CELLUCIDATE: Modeling by Community

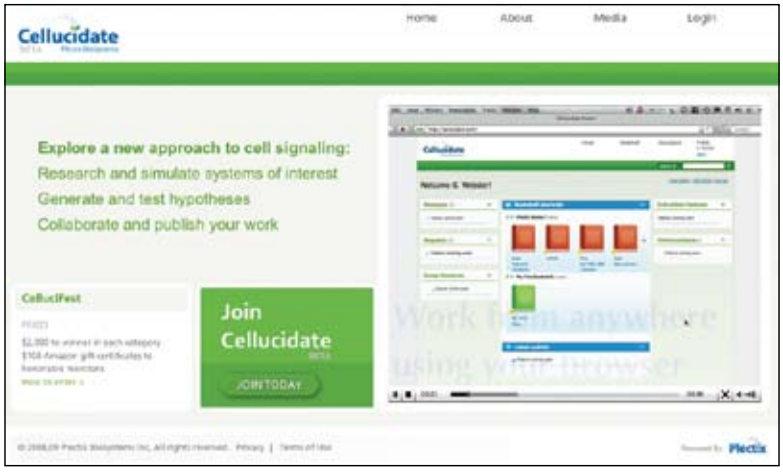
BY JOHN WHITFIELD

"CELLUCIDATE," SAYS SFI EXTERNAL PROFESSOR WALTER FONTANA, IS "FACEBOOK FOR PROTEINS." It allows biological molecules, via the people who study them, to form connections, find out what they have in common, and keep track of one another. And, he says, it can do the same for the scientists, helping them to communicate and collaborate. "It's also Facebook for researchers that deal with proteins."

More precisely, Cellucidate is a platform, accessed through the World Wide Web, for modeling and simulating networks of cell signalling. These molecular interactions between proteins and genes control and coordinate the workings within a cell and its communication with other cells to ensure every component does the right job as part of the whole body. It doesn't quite have Facebook's social networking reach yet: a beta version of cellucidate.com had a soft launch in December 2008, and as of April 2009, the site had 300 registered users. But what it lacks in numbers, the project makes up for in intellectual ambition: the website describes itself as "Gutenberg for models." Once someone has come up with a model, they can share it with the site's other users, who change it, critique it, or plug it into their model of a different part of the cell-signalling system.

PHOTO: NARISSA ESCANLAR

Walter Fontana explores ways to make models perspective-generating instruments; Background pattern: This depiction of a "hairball," by Funahashi H. Kitano and colleagues, conveys a sense of the complexity inherent in networks governing cellular information processing.



Through the Cellucidate website, researchers can work with the complexity of cells. Such a task is especially daunting because the facts keep changing (and growing), which requires that such networks be easy to edit and update.

And when new data comes along—as it does at an alarming rate, thanks to automated, high-throughput methods for analyzing protein interactions—the model can be updated. The whole business is intended to be simple enough to democratize the modeling process, allowing people with different backgrounds and different skills to talk to one another, and to help researchers focused on the workings of a single molecule see the bigger picture. "Cellucidate allows a community of researchers to look at the same facts in the same way," says Fontana. "It's capable of tracking knowledge scattered across the research community."

In many ways, this echoes Fontana's larger research goal. He has spent his entire career exploring different disciplines. His Ph.D. advisor at the University of Vienna, for example, was theoretical chemist Peter Schuster, an SFI external professor. With Schuster, Fontana trained in chemistry and molecular biology. But since his teens he had also harboured an interest in evolution, sparked by Jacob Monod's classic 1970 book *Chance and Necessity*. "It showed me that chemicals don't just change color in a beaker, or explode, or make smells," he says. "They organize information, and read and write. It was an epiphany." Working with Leo Buss of Yale University, a leader in evolutionary developmental biology, or evo-devo, Fontana pursued these evolutionary interests, looking at the links between biological information, form, and function at the molecular level, using models that were at times conceptual and at times very realistic, such as how sequences of RNA molecules relate to their structure.

In the early 1990s, Fontana spent two years at SFI as a postdoctoral fellow. He returned toward the end of that decade, giving up tenure at the University of Vienna to spend six years as a research professor at the Institute. The move allowed him to escape the problem of extracting space and money from institutions and funding agencies set up along traditional disciplinary boundaries. It's a dilemma all too familiar to researchers who not only straddle disciplines, but also bridge conceptually and empirically driven research.

