

NewsBytes

Pore Picture Construction

Like puzzles? Here's a tough one: Try figuring out the construction of a nearly 500-piece machine without blueprints or a complete picture. Biologists have now accomplished just such a feat, working out the protein-by-protein structure of an important cellular assembly called the nuclear pore complex. Their success depended on computationally combining incomplete imaging information with bits and pieces of structural data from all sorts of different experiments.

"It is as if we use many tiny lights, each of which shines from a different perspective, to illuminate every part of the whole structure," says **Andrej Sali, PhD**, a coauthor and professor of biopharmaceutical sciences and pharmaceutical chemistry at the University of California, San Francisco. "We are able to use information from many sources, even sources that haven't been traditionally used for structure determination." As described in *Nature* in November 2007, this gleaning strategy should be helpful in determining the structures of many hard-to-pin-down cellular complexes.

The nuclear pore complex (NPC) is a gatekeeper of the nucleus, a 456-protein assembly in the shape of a thick donut spanning the nuclear membrane. Scientists know the general structure of

eight spokes that make up the donut and have a roster of its proteins. But where each protein fits has been difficult to pin down.

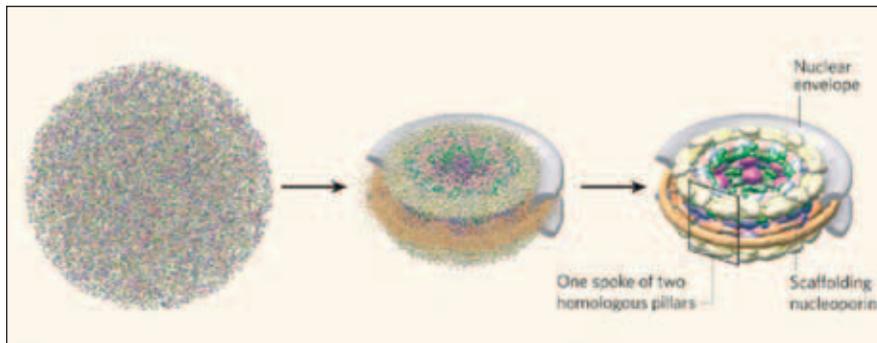
The challenge is that no one tool images the details of complexes this size. Electron microscopy reveals the overall shape and outline but not individual proteins. NMR spectroscopy and X-ray crystallography, on the other hand, show individual proteins in stark relief, yet can't be used on the whole assembly.

In collaboration with two groups of experimental biologists at Rockefeller University led by **Michael Rout, PhD**, and **Brian Chait, PhD**, Sali's team found a way of combing structural information from disparate sources and computationally putting all the bits together to create a low-resolution image of the entire complex. They included experimental data from, for example, affinity purification assays (which indicate interactions between proteins) and ultracentrifugation (which reports on protein shape). At least seven different experimental techniques were used to produce structural data.

With the data in hand, they translated each piece into a "spatial restraint"—a mathematical probability of the structure's geometry. For example, one restraint might indicate that protein A very likely interacts with protein W. Then the computer, starting with a random configuration of the proteins,

moved them step-by-step in a direction that minimized violations of restraints. This process was repeated until the group had acquired 1,000 optimized structures that each satisfied the restraints. (In total, that took 200,000

"We are able to use information from many sources, even sources that haven't been traditionally used for structure determination," says Andrej Sali.



Starting from a random mess of proteins (456 beads), experimenters ended with the structure of the nuclear pore complex. They did so by directing a computer to move the beads in any direction that minimized pre-programmed structural restraints—as if the proteins were gradually tugged towards proper placement. The final structure is an arrangement with the least cumulative "tug" or structural restraint. Courtesy of Andrej Sali. Reprinted by permission from Macmillan Publishers Ltd: Nature 450:621-622, 2007

trials run on 200 CPUs for 30 days.) The small variations between those 1,000 structures were then combined into a slightly blurry final image.

Sali was struck by the simplicity of the final structure. "If you look at electron scanning microscope pictures of the nuclear pore complex, and imagine how many proteins are involved, you think, 'This is a mess! How did this evolve?'" But once the scientists began analyzing the protein architecture, they noticed a number of symmetries and a simple three-layer architecture: one layer to hold the pore to the membrane, one layer to facilitate transport of molecules through the pore, and a final scaffolding layer to hold it all together. "It is not hard to imagine the evolution," Sali says.

