is fully parallelized and highly scalable, Taylor says. "The software has been run on computers with tens of thousands of processors and on a laptop," he says.

SimVascular distinguishes itself from other cardiovascular simulation programs because it accurately models blood pressure as well as blood flow velocity. "That's very unique—we can model not just the velocity of blood going through the arteries, but the actual realistic pressure wave propagating through the vascular system," Taylor says. "Blood pressure is obviously quite important, because how much the arteries deform is directly related to blood pressure."

Other cardiovascular simulation programs have ignored the deformability of arteries, treating them as rigid tubes; but SimVascular treats blood vessels realistically—as flexible, dynamic objects that interact with the blood. It is also unique in that it models the microcirculation—the tiny blood vessels at the

"We can model not just the velocity of blood going through the arteries, but the actual realistic pressure it's propagating through the vascular system," Charles Taylor says.

ends of arteries that cannot be seen through imaging but play a critical role in blood pressure, distribution, and flow. "In order to be able to predict the outcome of an intervention, you have to have a realistic model of the microcirculation," Taylor says.

Since the beta version of SimVascular was released in 2007 (see: *Biomedical Computation Review*, Spring 2007, p. 25), the code has been "tested and improved," he says. An updated version was released in summer of 2008 that yields more stable solutions for problems that previously caused the program to crash. "That basically enabled us to solve a much larger number of problems than we could solve before," Taylor says.

He and others are working on several new capabilities that may come online with future releases of SimVascular. For example, the "fluid-solid-growth" model connects SimVascular with software that models the growth and adaptation of blood vessels over time (developed by **Jay Humphrey, PhD**, professor of biomedical engineering at Texas A&M University). Currently, SimVascular simulates at most 50 cardiac cycles

(about 50 seconds). The fluid-solid-growth model can simulate events over longer time periods—such as the growth of an aneurysm over years, Taylor says.

Another major advance that Taylor's group is working on—which he is calling "image-based fluid-solid interaction"—enables SimVascular to build models from four-dimensional imaging data, such as time-resolved CT or MRI. Unlike three-dimensional imaging data, which is averaged over the cardiac cycle, four-dimensional data gives the instant-by-instant changes in geometry and would yield more realistic simulations, Taylor says.

Finally, the commercial components of SimVascular have posed some difficulties for users. "It's a challenge to compile right now because of all of those parts and issues of proprietary software," says **Alison Marsden, PhD**, an assistant professor of mechanical and aerospace engineering at the University of California, San
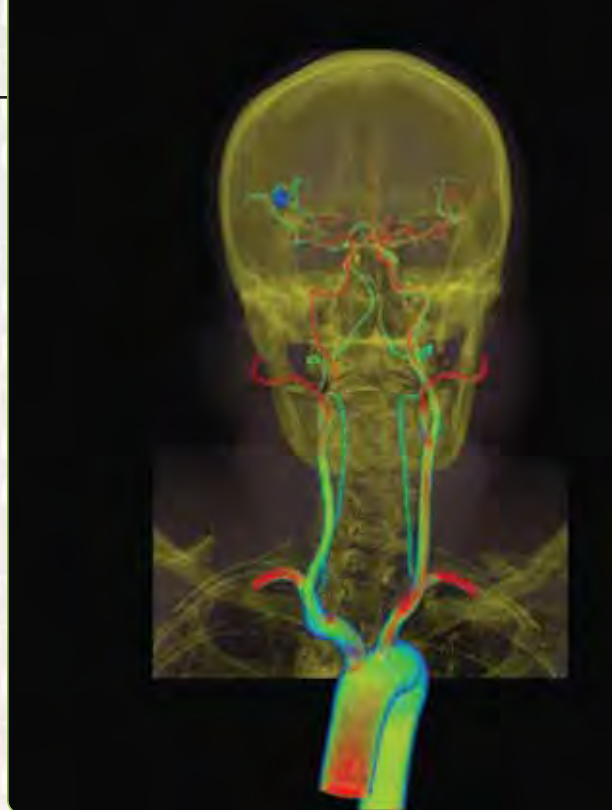
Virtual Vessels. A patient-specific model from SimVascular is superimposed on the computed tomography (CT) image data for the same patient. This patient has a cerebral aneurysm (shown in blue on the left side of this picture). Reprinted with permission from: Taylor CA and Figueroa CA. Patient-Specific Modeling of Cardiovascular Mechanics. **Annu Rev Biomed Eng** *2009; 11:109-134.*

Jay Humphrey at Texas A&M collaborates with Simbios on a fluid/solid/growth model of the cardiovascular system.

