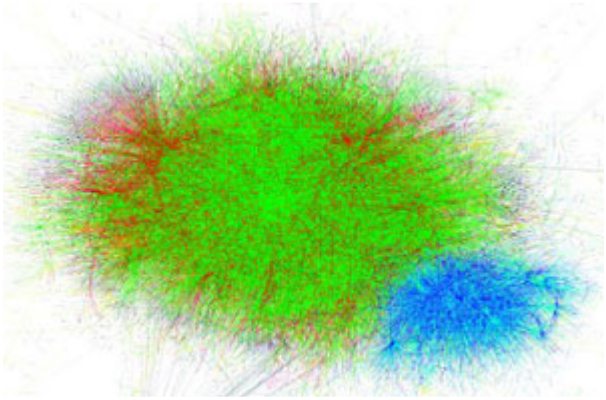
CHEMICAL & Engineering News

Cover Story

February 14, 2005
Volume 83, Number 7
pp. 47-55

SYSTEMS BIOLOGY

Measurement and modeling approaches bring a big-picture view of biology and may improve drug discovery and development



MAKING CONNECTIONS This network graph shows causal connections among 30,512 genes, 31,459 proteins, and 5,824 small molecules in Genstruct's model, which contains 136,362 causal connections that can be evaluated to explain the molecular state changes observed in large-scale systems biology experiments. The red connections represent inhibitions; green, activation; light blue, reaction; dark blue, a product; yellow, catalysis; orange, binding; and black, gene product relations.
COURTESY OF GENSTRUCT

CELIA M. HENRY, C&EN WASHINGTON

Most approaches to studying biology have been reductionist. Typically, researchers focus on one component of a biological system at a time—a gene or a protein, for example. They then take what they learn about the individual components and try to work up to the system level, whether that system is a tissue, an organ, an organ system, or the whole organism. The problem is that, in such a bottom-up approach, the data about the individual genes or proteins are often divorced from the context.

Drug discovery has been similarly reductionist, starting with how compounds interact with a drug target and then moving into in vitro models, an approach that is still removed from the ultimate context. Only in clinical trials do potential drugs finally encounter their true context–the patient. Often, it is only at that late point that previously unrecognized effectiveness and toxicity problems become apparent.

That's where systems biology comes in. Systems biology tries to bring a big-picture view–context–back to biology and especially drug discovery. Researchers combine data about

genes, proteins, and metabolites to generate a comprehensive picture of the connections between the different parts. In the context of drug discovery, the data can be used in computer models and simulations to identify the pathways involved in a particular disease–leading to new therapeutic targets–or to determine whether a drug is hitting the intended pathway.

Although some large pharmaceutical companies such as Eli Lilly & Co. and Novartis have made investments, most of the systems biology work is being done in small companies. (Because systems biology has been such an ambiguous term, some companies have shied away from that label, instead using descriptions such as computational biology or molecular fingerprinting to describe what they do.) These smaller companies then form partnerships with or provide services to the larger pharmaceutical companies.

Despite different approaches to systems biology, people generally agree that its aim is the understanding of the function of a biological system–tissue, organ, or organism–as a whole. Some people emphasize the large-scale measurements needed for systems biology, while others focus on the computer-modeling aspects. Many modeling companies rely on the literature and their pharmaceutical partners for the experimental data needed for their models. Other companies have developed measurement technologies to generate their own experimental data for modeling. Ultimately, both measurements and models are necessary to understand complex biological systems.

Such an understanding can have a tremendous impact on drug discovery and development. Initially, the effect is likely to be evident in the safety and efficacy of drugs. Later, it will probably reduce both the time and the cost of getting drugs to market.

"Right now, we spend so much money on targets and on drugs that are eventually going to fail in the clinic," explains Ellen L. Berg, chief scientific officer of BioSeek, Burlingame, Calif. "To get that failure rate down, we have to be able to predict unexpected activities earlier."

