Tech News

Proteomics

A Lot to Do

Proteomics—the systematic study of all the proteins produced by given cell types, tissues, or organisms—remains complicated by technical challenges. These include the large numbers of proteins that can be encoded by a given genome, the relatively small amounts of many of these proteins in cells or body fluids, and the variations in protein content and composition in cells of different types or stages of development. As proteomic-related analytic techniques are improved and more information is acquired, the computing software and power needed to analyze the resulting data will need to keep pace.

Developing Methods

Dr. Terry Gaasterland, a professor at Scripps Institution of Oceanography at the University of California, San Diego (San Diego, CA), has a new National Science Foundation (NSF) grant for her work with matrix-assisted laser desorption/ionization (MALDI) mass spectroscopy (MS). The goal is to image tissue samples to generate open-ended molecular snapshots of tissue regions. The proteomics-only part of the project analyzes proteins in tissue slices in order to create a high-resolution picture of the location of these proteins. Conventional localization of proteins in tissues relies on antibodies, but the drawbacks of this method are that antibodies can bind nonspecifically and can be used only a few at a time. Therefore, Gaasterland's group has developed a high-throughput MALDI MS imaging method, using the neural development of the medicinal leech as a model.

Gaasterland explains that the embryonic leech exhibits segmented growth from one end, adding segments daily for 12 days, so that it can be viewed as a time sequence of developmental events. It is also easy to grow and mount on slides, and its nervous system is large. The leech embryo is analyzed using MALDI MS at a 50- μ m resolution, generating a two-dimensional tiled array. The data generated can be used to determine which proteins are everywhere in the animal and which proteins have a unique location or appear at a unique time point during development. Size is a current limitation, as only peptides and small proteins can

Figure 1. A schematic diagram of the combined electrochemical-SPR-AFM instrumentation. Courtesy of Rigoberto Advincula, Department of Chemistry, University of Houston (Houston, TX).

be visualized. Gaasterland says that other groups are developing methods to do in situ trypsin digestion of proteins in tissues on slides. Resolving how to use trypsin in this way will allow her group eventually to analyze much larger proteins.

“I’ve learned, as a computer scientists working in biology, that I build better programs when I have a goal.”

There are other interesting questions that may also be answerable using these methods. For example, when ganglia in the adult leech are crushed, they regenerate. Gaasterland would like to know what proteins are involved in this process, and if new proteins are produced during regeneration. “I’ve learned, as a computer scientists working in biology, that I build better programs when I have a goal,” she says. As her group proceeds, they would like to look at brain sections from mice to confirm they can do similar types of measurements in this animal.

Gaasterland's group is also building new genome analysis software for proteome-directed genome assembly (ProDiGy) using high-throughput sequencing data and high-throughput MS data together to achieve results that cannot be done with DNA sequence analysis data alone.

Combine and Conquer

Rigoberto Advincula, a professor in the Department of Chemistry at the University of Houston (Houston, TX), has also recently received an NSF grant for his project, “Label-free protein arrays based on linear dendron macromolecular layers and in situ real time EC-SPR-AFM methods.” This work will further his group's effort to improve the specificity and sensitivity of protein arrays using a materials approach.

Their approach uses a synthetic, pegylated linear dendrimer, one end of which can be functionalized to capture specific proteins via either a histidine tag or biotinylation. The other end (dendron) can be anchored on a surface. This system can be designed to overcome the potential problem of nonspecific protein binding. “Making the binding highly specific is half the battle won,” Advincula observes. His group is using techniques such as the Langmuir-Blodgett method to prepare single (or very thin) layer films to direct protein molecules to bind in arrays in a whole-molecule layering approach.

The analytic instrumentation combines surface plasmon resonance (SPR) spectroscopy, which generates optical data, with atomic force microscopy (AFM). AFM enables imaging of peptides or proteins on the array surface, and provides additional information on conformational stability (i.e., whether the proteins are denatured or properly folded). The third component in the instrumentation adds electrochemical (EC) measurements. Whether the proteins of interest are oxidized or reduced can be determined by applying controlled voltages as positive or negative charges (Figure 1).