A new model of arteries that simultaneously simulates fluid, solid, and growth mechanics could eventually help prevent or treat the rupture of aneurysms. "Our goal is to predict natural disease progression and ultimately how a particular lesion or artery will respond to a clinical intervention that changes the mechanical loading," says **Jay Humphrey**, **PhD**, professor of biomedical engineering at Texas A&M University.

To create the model, Humphrey obtained a collaborating R01 grant from NIH to work with the Stanford team of **Charles Taylor**, **PhD**, and **Chris Zarins**, **MD**, that developed SimVascular. "We can't just look at the fluid mechanics (which is what has been done traditionally), the solid mechanics (done by some), or the growth mechanics (done by few), we have to look at these factors together," Humphrey says. SimVascular takes care of the fluid part. **Gerhard Holzapfel**, **PhD**, at Graz University of Technology and **David Vorp**, **PhD**, at the University of Pittsburgh, handle the solid part. Humphrey and his team work on the growth part.

SimVascular takes information from medical images of patients' aneurysms and translates it into model inputs including system geometry, local pressures, and local flows. Then Humphrey's code predicts how the vessel will respond to these pressures and flows over days or weeks. In turn, SimVascular predicts changes in fluid dynamics based on that growth. "And we'd do that iteratively until you predict that this vessel is likely to rupture or that it may stabilize and just need to be monitored for an extended period," Humphrey says.

The fluid/solid part of the model is nearly finished, and the collaborators are working on implementing the growth part. Eventually, the entire package will be widely distributed just as SimVascular has been, Humphrey says.

Diego. Users also have to pay license fees for the commercial components or search for suitable open source alternatives, which aren't readily available. So, Mardsen and others are hoping to develop open-source alternatives and contribute them back to the code.

THE BIOLOGY

Among other applications, Taylor's team is using SimVascular to study interventions for abdominal aortic aneurysm, a potentially deadly condition in which the abdominal aorta weakens, bulges, and might eventually rupture.

The condition is usually treated with a stent graft: doctors insert a fabric-covered metal tube into the diseased vessel, and blood flow is diverted through the tube, removing pressure from the aneurysm. Unfortunately, some stent grafts eventually dislodge and move over time, letting blood back into the aneurysm and putting the patient at risk for rupture.

Doctors have always assumed that the grafts move parallel with the blood vessel—flowing downstream. But in two papers in the June 2009 issue of the *Journal of Endovascular Therapy*, Alberto Figueroa (a Research Associate in the Bioengineering Department at

Stanford), Christopher Zarins (a Professor of Surgery at Stanford) and Taylor demonstrated that the forces acting on stent grafts are in fact pushing the grafts sideways (perpendicular to blood flow) in the majority of cases. These simulation results obtained with Simvascular help to explain the observed lateral movement of stent grafts recently described by Zarins.

Using SimVascular, Taylor's team built three-dimensional models of the abdominal aorta based on imaging data from individual patients with aneurysms. Then they placed virtual stent grafts in

these models and calculated the magnitude and direction of the forces acting on the grafts. Surprisingly, the majority of the forces were acting perpendicular, rather than parallel, to the graft. The more curved the artery, the bigger the sideways force—and "most aortic aneurysms are very tortuous," Taylor says. The finding has huge consequences for device design, Taylor says. "So that's had a lot of repercussions in the medical device industry." He predicts that the two papers will be highly cited.

Taylor's team is also involved in a rare pairing of computational simula-

> SimVascular is playing
> an integral part in a large,
> ongoing trial to test the potential
> for exercise to slow the growth
> of abdominal aortic aneurysms.

University of California, San Diego's Alison Marsden uses SimVascular to do patient-specific modeling of blood flow for surgical applications.

Alison Marsden, **PhD**, an assistant professor of mechanical and aerospace engineering at the University of California, San Diego, has used SimVascular since she was a post-doc at Stanford. "These tools let us go directly from medical image data to simulation results," she says. "SimVascular has capabilities that are not available in commercial packages—things specifically tailored for cardiovascular applications."