PLATFORM WITH MANY USES

As well as supporting individual researchers, Cellucidate may aid pharmaceutical companies. Many diseases are the consequence of malfunctions in cell signalling. "Signalling systems are the basis of virtually all cancers," says Fontana. And such systems are the target of many potential drugs. The pharmaceutical industry is always looking for any edge in the search for drug targets or possible side effects, and modeling signalling might help them.

Subscriptions from such companies will likely form the basis of Cellucidate's paid users, though currently, registration for the platform is free. Eventually, it will have to make money for its parent company, a start-up co-founded by Fontana called Plectix BioSystems (www.

plectix.com). For paying customers, Cellucidate will include layers of privacy to allow commercial users to work in confidentiality. But Fontana envisages that most models will remain open-source projects—as will the underlying software—with a view toward maximizing benefits of the networks between proteins, models, and researchers.

But Cellucidate is more than a tool and a business. It's the embodiment of Fontana's attempts to build links between computing and cell biology. Fontana, who moved to Harvard Medical School's Department of Systems Biology in 2004, believes that computation should be what he calls the "third pillar" of systems biology besides chemistry and physics. This means going beyond using computers to crunch data, and even beyond trading metaphors between these disciplines, as with genetic algorithms or DNA computing. Instead, Fontana sees computer science as a basic science that will become increasingly important to biology, as a source of ideas, formal techniques, and explanations.

UNRAVELLING THE HAIRBALL

Even within a single cell, the systems created by signalling molecules are head-spinningly complex. There are about a dozen major systems controlling processes such as metabolism, DNA replication and repair, growth, and cell division. Each has a large number of components, and all of them overlap and interact. One of the main products of systems biology up until now has been diagrams that map the components and connections of a particular system. You can get an idea of what these look like by the fact that Fontana calls them "hairballs"—if you weren't a systems biologist, all you'd take from looking at such a diagram is that the cell is a very, very complicated place.

And yet even the hairballs don't capture the reality. They are static, like a subway map, when in fact the links in signalling networks are constantly changing, winking on and off, varying in strength, and feeding back on one another.

All in all, a cell may have more potential states than there are stars in the universe. "The number of possible objects that the hairball can in principle enable is astronomic—it could be 10^{40}," says Fontana. Even seemingly identical cells can, within seconds, drift into very different states. Understanding such processes requires new tools: "There's no way that you can represent the dynamics of that map with traditional methods."

But, he points out, the signalling processes within a cell have a lot in common with computational systems such as the Web (or the brain). They also share properties with distributed computing projects, such as traffic control, where local communication between many autonomous components gives rise to coherent behavior at the system level.

Fontana taught himself computer science, and he devised Cellucidate in collaboration with a group of specialist computer scientists in an attempt to make sense of this complexity, and let other people do the same. The system uses plain English to describe which proteins bind to what. Then, using a computer language called Kappa, it turns these descriptions into a model of cell signalling that is also a program, thus making the connection between cell signalling and computation explicit. When this program is run, it becomes a simulation that provides an idea of how this signalling system plays out at a cellular level. (You can see a demonstration video at www.cellucidate.com.) It's still a work in progress. "The software engines are really good," Fontana says. "Now the question is how to make the best interface."

GROPING TOWARD KNOWLEDGE

As one of the founding professors in Harvard Medical School's Department of Systems Biology, Fontana believes that life is getting easier for interdisciplinary researchers; although, he stresses, that doesn't mean that everyone has the same scientific worldview. Still, Fontana continues to promote modeling, especially a type that owes more to computation than physics. But getting others in systems biology on board can be an uphill struggle. "People are very sceptical about modeling. They say we can't model, because we don't know everything yet." But this, he says, "is precisely why we need to model." Rather than describing something we already understand, a tool such as Cellucidate helps in approaching understanding. "The model is becoming more of a reasoning instrument," he says. "It's a way of arguing and groping towards consensus knowledge." ◀

John Whitfield is a London-based science writer and author of In the Beat of a Heart: Life, Energy, and the Unity of Nature. *From September through December 2007, he was science writer in residence at SFI.*