Biomarker Challenge

David Speicher is head of the proteomics core facility and director of the Center for Systems and Computational Biology at the Wistar Institute (Philadelphia, PA). His group is working on several oncology-related proteomics projects. One project uses proteomics to apply a systems biology approach to identify pathways and networks involved in the oncogenic process—particularly the transition to the metastatic phenotype. Another major project involves the discovery and validation of protein biomarkers in cancer that could be used for early diagnosis, predicting clinical outcomes, and monitoring response to therapy. “Biomarker discovery is a challenge

for us and others,” he says, observing that there is some skepticism in the field, in part because some researchers didn’t realize how hard it would be and initial expectations were unrealistic. “We are optimistic that new biomarkers will be discovered and used for diagnosis and disease management in a personalized medicine approach,” he says.

Speicher’s group is looking for biomarkers in blood plasma, and he points out that a big problem is that specific disease markers are present in low abundance among the large, complex mix of proteins in plasma. Nonspecific proteins such as acute phase reaction proteins and markers for inflammation are present in the plasma in much higher amounts than specific cancer biomarkers, and these proteins often show quantitative changes in plasma from cancer patients compared with normal controls. The majority of known cancer biomarkers, says Speicher, are present in nanogram-per-milliliter amounts or less. It’s also not clear how many proteins of all kinds are present in plasma. Assuming that most cells or tissues in the body are shedding proteins into the blood, then the proteins in the plasma would be expected to vary from person to person and with an individual’s state of health, he explains. Therefore, there is no such thing as one plasma proteome.

The use of MS techniques has helped in plasma biomarker discovery, but it remains difficult to set up experiments testing low abundance proteins in samples from large numbers of patients and validating them along with the appropriate controls. Validating biomarkers, Speicher points out, is a separate issue from identifying candidate biomarkers, and has its own distinct challenges. One issue is the difficulty achieving sufficient sensitivity in a cost- and time-effective manner. His group is working on a label-free approach that allows them to rapidly and economically set up validation screens for hundreds of candidate biomarkers. This is part of a multi-tiered strategy to enable selection of the most promising candidates, which can subsequently be confirmed with more quantitatively robust assays using specifically synthesized internal-standard peptides.

The ability to analyze large data sets has lagged behind their acquisition, Speicher says. However, he notes that his group is not running into the same challenges that are associated with high-throughput technologies such as DNA sequencing. Analysis of genomic data involves a limited number of approaches. Proteomics involves a wider variety of techniques, and the complexity of data and how to process and interpret them provide challenges at every step. One

issue concerns relating liquid chromatography–coupled mass spectroscopy (LC-MS) data to the protein sequence to determine how peptides in the sample correlate to protein databases. In the search for plasma biomarkers, analyses can produce several million MS spectra, and some platforms can’t handle this volume of data. But Speicher says this aspect doesn’t have to be a bottleneck, as it can be overcome by the addition of more computing capacity. Tools for meaningful interpretation of the data, on the other hand, are not as robust as he would like them to be: his group is developing more powerful software tools to suit their unique applications.

Future Directions

Speicher says that it is not unrealistic to anticipate the eventual development of a point-of-care monitor to quantitate a series of proteins that could be used for diagnostic and/or prognostic purposes. He explains that this would require the identification of informative proteins and the development of assays. His group is using MS to discover and validate biomarkers, but enzyme-linked immunosorbent assay (ELISA) and its equivalents are likely to remain the gold standard for protein diagnostic tests. In this context, Speicher says, he can envision multiplexed quantum dot or bead-based assays in the near future, to monitor these proteins. “I wouldn’t want to predict whether this might be in five, ten, or fifteen years,” he concedes, “because substantial challenges remain.”

Written by Lynne Lederman, Ph.D.

BioTechniques 47:727-729 (September 2009)
doi 10.2144/000113227

Free Subscription

BioTechniques is the first peer-reviewed journal with open access to the whole of the life science community.

100% life science
100% peer-reviewed methods
100% “everyday” practical information

Sign up for free at
www.BioTechniques.com/subscribe