Using SimVascular, Marsden and her colleagues designed a new "Y-graft" modification of the Fontan surgery, which is done to treat children who lack one of the ventricles of the heart. After modeling the geometry of the traditional surgical correction from a specific patient's surgery, they used SimVascular to test alternative approaches and came up with the Y-graft. "It appears to be a better technique," she says. She expects it to be put into use in an actual surgery within the next six months.

Marsden's group also recently linked SimVascular to an optimization algorithm her team created. The tool tests a series of different potential surgical designs in an automated way. "We set constraints and allowable bounds on the geometry by talking to surgeons and then the computer, given those constraints, will decide which designs to try." Eventually, this tool may become open source and available as part of SimVascular.

tion with a clinical trial. SimVascular is playing an integral part in a large, ongoing trial to test the potential for exercise to slow the growth of abdominal aortic aneurysms. The clinical results from the three-year trial are not due out until next year, but the simulation results are already yielding insight into the hemodynamic effects of exercise.

Using imaging data from patients in the trial, Taylor's team simulated the effect of exercise on aneurysms with a variety of different geometries—and found that they all derived benefits. "Light exercise was sufficient to eliminate areas of chronic blood flow stagnation, which are hypothesized to lead to the progression of aneurysms," Taylor says. A paper describing the results is under review.

THE COMMUNITY

SimVascular is impacting the cardiovascular research community in a variety of different spheres. To teach the community about SimVascular, Stanford has hosted two short courses (in 2007 and 2008), which were attended by participants from both academia and industry. Taylor has also co-founded Cardiovascular Simulation, Inc., a company that will use SimVascular in-house. The company will partner with pharmaceutical companies, medical device companies, and doctors to help bring the capabilities of SimVascular—such as planning better surgeries and designing better devices—outside of academia.

Taylor and Humphrey are also leading the vascular part of the Physiome project, which is an international effort to provide an integrated framework for modeling human physiology at the cellular, tissue, organ, and total-body levels. SimVascular will play a vital part in this effort. "The Physiome is a global project, with global outreach. That's been a major focus," Taylor says.

Jeanette Schmidt, PhD, is the executive director of Simbios.

Simbios:
Packaging Research Tools for Your Fingertips

Simbios does first-rate science, says **Jeanette Schmidt**, **PhD**, executive director of Simbios. But what distinguishes Simbios from other research projects at large research universities is the dissemination effort. "Research universities don't often do such a good job of disseminating software tools," she says. "But at Simbios, we spend a lot of time packaging these things up so that other people can actually use them."

In its first five years, in addition to releasing OpenSim, OpenMM and SimVascular, as described above, Simbios released tools (and produced significant research) related to two other major efforts—the RNA Folding DBP and the Myosin Dynamics DBP.

In providing these tools, Simbios went the extra mile to make them user-friendly. It also fulfilled its aim to provide simulation tools across a range of scales, Schmidt says, with a shared underlying

> "I really think we have built the right tools," says Jeanette Schmidt.

toolkit (SimTK) used at both the neuromuscular and the molecular scales.

With this foundation in place, Schmidt says Simbios is well-primed for

References

1. Friedrichs S, Eastman P, Vaidyanathan V, Houston M, LeGrand S, Beberg AL, Ensign DL, Bruns CM, Pande VS. Accelerating Molecular Dynamic Simulation on Graphics Processing Units *J Comp Chem* 2009;30:864-872.

2. Bowman GR, Huang X, Pande VS. Using generalized ensemble simulations and Markov state models to identify conformational states. *Methods* (article in press, 2009).

3. Delp SL, Anderson FC, Arnold AS, Loan P, Habib A, John CT, Guendelman E, Thelen DG. OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement. *IEEE Transactions on Biomedical Engineering* 2007; 54: 1940-1950.

the next five years. "I really think we have built the right tools," she says. "But there's still significant refinement and additional functionality needed in order to make them applicable to more areas."

The plan, then, is to emphasize more sophisticated use of the tools Simbios has already built—to exploit what has already been done, improve it and expand it into even more research areas, Schmidt says. "The opportunity is really there for the fruits to be reaped." □

Through RNABuilder, Simbios brings computational modeling to Rick Russell's lab at the University of Texas.