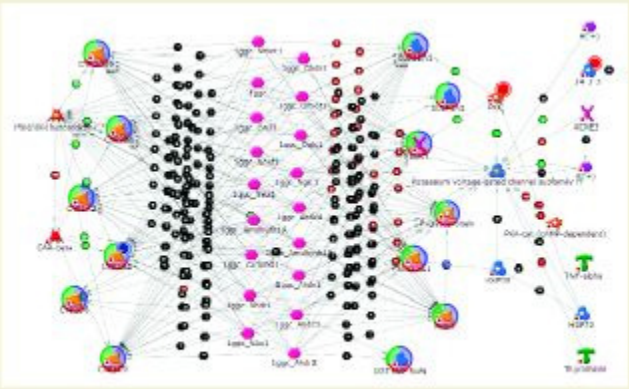
J. Michael French, chief business officer at Entelos, Foster City, Calif., believes that the key to increasing efficiency and reducing failures is decision-making. "You can equate failure rates to bad outcomes based on poorly informed decisions," he says. Use of modeling and simulation gives people the information they need to make better informed decisions, resulting in more successful clinical programs, he says.

"Big pharma is by its very nature extremely conservative," says Keith O. Elliston, president and chief executive officer of Genstruct, Cambridge, Mass. During the genomics era, technologies promised more than they could quickly deliver, and pharmaceutical companies ended up feeling burned.

Such experiences have created skepticism toward new technologies, but a "healthy amount" of skepticism may be good for systems biology, according to Colin Hill, president and CEO of Gene Network Sciences, Ithaca, N.Y. "It keeps us in check and makes us work that much harder to make the science work and to articulate our message and value proposition."

However cautious big pharma might be, the smaller companies are moving forward with systems biology and are applying it to all stages of drug discovery and development, from target discovery and validation to preclinical research and clinical trials.

One area in which systems biology is expected to have a large impact is in ADME/Tox (adsorption, distribution, metabolism, excretion, and toxicity) evaluations.



ON THE PATH **GeneGo uses its software and databases to generate maps of pathways relevant to drug discovery and development.**
COURTESY OF GENEGO

GeneGo, based in St. Joseph, Mich., has two databases specifically for such evaluations: MetaCore and MetaDrug. MetaCore contains information about human biological pathways gleaned from the experimental literature by the company's scientists. The content serves as the backbone for generating networks and maps of the connections between system components. The database itself has a novel architecture designed to handle ambiguous data and data of different degrees of quality and completeness.

MetaDrug predicts the metabolites and interactions for a particular drug structure that might be important in toxicology. The software offers the ability to make theoretical predictions of metabolites and toxicity as well as to incorporate experimental measurements of metabolites to visualize preclinical and clinical data in the context of the complete biological system.

Like much else in drug discovery, toxicology studies have been reductionist, points out Sean Ekins, GeneGo's vice president for computational biology. For example, toxicology studies have concentrated on the interactions of drugs with individual enzymes, particularly various cytochrome P450s. "We need to look at all the potential interactions a drug might have, not just with the drug-metabolizing enzymes" but also with other proteins such as nuclear receptors and transporters, Ekins says. "We have to open our eyes to all the information that can be generated from a molecule once it comes into contact with these receptors, transporters, and enzymes."

TWO PATHS
Measurement Is Giving Way To Modeling

The real differences in systems biology come in the approaches that people take, not the final goal. There are two main types of approaches: those based on measurement of the system and those based on computer modeling. Few companies use one approach to the exclusion of the other. The measurements typically consist of the various "omics" data sets that people have become familiar with. In modeling, those data are incorporated into computer representations of biological systems and pathways.

"If you look at measurements versus modeling, the key limitation with measurements is that it doesn't matter if I can measure everything if I can't model and conceptualize it," says Keith O. Elliston, president and chief executive officer at Genstruct, Cambridge, Mass. Basically, modeling serves as a way to extract meaning from the data and is now the key bottleneck.

Elliston says.

"In the beginning, people were more oriented toward experimental data collection in a high-throughput way," says Colin Hill, president and CEO of Gene Network Sciences, Ithaca, N.Y. "Some of them have come around to realizing that they need computer models to make sense of all the data they've been collecting."

For example, Target Discovery, in Palo Alto, Calif., started out as a measurement-based company but has moved more toward modeling, which the company prefers to call computational biology. At first, the company focused on measurement technology because the existing technology could not produce data of high enough quality to distinguish between competing model predictions, according to Luke V. Schneider, chief scientific officer. "Now that we've got better measurements, we need to refine our ability to build better and alternative models, or better and alternative hypotheses," he says.

