

Proteomics: One small step for a digital computer, one giant leap for humankind

Michael J. Geisow

Borrowing heavily from proteins and genomics, proteomics attempts to bridge the knowledge gap between genomes and living cells, or rather, that constellation of proteins that gives living cells structure and function. Proteomics is already the next biological bottleneck in understanding such organisms as *Saccharomyces cerevisiae* and *Hemophilus influenzae*, whose genomes are completely sequenced. Just attempting to think about the functions and interactions of 6,000 yeast proteins would make substantial demands on the most experienced biochemist. The order of magnitude increase in higher species such as

humans takes this problem firmly into the domains of high-throughput screening and bioinformatics.

Two key technologies in experimental proteomics are two-dimensional polyacrylamide gel electrophoresis (2D PAGE) and electro-spray or laser desorption mass spectrometry. The resolving power of 2D PAGE, combined with the sensitivity and accuracy of mass spectrometry, allows individual proteins in mixed 2D gel spots to be identified. These two technologies have reached such a high degree of refinement, including partial or complete sequence analysis on femtomole amounts of material, that even subdomains within an otherwise unknown protein can be recognized and mapped to known sequence and function.

Pharmaceutical companies are keenly interested in exploiting proteomics to pin-

point drug targets by matching differential protein expression with disease states. Hand in hand with this goes pressure to understand the functions and patterns of expression of all gene products. Other companies see their opportunity in developing and selling informatics solutions or very large-scale libraries of genes or gene products for screening.

As with all aspects of biology, the World Wide Web is very helpful for locating information. The sites below include locations for completed microorganism genomes, 2D PAGE gel image libraries of the "click and tell" variety, and online tools that return probable protein identities when fed peptide mass data. A recent Internet search on the term *proteomics* returned a meager 2000 documents; however, this will soon change as the field expands.

Michael J. Geisow is a director of Biodigm Ltd. and secretary of the EU Structural Biology Industrial Platform (mike@biodigm.com).

A sampler of websites related to proteomics

Genomic Databases

Fully sequenced genomes Present In The Public Databases

<http://geta.life.uiuc.edu/~nikos/genomes.html>

This well-maintained resource catalogs and links the completed genomic sequences for viruses, bacteria, and other organisms. The *Genome* page of the EMBL Peptide and Protein Group (<http://mac-mann6.embl-heidelberg.de/ProteinResearch/Genome.html>) is also helpful.

Families, Function, and Expression

<http://info.gdb.org/~avoltz/family.html>

This valuable page at the Genome Data Bank highlights and links protein family databases and structural classifications.

YPD (Yeast Protein Databank)

<http://www.proteome.com/YPDhome.html>

Proteome Inc. (Beverly, MA) maintains the YPD and many helpful associated resources. A new feature at this site is invited minireviews on protein families.

2D Gel Databases

Links to 2D gel databases on the Internet

<http://www.phoretix.com/links.htm>

A comprehensive table of links maintained by the author at the Phoretix Ltd. (Newcastle upon Tyne, UK) website. Two particular examples are highlighted here:

2DWG catalog of 2D gel images

<http://www-lecb.ncifcrf.gov/2dwgDB/>

This is a metadatabase (i.e., it contains information about other databases). Data for each gel includes tissue, species, cell-line, image URL, database URL, gel protocol, organization URL, image properties, and map URL (if it exists).

Select Pair of 2D Gel Images to Flicker Compare from 2DWG or your URLs

<http://www-lmmb.ncifcrf.gov/flicker/totFlkPair.html>

Very neat way of spotting differences between two gels. The authors have written a powerful Java Applet.

Tools

EMBL Peptide and Protein Group Home Page (peptide search)

<http://www.mann.embl-heidelberg.de/Default.html>

Matthias Mann's group at EMBL has one of the best mass spectrometry sites, with many resources for proteomics. This site is one of several permitting online matching of peptide masses against the SwissProt Protein Databank. Other search tools are listed at <http://www.phoretix.com>.