Rick Russell, PhD, associate professor of chemistry and biochemistry at the University of Texas, Austin, studies RNA structure formation in the laboratory. Recognizing that his experimental data were not sufficient to give a structural picture of RNA folding, he decided to turn to modeling for help. Last winter, he asked graduate student **Yaqi Wan** to scour the Internet for RNA modeling tools. She tried for a few weeks but found nothing that was user-friendly enough for a graduate student without training. Fortunately, in the spring, **Sam Flores**, PhD, a post-doc in Simbios, visited Russell's lab to demonstrate Simbios' RNABuilder software (see *Biomedical Computation Review*, Summer 2009, p.32), which can quickly model possible RNA structures based on limited experimental information. "We immediately saw this as the opportunity we were looking for," Russell says.

Since then, Wan has learned how to use RNABuilder and is helping Flores add functionality to the software—allowing it to model RNAs based on homology to related RNAs. "That hadn't been done before with RNABuilder and really hadn't been done much for RNA at all," says Russell. Russell and Wan then successfully used this new capability on a test case, generating a model of a fairly complex, 200-nucleotide RNA without reference to the known crystal structure. They plan to use RNABuilder to model a related RNA for which the structure of extensive regions is unknown.

"This is basically a first for my lab, to work with computational models," Russell says. "It's quite helpful to have such knowledgeable and interactive collaborators."

Li Niu of the University of Albany works with Simbios to understand an unusual RNA.

Li Niu, PhD, associate professor of chemistry at the University of Albany, SUNY found an interesting RNA while selecting aptamers against glutamate ion channel receptors from a very large RNA library. "We only stumbled on this interesting problem which now has a life of its own," he says. This RNA—which can inhibit glutamate ion channel receptors—was confounding: the same sequence can fold into three different stable, functional structures. Two of the forms, M1 and M2, are generated only by enzymatic transcription and must act together to inhibit the receptor. Yet they cannot convert from one to the other under any existing denaturing conditions. "Once they are made by the enzymes, they are what they are. There's no way to convert them." Niu says. "That is unprecedented. The structures are not conformations, because they cannot change. RNA is a whole lot smarter than we thought."

That's where Simbios comes in. Can physics-based simulation explain how a single RNA sequence can produce three stable RNA structures? A few months ago, **Magda Jonikas**, a graduate student in Russ Altman's lab, decided to have a look using NAST, the Nucleic Acid Simulation Tool (see: *Biomedical Computation Review*, Spring 2009, p. 4) she developed with others on the Simbios team. NAST attempts to automatically predict possible 3-D RNA structures from the primary RNA sequence coupled with experimental evidence and known constraints based on the secondary structures Niu's lab produced. "And she generated some interesting structures," Niu says. "Hopefully, we'll be able to collect more chemical and enzymatic probing data to provide better constraints for 3-D modeling. I'm eager to see what they can generate out of this."

With NAST in the public domain, it will become possible for experimental researchers like Niu to generate possible RNA structures on their own—and these structures may also provide clues to guide further experiments. "I would be happier if NAST could make a complete prediction of a 3-D structure from a primary sequence, and then we could experimentally verify these structures," Niu says. "The field is not there yet, but that's the goal."

Columbia's Jung-Chi Liao seeks pathways within proteins using AlloPathFinder, a Simbios tool he co-developed while at Stanford.

As a Simbios post-doc, **Jung-Chi Liao, PhD**, sought to understand how a conformational change in one part of the myosin protein (at the ATP binding site) triggers a change in the other part (the functional site), allowing movement. Understanding this "allosteric pathway" can help researchers understand why certain mutations alter the protein's function. So Liao developed AlloPathFinder with **Susan Tang**, a Simbios Masters student, Russ Altman, and Jeanette Schmidt, executive director of Simbios. The tool was developed for and tested on myosin proteins, Liao says, "But we aimed to apply it to other proteins."