Establishing the protein architecture

is also a huge step in coming to a better understanding of how the NPC facilitates controlled transport of molecules in and out of the nucleus. The group of **Klaus Schulten, PhD**, director of the theoretical biophysics group at the University of Illinois at Urbana-Champaign, is using this structure to study the transport mechanism. “The recipe that the investigators found for combining many experiments into one picture worked so consistently and so coherently across many independent trial predictions that the results must be true,” Schulten says. “Already now the relatively low resolution structure helps us to understand much better how the NPC organizes its complex function.” Chait, Rout, and Sali are now working on a high-resolution structure, with detail down to the atomic level.

—By **Louisa Dalton**

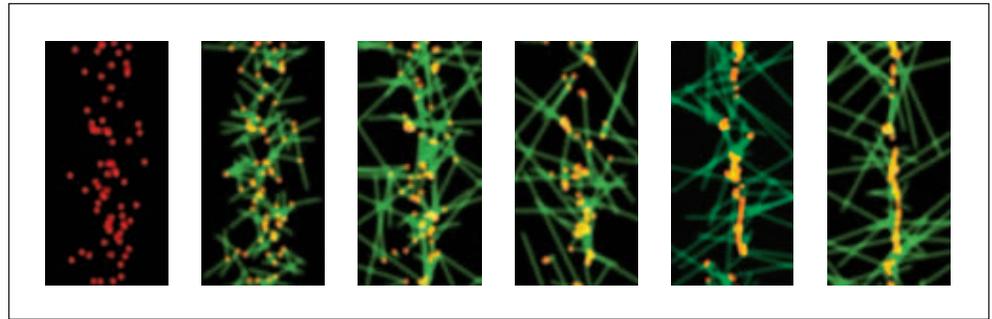
Cell Division’s Surprise Twist

During the final step of cell division, a ring of proteins pinches the cell in two—a process often likened to a purse string drawing shut. The analogy evokes a picture of thread-like proteins wrapping around the cell’s middle in an orderly fashion. But the mechanics of

Ring assembly in fission yeast follows a dynamic “search, capture, pull, and release” mechanism.

this “contractile ring”—detailed for the first time in the January 4th issue of *Science*—turn out to be far more intricate and chaotic.

“The answer—which is very exciting and surprising—is that it’s a completely random, unguided process that works perfectly,” says **Thomas D. Pollard, MD**, professor of molecular, cellular, and developmental biology at Yale



Simulations of the assembly of the contractile ring in fission yeast: Nodes (red) sprout actin filaments (green) in a random network. Myosin proteins in one node randomly encounter, capture, pull on, and then release actin filaments growing from another node. Repeated iterations of this process eventually draw the nodes together in a ring. Courtesy of Thomas Pollard.

University. His team used a combination of computer modeling and high-resolution microscopy to show that ring assembly in fission yeast follows a dynamic “search, capture, pull, and release” mechanism. The general principles are likely to be the same in higher organisms, Pollard says.

Their work follows decades of scientific exploration on the topic, he says. Experiments in the 1970s revealed that myosin and actin—the same proteins that make muscles contract—are key players. Genetic studies later identified a complete “parts list” of proteins required (about 50). Recently, scientists observed

Nodes grow actin filaments that are captured by myosins in neighboring nodes to make a continuous chain; then the myosins pull the chain closed. But, a Monte Carlo simulation of the scenario gave disappointing results—instead of forming a ring, the proteins disbanded into isolated clumps. “So we were missing something,” Pollard says.

Back in the lab, they carefully measured the movements of fluorescently tagged actin and myosin using high-resolution time-lapse microscopy in live cells. What they saw was unexpected: “The nodes move around in a completely crazy way,” Pollard says, “They go at almost 360 degrees. They don’t all head to the equator at all. They start and stop.”

This suggested a different model of ring assembly where the nodes form transient rather than permanent connections: nodes sprout actin filaments in random directions; these filaments encounter myosins in nearby nodes; the myosins capture, pull on, and then release the actin. Repeated iterations eventually draw the nodes together in a ring.

“You’d swear after two minutes of this 10-minute process, this thing was never going to get there. Even after five minutes, even after seven minutes, it’s a mess,” Pollard says. “But it turns out that just by this completely random process of searching, getting captured, moving intermittently, and then breaking connections, it always works.”