Similarly, BG Medicine, formerly known as Beyond Genomics, originally concentrated on developing technology for comprehensive biological measurements. Now that those technologies are in place, the company will put "less emphasis on the technology and more emphasis on the integrated output of those technologies," says Stephen A. Martin, chief technical officer.

Although data are important, models can be generated with less information than people might think. "We don't need to measure everything in a cell to understand its critical activities," Elliston says.

Also, models need not be exact; they just have to be predictive and, according to Schneider, "predictive enough that you don't kill somebody, or you certainly don't make them worse."

TARGET DISCOVERY of Palo Alto, Calif., is another company that is using systems biology in ADME/Tox. Its MetaSIRMS platform uses stable-isotope, rather than radioisotope, labeling for protein measurements by mass spectrometry. With stable isotopes, which can be administered to people, the assays designed for in vitro animal models can be translated directly into clinical studies. "You can develop a model [in animals] and then validate whether that model holds or doesn't hold in humans," says Luke V. Schneider, chief scientific officer. Schneider estimates that this approach would lower the cost of experimental work by one order of magnitude. In addition, being able to validate the animal model directly should speed up development and might even reduce the amount of preclinical work required, he says.

"The question with every drug product is, 'Is it efficacious before it's toxic?'" Schneider points out. "That critical question really drives the whole lead optimization process. You don't want to go into clinical trials–where 80% of your cost is–until you've answered that question adequately."

At any point in drug development, a major goal of many systems biology projects is to identify biomarkers–proteins or metabolites, alone or in combination, that can report on the behavior of the system. Biomarkers can help deal with issues of efficacy and unexpected toxicity, the main causes of failures in late-stage drug development. Biomarkers, preferably ones found in easily accessed body fluids, can indicate whether a drug is working or whether it is toxic even before damage occurs.

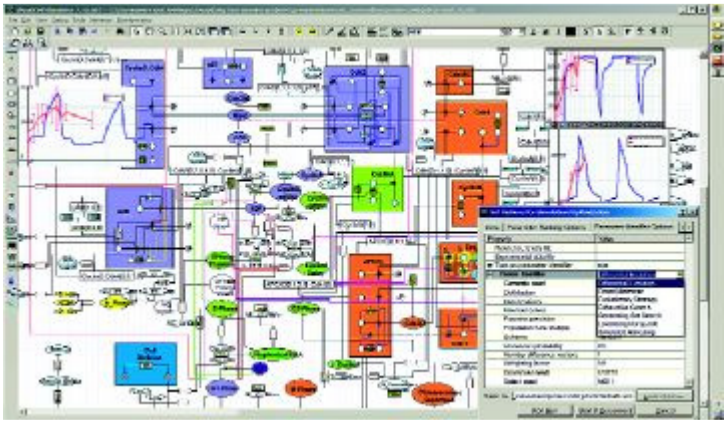
"If you have a biomarker, something readily measurable in serum or urine, that can tell you whether your compound is working, you have a great tool for giving your drug to the right patients," Genstruct's Elliston says. Even the Food & Drug Administration sees value in biomarkers. In the report "Innovation or Stagnation: Challenge and Opportunity on the Critical

Path to New Medical Products," the agency outlines what it sees as the opportunities to create an improved "tool kit" for assessing safety and efficacy with biomarkers, among other approaches.

Genstruct is one of various companies using systems biology to identify biomarkers that can be used as indicators of drug efficacy and toxicity. It finds these biomarkers through techniques that represent biological knowledge in a format that permits logical modeling, as opposed to discrete mathematical modeling. The key pieces of information are the drug's mechanism of action and mechanism of toxicity–not just which proteins the drug binds but its biological impact at a mechanistic level, Elliston explains.

Elliston also believes that the problems of efficacy and toxicity are more tractable from a modeling perspective than those related to underlying disease biology. "I've got a pretty simple system. I've got the same cells, the same tissues, the same organism. The only difference is the drug. I've got a very simple perturbation to model," he says. "All the effects that I see at a systems level–gene changes, protein changes, phosphorylation changes–all relate back to that one perturbation." In contrast, disease biology is more difficult to model. For example, cancer is caused by a number of mutations occurring all at once.

