

characterize thousands of proteins active at any time in various parts of the cell. Michael Washburn and Dirk Wolters at Syngenta Agricultural Discovery Institute in San Diego and John Yates at the Scripps Research Institute in La Jolla, California, have devised a system to separate and identify 1,484 proteins from the proteome of the yeast *Saccharomyces cerevisiae* (see *Nature Biotechnol.* **19**, 242–247; 2001). But that relatively low number in the humble yeast doesn't begin to reveal the complexity in humans. For example, there are a thousand or more proteins involved in the G-signalling pathway, which regulates everything from the most basic activities of the cell (division, motility) to the most specialized

ones (secretion, electrical excitability).

Perhaps the biggest hurdle is not in designing the equipment but in the conceptual realm. Researchers might know individual elements in a signal cascade, understand something about their function, and perhaps even have obtained their structure. But, explains Ehud Isacoff, a biophysicist at the University of California, Berkeley, scientists are still encumbered by a bias to view the overall picture as if it were made up of discrete events, with one protein handing a signal to another sequentially, in a series of 'stills'.

What is really happening in the cell, Isacoff continues, "is that proteins are very localized, and dock against one another very precisely in assemblies, and signalling happens by molecular motions that propagate from one subunit to another". New methodologies and systems of notation must be devised to describe these things, and a new breed of student has to be recruited who can think about them as concrete objects with specific structures and interactions.

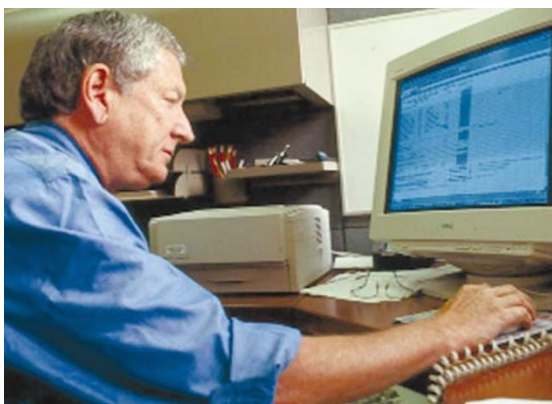
In fact, these needs are being recognized and the integrative effort is under way on several fronts. Leroy Hood's Institute for Systems Biology in Seattle has been in existence since early last year (see *Nature* **407**, 828–829; 2000), and Al Gilman's Alliance for Cell Signalling at Dallas set up shop a year ago (see *Nature* **407**, 7; 2000).

They aim at a holistic understanding of the cell in all of its pathways and interactions. New methodology — and, perhaps, improved equipment — may emerge from such efforts.

And a Clinical Proteomics Initiative, under the aegis of the US National Institutes of Health, started seeking grant applications last month. One of its key elements will be the antibody consortium, says Lance Liotta of the National Cancer Institute and one of those engineering the enterprise. This will be modelled on the open-access but industry-supported SNP consortium that is mapping simple genetic variations. Support — both in terms of finance and willingness to donate antibodies — from industrial and academic groups is very enthusiastic, says Liotta. The consortium's ultimate goal is to develop and make available arrays of every antibody and every ligand in existence.

Other aspects of the NIH initiative are looking for new approaches to existing techniques. However, it's unlikely that any new technology will completely replace an old one. Instead, innovations arising from the initiative will probably occur alongside the stalwarts of electrophoresis, mass spectrometry and chromatography — further complicating the ever-changing face of proteomics. ■

Potter Wickware is a science writer based in San Francisco; **Paul Smaglik** is *NatureJobs* editor.



Al Gilman: seeking the cell's secrets.

SOFTWARE

In the realm of software and databases, there is a real opportunity for integration, but instead developers have tended to go off in their own directions. Great strides have been made in areas such as image analysis and peak-picking tools for mass spectrometry with software packages including Tycho, Melanie and Quest. Software developed by Nonlinear Dynamics of Newcastle upon Tyne, UK, aids in spot detection on gels and also helps in quantitative analysis of those spots once they are picked. Major



Tony Pawson.

equipment manufacturers Amersham Biosciences and PerkinElmer have already signed on to bundle this program, called Progenesis, with some of their instruments. But, according to Patsy Babbitt, a protein informaticist at the University of California, San Francisco, the software side is fragmented. "It's a big problem," Babbitt says.

Organizations such as the Bio-Ontologies Consortium aim to clarify the picture with

standards and nomenclature, but perhaps what is lacking are new ways of thinking about the information generated in proteomics — classical bioinformatics is built around pattern-matching algorithms.

Tony Pawson and Chris Hogue at the University of Toronto have been thinking about the informatics side of proteomics. They have developed the Biomolecular Interaction Network Database (BIND), which indexes interactions between DNA, RNA, proteins and small molecules, as well as temporal and compartmental information. As BIND's content grows, "we'll be the GenBank of interactions", predicts Francis Ouellette, of the University of British Columbia in Vancouver, one of the resource's developers. Other databases for proteomics include the Database of Interacting Proteins at the University of California, Los Angeles, Large Scale Biology's Human Protein Index and Atlas Base, by the San Diego company Accelrys, which contains protein structures.



Chris Hogue.

Setting standards