A simulation of this model formed a virtual ring in the same time it takes a

that the process begins with a broad band of dots—or “nodes”—appearing around the equator of the cell. Pollard’s team pinpointed the composition of these nodes; among other proteins, they contain formin, which polymerizes actin filaments, and myosin, which interacts with actin.

Their observations suggested a simple and elegant model for ring assembly:

live cell. “The gratifying thing is that not only does it make a ring, but it makes it in 10 minutes—which is actually a big constraint,” Pollard says.

“It’s fascinating work,” comments **Alex Mogilner, PhD**, professor of neurobiology, physiology and behavior and of mathematics at the University of California, Davis. “I think there will be more surprises in the future,” he says, “but they nailed the essence of what’s going on.”

—By *Kristin Sainani, PhD*

Modeling the Deformable Body

August 2007 saw a surge of new open-source software for simulating musculoskeletal movement. In addition to OpenSim 1.0 (described in the Fall 2007 issue of this magazine), FEBio arrived on the scene. While OpenSim uses rigid body mechanics—simulating the body moving essentially as a series of segments attached at joints—FEBio (Finite Elements for Biomechanics) addresses the other part of the problem. It can simulate how movement deforms and places stresses upon solid parts of the body such as muscles, tendons, ligaments,

cartilage and bone.

Created by **Jeff Weiss, PhD**, associate professor of bioengineering at the University of Utah, and his colleagues, FEBio already has 200 to 250 users. “Initially we developed FEBio for our use in-house,” says Weiss, “but we saw the potential for it to be a really popular tool in the research community and decided to make it available to everyone.”

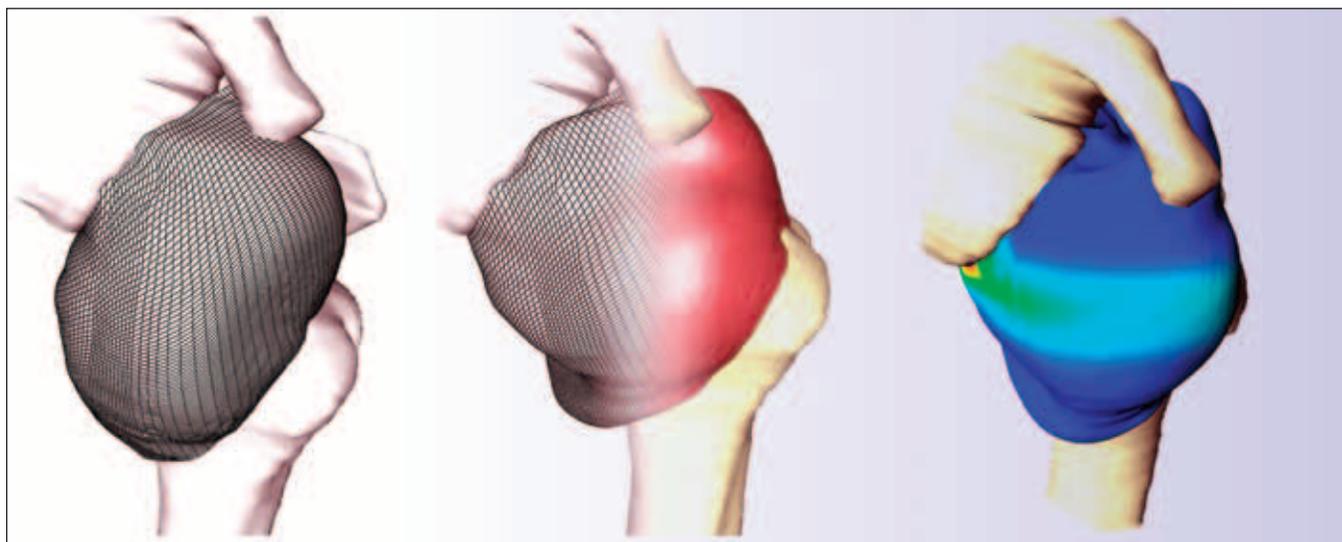
Before now, biomechanics researchers studying the solid mechanics of soft tissue have relied upon costly general-purpose finite-element programs such as Abaqus or LS-DYNA. But because these programs are proprietary, it’s hard to add new features to the code. “We saw that as a major shortcoming in our field,” says Weiss. So he and his colleagues tailored FEBio to address the kinds of problems that come up in biomechanics.