The company uses measurements of mRNAs, proteins, and metabolites in conjunction with its models to find surrogate biomarkers for efficacy and toxicity. So far, Genstruct has worked in discovery and safety assessment, but Elliston expects the firm to start its first project involving clinical data toward the end of this year.



IN SILICO In this model of the cell cycle generated by Gene Network Sciences, the model parameters were optimized for phosphoprotein and cell cycle distribution data.
COURTESY OF GENE NETWORK SCIENCES

ANOTHER COMPANY identifying biomarkers is BioSeek, which uses its BioMAP cell-based assays everywhere from target discovery to clinical trials. BioSeek focuses primarily on inflammation biology, but its assays are also useful in studying metabolic disease, cardiovascular disease, and cancer.

The assays mimic the biology of human disease by incorporating multiple cell types and multiple pathways that are turned on in combination, according to Berg. "We use these very sensitive, very powerful assays to test-drive drugs in these complex environments."

The assays use only human primary cells–ones taken directly from the body as opposed to being derived from cell lines. Cell combinations and culture conditions are carefully optimized so that the cells behave as they would in the body. In each assay, eight to 12 proteins that

have been identified as being the most informative about that particular assay are measured.

"Once we have assay conditions that we are satisfied with, we test-drive [the assay] with known therapeutics," says Ivan Plavec, senior director of technology development at BioSeek. "If our assay is correct, then it should respond to known therapeutics the same way as has been recorded in clinical trials."

The company maintains a database of the profiles of the drugs it runs through the BioMAP systems. From this database, BioSeek's scientists are able to see differences between drugs that target the same pathway, COX-2 inhibitors, for example, or steroids. Despite being designed for the same clinical indication, individual drugs of a class still differ in their overall biology. Such differences have been in the spotlight recently with the high-profile withdrawal from the market of Vioxx, Merck's COX-2 inhibitor. BioSeek's database also can be used to identify the mechanism of action of a new drug by seeing which pathways it hits and looking for similarities with other drugs. In addition, the assays can identify secondary and off-target effects, Berg says.

Icoria, based in Research Triangle Park, N.C., is interested in biomarkers that serve as a "bridge" between events happening in a target organ and an easily accessible body fluid such as serum or urine. "If you can identify proteins or biochemicals that change in association with a disease or target, you're on your way toward having a marker that will help you throughout the discovery and development cycle, from the development of assays to tracking pharmacologic activity to developing tests that might help you select patient populations," says Thomas J. Colatsky, vice president of human health research at Icoria.

Icoria uses the concept of data coherence to help simplify the vast data sets that are generated by the high-throughput measurements used in systems biology. Data coherence refers to the ability to take complex data sets and reduce them to sets that can be analyzed in a single dimension. The treatment removes bias from individual data streams–genomic, proteomic, and metabolomic–so that no data source overwhelms the others. One way such bias elimination is accomplished is through purely statistical approaches that look for sources of variance in the data. Another way, a nonstatistical approach, is to check the data to see where they belong and look for "hooks that link the data sets together," Colatsky says.

Whereas most companies approach systems biology from the bottom–starting with individual genes, proteins, and metabolites–and build their way back up to the patient, Entelos starts at the clinical level and works its way down to the molecular level to find clinically relevant targets.

Entelos' PhysioLab technology is a computer-based representation of human physiology. The company has models for metabolism, asthma, and rheumatoid arthritis. Within each of these platforms, "virtual patients" represent particular combinations of genetic, lifestyle, and environmental factors. There might be more than 100 of these virtual patients within a given platform. The virtual patients are used to predict the outcomes of particular therapies.

Meanwhile, Gene Network Sciences focuses on drugs that are in the later stages of drug development, even as late as Phase III clinical trials. "That is the pain point of the pharmaceutical industry," Hill says. "It's not so much about finding additional genes and targets to go after, but it's about being able to bring compounds that hit those novel targets through the clinic." The company integrates preclinical and clinical data into computer simulations of human cancer cells and the heart. These models are then used in drug development alliances with pharmaceutical companies to determine the mechanism of action of new drugs and biomarkers for drug efficacy and toxicity.