Now an assistant professor of mechanical engineering at Columbia University, Liao is applying AlloPathFinder to another family of proteins: helicases, which unwind the DNA helix. Like myosin, they use the breakdown of ATP to do mechanical work: ATP binding to one part of the helicase triggers it to move one base pair along the DNA. Liao is trying to identify the specific amino acids that bridge ATP binding to changes at the DNA binding site. Knowing which amino acids are important will help researchers generate testable hypotheses for mutational experiments.

Aside from the Simbios tools he continues to use, Liao says he valued being part of Simbios because the team nurtured a spirit of collaboration. "Scott and Russ gave me a lot of opportunities to lead projects, talk to many people, and to generate ideas with collaboration in mind."

Alain Laederach of the Wadsworth Center counts on Simtk.org as a long term software and data repository and says Simbios' dissemination efforts will pay off.

Alain Laederach, PhD, a research scientist at the Wadsworth Center and an assistant professor in the School of Public Health at the University at Albany, was a post-doc in Russ Altman's lab when the Simbios grant was written. He left Stanford in 2008 but continues to use various Simbios tools including the Simtk.org repository (see: *Biomedical Computation Review*, Winter 2008/09, p.3-4).

"We use Simtk.org as our model for all software distribution," Laederach says. In fact, he includes Simtk.org in his grant applications. "When you write a grant to propose software development, you have to explain how—past the grant—the software will live somewhere. Simtk.org is a way to satisfy that requirement for future projects because there's an intent that it will live forever—beyond Simbios."

In several of his collaborations, Laederach uses Simbios tools—NAST or RNABuilder (which themselves are built on top of SimTK libraries)—for 3-D modeling of RNA because he's familiar with them. The spread of new tools, Laederach says, is fundamentally grassroots. "People use tools because they've used them before or they know someone who has used them."

That's why the recent unveiling of NAST and RNABuilder at the RNA Society Annual Meeting in Madison, Wisconsin was so important. It attracted more than 100 people from beyond the Simbios sphere, "putting all the people who do RNA 3-D structural modeling in one room," Laederach says. "So I think we'll see more users in the future."

Under TheHood

BY MICHAEL SHERMAN

Putting Technology In Its Place

When you step on the gas pedal, you expect acceleration (and lots of it). Stomp on the brake to come safely to a stop in the rain. Finger the power-assisted steering wheel and the car obeys. Make a serious mistake and find yourself safely ensconced in deflating airbags. There is an immense amount of technology at your service as a driver, most notable for being practically invisible as you use it to get where you want to go. Each new generation of automobile is more

need to employ the immensely complex and ever-improving computational technologies of physics-based simulation. They want to step on the computational gas pedal and investigate biomolecular structure formation, neuromuscular gait disorders, or the turbulent flow of blood through the circulatory system. They need to get there fast, reliably, in reasonable comfort, and confident that inevitable errors will result in deployment of gentle safety devices rather than

Granny is off to the store.

To get to that level of simplicity and utility takes the intervention of vehicle designers who understand both drivers' needs and the available technology. Designers don't have to understand the

---

BOX 1. TYPICAL ABSTRACTIONS AVAILABLE TO VEHICLE DESIGNERS.

- **Body, Chassis, Paint, Glass**
- **Engine, Transmission**
- **Electrical, Cooling, Safety**
- **Tires, Axles, Suspension**
- **Brakes, Steering**
- **Fuel, Exhaust, Emission Controls**
- **Security, Comfort, Entertainment**

---

internal details of every technology; instead, they work with a different set of more sophisticated abstractions, such as those listed in Box 1. Similarly, for us to best serve the varied needs of biomedical researchers, it is the developers of biomedical application software we need to reach. They are the "vehicle designers" in our field, uniquely situated to understand both the needs of researchers in a particular biomedical domain, and the appropriate uses of physics-based simulation technology to serve those researchers. Application developers need to work with meaningful abstractions of our technology rather than its ever-changing intricate details. The applications they produce can in turn provide tightly focused interfaces presenting the abstractions that make sense to the people for whom they were designed. SimTK, the Simbios Toolkit (including Simbody, OpenMM, Molmodel, OpenSim Core, and other tools), implements such a set of technology abstractions for application designers, who can then effectively employ physics-based simulation technologies without having to be specialists in them. In turn, they hide these "under

sophisticated and complex than the last, chasing the ever-improving technological state of the art. Yet as a driver all you see is better transportation.