In addition to FEBio itself, Weiss and his colleagues also released programs that allow users to prepare their models in advance of using FEBio (PreView) and to analyze and visualize the results of an FEBio simulation (PostView). “That’s one of the advantages of FEBio,” says **Steve Maas**, a software developer who works with Weiss. “You can do your model creation and post-

“Initially we developed FEBio for our use in-house,” says Jeff Weiss, “but we saw the potential for it to be a really popular tool in the research community...”

processing on your own computer and use a high performance computer only for the FEBio step.”

FEBio’s users come from many different disciplines including orthopedics, ophthalmology and cardiovascular mechanics. Weiss himself has used FEBio for a variety of research projects including a study of hip stresses in



In this FEBio model of a shoulder capsule—the soft-tissue envelope that surrounds the shoulder joint—the upper arm bone (the humerus) is moved upward and then rotated around its axis. The left image shows the initial undeformed mesh, the middle image shows an intermediate state, and the rightmost image shows the stresses on the capsule in the final deformed state (blue means low stress, red means high stress). During a shoulder examination, clinicians typically move the shoulder in various ways in an attempt to determine the source of a problem. Models of this kind could eventually help clinicians better understand the results of such tests. Courtesy of Jeff Weiss and Steve Maas.

people with displasia and a study of the shoulder capsule. He and his colleagues are also continuing to add new features to FEBio.

Weiss and one of the OpenSim creators **Scott Delp, PhD**, a professor of bioengineering at Stanford University, have begun a collaboration to link the two programs to address problems that can't be handled by either program alone. Although Delp's group has combined dynamics with finite element approaches in previous work (for example in a study of knee pain), "development of advanced methods in biomechanics would be accelerated if one could use two open-source programs connected in a straightforward way. I'm looking forward to that day," Delp says.

—By *Katharine Miller*

Discovering The Bugs Within

We are crawling with bugs. It might even be better to say that we are bugs. For every human cell in our bodies there may be ten or even a hundred other cells that aren't human at all. Yet many of these microbes are entirely unknown to science. To change that, the National Institutes of Health has just begun a five-year, \$115 million Roadmap initiative called the Human Microbiome Project. It aims to find out what these bacteria, viruses, archaea and fungi are, how they function, and the ways they can keep us healthy or make us sick.

"There have been some tantalizing findings that gut flora influence things like obesity and irritable bowel disorder," says **Jane Peterson, PhD**, associate director of the Division of Extramural Research at the National Human Genome Research Institute and a program director for the project. "Ultimately, what we really want to understand is health as well as disease. What makes us healthy? Our microbes are a part of that."

But learning about these bugs has seemed like an overwhelming undertaking. Part of the problem is simply numbers: thousands of different species of microbes swarm on and in our bodies.



A human gut microbe. This bacterium, Enterococcus faecalis, which lives in the human gut, is just one type of microbe that will be studied as part of NIH's Human Microbiome Project. Courtesy: United States Department of Agriculture

The most obvious way to find out what they are is to understand their genomes. Unfortunately, sequencing these microbes is even harder than sequencing our own genome because most of the microbes have an obstreperous unwillingness to grow in isolation in a lab. They will only grow in the particular conditions of, say, our teeth, where they commune with a particular group of other microbes that create an agreeable environment.

Sequencing technology has been improving rapidly, however, bringing the task within reach now. "Metagenomic" techniques have been developed to study the genomes of many different microbes simultaneously, making it unnecessary to culture the microbes in the lab. In addition, modern sequencing machines can now produce millions of sequences in a day, compared to a few thousand in the past, and they do it less expensively.

The Human Microbiome Project has already awarded \$8.2 million to research groups around the country in 2007, and they currently have six requests for proposals out, due between February

and May.

Analyzing the data from all these far-flung groups will require the development of new computational techniques. Genomic analysis already produces such

"Ultimately, what we really want to understand is health as well as disease. What makes us healthy? Our microbes are a part of that," says Jane Peterson.

vast quantities of data that it has pushed the computational capacity to make sense of it all, and the Human Microbiome Project will produce an order of magnitude more data than that. The project aims to coordinate the results from all the different groups, producing a single, publicly-available dataset.