One of the ultimate results of systems biology may be personalized medicine. Including information about individuals and ethnic groups in the models generated through systems biology will allow drug companies to go after diseases with smaller patient populations. "We'll

also get into niche diseases by using some of these systems biology tools because they will make [discovery and development] faster and cheaper," says Julie Bryant, vice president for business development and sales at GeneGo.

Another application of systems biology is in the evaluation of so-called combination products, in which drugs are administered together to achieve a synergistic effect, according to Pieter Muntendam, president and chief operating officer of BG Medicine, Waltham, Mass. Standard clinical trials are often ill equipped to demonstrate that the combination is greater than the sum of its parts, he says. A systems approach can aid in the selection of doses of the agents. "You use [clinical trials] to validate your findings from the systems pharmacology rather than as an attempt to discover whether the synergy exists," says Robert N. McBurney, the company's senior vice president of R&D and chief scientific officer.

COMPUTER CENTRAL Lots of computer power is required to handle the data sets and models involved in systems biology. BG Medicine's bioinformatic computer farm of Linux servers is shown here.

PHOTO BY STACEY HORRIGAN/BG MEDICINE

One result of systems biology is the emergence of large-scale hypothesis-driven experimentation. "People were just a bit drunk with the power of doing things in a high-throughput manner, without really being so clever about the kinds of questions they were trying to answer," Hill says. "This integrated systems biology approach allows us to go full circle back to hypothesis-driven biology and drug development."

Elliston, however, points out that the hardest part of systems biology is the development of testable hypotheses, something that is relatively straightforward in single gene or protein experiments. Looking at thousands of proteins simultaneously makes it next to impossible to develop rational hypotheses in the absence of models. "The root of the scientific part of systems biology is developing the hypothesis," he says. "Once you've developed the hypothesis, testing it is relatively straightforward."

The fruits of systems biology will take a while to make their way to the market in the form of an approved drug. Even then, it may be difficult to assess the contribution made by systems biology. "I'm not sure you'll ever be able to quantify specifically how [much] one technology contributes to the overall savings in the entire discovery and development pipeline," French says.

Elliston, a molecular geneticist by training, believes that the past 100 years of drug discovery have been a century of chemistry but that the problems now confronting drug discovery are ones of biology. "The next 100 years are going to be the century of biology within pharmaceutical discovery," he says. "Understanding the impact and the functioning of biology is going to rule how people can build not just single drugs but portfolios of compounds that have beneficial therapeutic impact in treating disease."

BIOMARKERS
Looking For Causes, Not Just Correlations

One of the hopes of genomic and proteomic analyses is that they will identify biomarkers–that is, a panel of easily measured genes and proteins that will directly correspond to the state of the biological system. However, biomarkers generated this way are correlative rather than causal, according to Keith O. Elliston, president and chief executive officer at Genstruct, Cambridge, Mass. Unless biomarkers are causally linked to a change in the systems, they can lead researchers astray as the researchers attempt to interpret the data. "Correlation not only doesn't prove causality, but it's also not scalable. A

correlation breaks down" when the number of samples becomes too large, he says. On the other hand, causality is related 1 to 1.

Causality can be discerned through cause-and-effect modeling, which identifies "the series of events that can most rationally explain outcomes. That [information] scales very readily because it's a biological effect," Elliston says.

Genstruct uses its models to generate hypotheses that will test a particular disease mechanism rather than simply generate a disease signature. Elliston claims that the biomarkers identified this way are causal, not just correlative.

To illustrate the difficulties of gene expression profiling and the advantages of cause-and-effect modeling, Elliston cites an example related to type 2 diabetes. In two similar experiments looking at the effects of a treatment on the expression of certain genes from different individuals, the expression levels of 700 genes changed. Of those, only 20 were common to the two sets tested, but those 20 could not form a disease signature because two-thirds were regulated in opposite directions.

The researchers found through cause-and-effect modeling that a common mechanism of action was involved in the expression changes of those 20 genes. However, although the mechanism was the same, the way the genes were responding to it varied from person to person. The 20 genes are unlikely to be useful as biomarkers, but the modeling has led to hypotheses about type 2 diabetes that the company is in the process of testing.

Chemical & Engineering News
ISSN 0009-2347
Copyright © 2005