Would that software could be like that! At the Simbios Center, like all the NCBCs, we support researchers who are going important places in biomedical research. To get there, our "drivers"

bloody career-mangling disasters.

In this column I'm going to talk about how we at Simbios approach the problem of delivering powerful, high-performance computational software to users who have better things to do than worry about what's under the physics-based simulation hood. The key to keeping technology in its place, whether for vehicles or software, is: *abstraction.* Users shouldn't interact with the technology itself but with suitable abstractions of that technology. In a car, the engine, transmission, and fuel system are reduced to a few amazingly simple concepts: ignition (on or off), gear (reverse and drive), and gas pedal (more and less). Add in similar abstractions for steering and brakes and
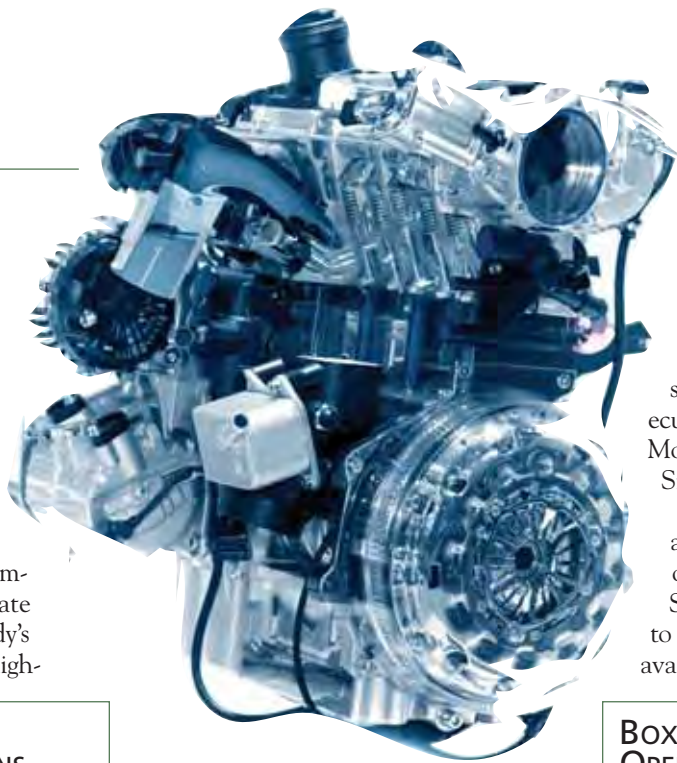

DETAILS

Michael Sherman is Chief Software Architect for the Simbios Center.

For more information about the Simbios Toolkit, go to https://simtk.org/home/ {Simbody, OpenMM,Molmodel,OpenSim}

the hood" in their applications, to best serve their own research communities.

For example, multibody dynamics is a highly specialized discipline necessary for both musculoskeletal dynamics and internal coordinate molecular dynamics. Simbody's abstractions (Box 2) enable high-

performance multibody simulation of complex systems of interconnected articulated bodies at any scale.

Graphics Processing Units (GPUs) offer an immense amount of low-cost numerical computation hardware, with year-to-year speed improvements substantially outstripping CPUs. Yet programming GPUs for scientific computation is an exacting, time-consuming, and esoteric skill, differing among GPU manufacturers and even from version to version of the same product line. OpenMM's abstractions (Box 3) allow the application developer to express molecular dynamics problems in familiar terms, while hiding the details of the particular hardware platform on which the problem is currently being solved. The result is dramatically accelerated molecular dynamics computations that take advantage of whatever GPU hardware is available, and can easily be incorporated into new applications and retrofitted into existing ones as we have done with Gromacs.

Application developers need to work with meaningful abstractions of our technology rather than its ever-changing intricate details. The applications they produce can in turn provide tightly focused interfaces presenting the abstractions that make sense to the people for whom they were designed.

Our higher-level modeling libraries for articulated skeletal and molecular systems also use abstractions to manage complexity (Box 3). OpenSim Core enables modeling and analysis of neuromuscular systems composed of abstractions like skeleton, muscle, tendon, and neurological control, using Simbody to calculate the articulated dynamics. Molmodel handles modeling of complex biopolymers as coarse grained articulated

systems with abstractions like molecule, bond, residue, and protein. Molmodel in turn employs both Simbody and OpenMM for speed.