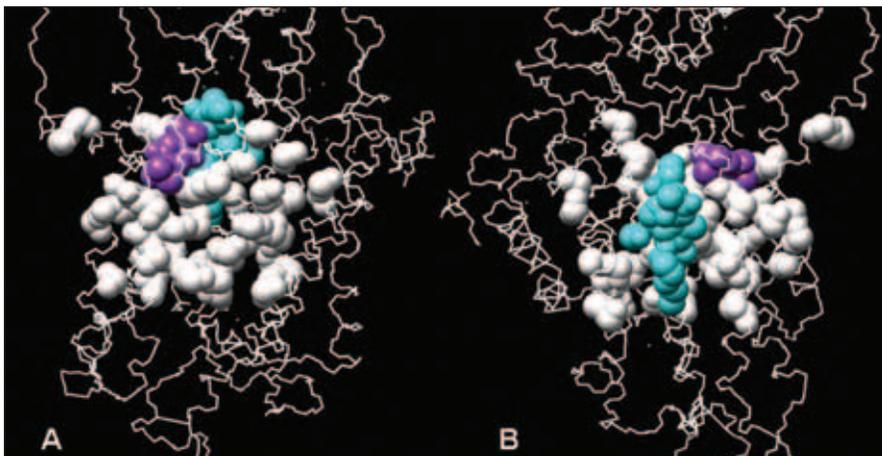
The researchers involved in the project say the most exciting part is that they simply don't know what they're going to find. "You have to expect that there will be very many ways microbes are impacting our health that we don't know and maybe can't imagine at this point," says **George Weinstock, PhD** of Washington University in St. Louis. "We're hopeful it will have an impact on the level of the human genome project."
—By **Julie J. Rehmeier**

Side Effects *in silico*

Many new drugs carry a risk that they will cause more problems than they cure. That's because a drug intended to bind one protein might also bind others. In an effort to address that problem, researchers have developed a new computational approach that can potentially predict the protein interactions that cause drug side effects. The new algorithm has already provided a possible explanation for some side effects caused by the widely-used anti-cancer drug Tamoxifen. The same approach may also help find new targets for commercially-available drugs.

Traditional drug discovery searches for possible drugs that can bind to a known receptor protein. "We're doing essentially the reverse of that," says **Philip Bourne, PhD**, professor of pharmacology at the University of California, San Diego and lead author of the work published in the November 2007 issue of *PLoS Computational Biology*. "We've already got something that binds to a receptor. The issue is that it doesn't necessarily bind only to that receptor."

To find out what else the compound is binding, Bourne and his colleagues



The algorithm created by Bourne and his colleagues identified a possible Tamoxifen binding site (white spheres) on a protein called SERCA that regulates calcium levels within muscle cells. They also found that two known inhibitors of SERCA bind to areas (shown in purple and blue) within the same zone. This suggests that a side effect of Tamoxifen could be inhibition of this protein, Courtesy of Philip Bourne.

start with a database of potential receptors—what they call the “druggable proteome.” They then test whether the compound binds to one or more secondary sites in receptors other than the primary target. Previous attempts to predict such drug-protein interactions have met with limited success. But **Lei Xie, PhD**, a member of Bourne's team, developed a novel algorithm that considers the evolutionary relationship among potential binding sites and also allows the receptor proteins to bend and move.

Combining these new parameters with an analysis of the receptors' shapes and binding characteristics yielded a powerful search tool capable of discovering off-target proteins missed by previous algorithms. Bourne's team then looked at whether the known functions of those off-target proteins could provide a logical explanation for a drug's known side-effects.

Bourne's team applied their algorithm to a family of cancer drugs that includes Tamoxifen. Known as selective estrogen receptor modulators (SERMs), this clan of drugs often causes unwanted side effects such as heart disease and ocular degeneration, both of which involve a disruption in cells' calcium balance. So Bourne's team was not surprised when their algorithm found

Tamoxifen could bind a protein that regulates calcium levels within muscle cells (Sarcoplasmic Reticulum Calcium ion channel ATPase protein (SERCA)). Specifically, the algorithm predicted that Tamoxifen inhibits SERCA's action (by binding near natural inhibitors' binding sites).

Bourne hopes the algorithm will help identify potential side effects of new compounds before they reach clinical trials, saving enormous amounts of money and time. In addition, the algorithm could help researchers design drugs with fewer side effects and find new targets for already-approved drugs. Indeed, Bourne's group has already found that existing Parkinson's disease drugs may help treat extreme drug-resistant tuberculosis.

“The potential value is huge if one could do this reliably,” says **Robert Stroud, PhD**, a professor of biophysics and biochemistry at the University of California, San Francisco. Stroud cautioned, however, that more examples of the algorithm's ability to successfully identify off-target proteins are necessary before any definite conclusions can be drawn.

—**Matthew Busse, PhD** □