HHMI
HOWARD HUGHES MEDICAL INSTITUTE

RESEARCH NEWS

FEBRUARY 12, 2001

Genome Wonderland Gives Way to Daunting Challenges of the Proteome

With the prodigious task of sequencing the human genome largely behind them, researchers now face the more daunting challenge of understanding the proteomeall of the proteins expressed in a cell, according to Howard Hughes Medical Institute investigator Stanley Fields.

In an article titled, "Proteomics in Genomeland" in the February 16, 2001, issue of the journal *Science*, Fields writes, "In the wonderland of complete sequences, there is much that genomics cannot do, and so the future belongs to proteomics, the analysis of complete complements of proteins."

Protein analysis, says Fields, is more complicated than figuring out the linear sequence of DNA genes because researchers must carry their analysis much further. "Proteomics includes not only the identifying and quantifying of proteins, but also determining their localization, modifications, interactions, activities, and, ultimately, defining their function," wrote Fields, who is at the University of Washington. Unlike DNA, proteins undergo complex biochemical modifications. A single gene, for example, can encode multiple proteins by means of alternative splicing of the messenger RNA. "All of these possibilities result in a proteome estimated to be an order of magnitude more complex than the genome," wrote Fields.

> "These technologies must be made widely available so that no community of biologists feels disenfranchised. At least some biologists see the genomics and proteomics revolution as a real threat, and widespread availability of reagents and equipment will help alleviate that concern."
>
> **- Stanley Fields**

In an interview about the *Science* article, Fields cautioned that scientists contemplating proteomics experiments will face major challenges as they learn new laboratory techniques and establish productive research collaborations. However, he expressed confidence that agencies that fund research are up to the task of fostering the necessary technologies and research infrastructure.

While researchers have gained a better understanding of thousands of proteins based on studies of their biological activity in the cell, Fields says that the huge number of proteins whose function is unknown provides a daunting challenge. Assigning roles to these proteins demands a new style of experimentation that "does not replace, but will increasingly operate in company with, the traditional way biology has been done," wrote Fields.

Fortunately, new technologies are enabling this work to proceed rapidly. For example, identifying small networks of proteins that work together has been aided by powerful analytical techniques such as mass spectrometrywhich sorts and identifies molecules based on their mass. And sophisticated genetic analyses using yeast are pinpointing associations between proteins, which can yield important clues about protein function. New biochemical tagging techniques have greatly aided attempts to determine the cellular location of proteins. Likewise, computer algorithms are being used to analyze protein sequences to try to identify proteins that have evolved together and thus are likely to act in the same cellular process, according to Fields.

The next major advances in proteomics, says Fields, will be technologies that measure how protein levels change dynamically during cell processes. Techniques of mass spectrometry analysis of complex protein mixtures "may allow human tissues to be used as the protein source and makes feasible the discovery of early disease markers by comparing the protein content of pathogenic cells with that of their normal counterparts," wrote Fields. He also cited protein arrays of tagged molecules as promising tools for exploring protein activity and function.

"For a field so laden with razzmatazz methods, it is striking that the number one need of proteomics may be new technology," wrote Fields. He cited, in particular, the need for streamlined assays to automate analysis of massive numbers of proteins. He also emphasized the need for widespread distribution of proteomic technologies. "Only when every laboratory is comfortable doing proteomics will its power be exploited fully," he wrote.

In an interview, Fields added that "these technologies must be made widely available so that no community of biologists feels disenfranchised. At least some biologists see the genomics and proteomics revolution as a real threat, and widespread availability of reagents and equipment will help alleviate that concern."

Also, he said, many issues remain unresolved about how to organize interdisciplinary collaborations required for advances in proteomics. "How do you get computer scientists to work with protein chemists to work with geneticists to work with combinatorial chemists, and so on?" he asked. "I don’t think that has been sorted out at all."

While plans to create interdisciplinary proteomics centers "have tremendous potential, it's not something that happens magically just because you put different types of scientists together," he said. Thus, he cautioned, "these kinds of interdisciplinary campuses, institutes, or buildings have the potential to be fantastic but also the potential to be unsatisfying. The scientists must speak a common language and develop problems that they all want to work on."

Proteomics will also change the scale of biological research, said Fields. "It is a major challenge to take scientists who are trained in a traditional approach to biological research of small laboratories and small-scale biology and introduce them to technologies that are extremely powerful, but often expensive and centered in proteomics centers or large laboratories," he said.

However, Fields said that he believes that federal funding agencies and foundations are "increasingly aware of the complexities, the problems, the need for technologies, and the funding to meet all these challenges."

Fields emphasized that the massive proteomics effort will be well worth it, because basic discoveries about the cell's protein machinery will likely yield rapid clinical applications.

"My guess is that the timeframe for a new proteomic-based diagnostic or treatment to go from the academic laboratory to prototype to engineering to commercial application will be rapid and will continue to shrink," he said.