There are excellent toolkits already available for many aspects of application development; Simbios was specifically chartered to tackle a substantial gap in the availability of suitable high-perform-

ance, high-quality, open source toolkits for physics-based simulation of biological structures. SimTK addresses the computationally intense, physics-based simulation aspects of biomedical applications. These are extremely involved technologies, but surely no worse than the thousands of moving parts and dozens of computers that Toyota already packs meekly into a Prius. By putting meaningful abstractions of our technology into the hands of domain-knowledgeable application developers both within Simbios and in the community at large, we have made significant progress toward the goal of delivering them into the hands of biomedical researchers who can now navigate the open road ahead rather than struggle to get the engine running. □

## s e e i n g   s c i e n c e

SeeingScience

BY JOY KU, PhD

Dealing With a Flood of Conformations Using MSMBuilder

In the world of molecular dynamics (MD), researchers often need to analyze and extract meaningful results from millions of conformations. Existing software for accomplishing this task focuses on geometric metrics—organizing the data only by the structural information. But **Greg Bowman**, a graduate student in biophysics at Stanford University, and his colleagues saw the need to also incorporate dynamic behavior—as a way to access important kinetic properties of the system. Their solution, MSMBuilder, automatically divides an MD trajectory data set into groups, or states, based on similarities in their kinetic behavior. It then identifies relationships between these states, resulting in a Markov state model (MSM). It is now freely available for download from http://simtk.org/home/msmbuilder.

"MSMBuilder is the first available software for kinetic clustering," says **Xuhui Huang**, **PhD**, a research associate with Simbios at Stanford University and a co-developer of MSMBuilder. Moreover, Bowman says, MSMBuilder ensures the analysis is done on a representative sampling of the conformations' true distribution. "Without a tool like MSMBuilder that helps provide good sampling, you can spend lots of limited computing resources doing the same thing over and over again and not get any more information." □
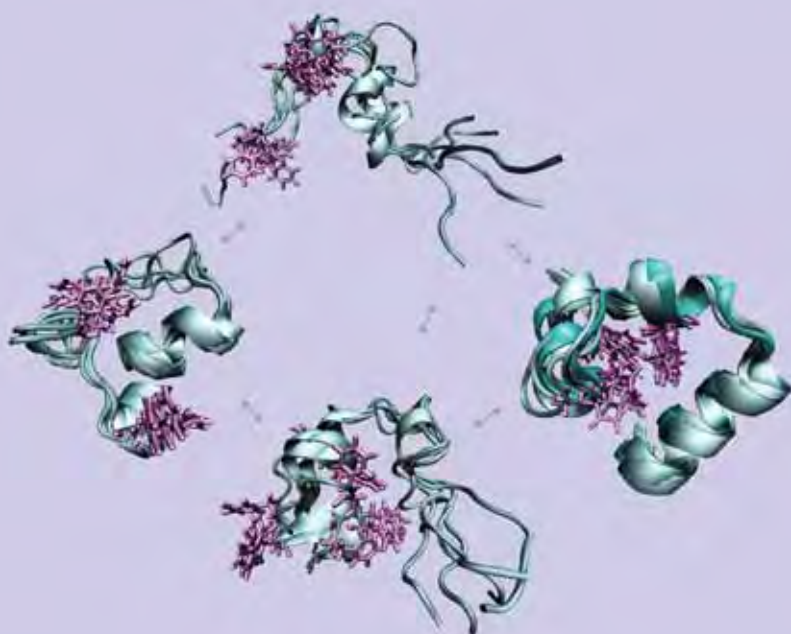
Bowman and his colleagues' first test of MSMBuilder's predictive capabilities sought to automatically identify the native state of the villin protein from an MD data set with nearly 8 million conformations. Shown here are 4 clusters of conformations automatically identified by MSMBuilder. Each cluster represents a state of the villin protein. Arrows indicate transitions between states, also identified by MSMBuilder. The group or cluster representing the native state (right-most) was accurately identified. Its members match the crystal structure (shown in darker blue and magenta) with an average root mean square deviation (RMSD) of 1.8 Angstroms. MSMBuilder is freely available for download from http://simtk.org/home/msmbuilder.