Tool for proteomics

An automated approach to protein fractionation

BY J. S. HOBBS, M. H. SIMONIAN, E. BETGOVARGEZ, V. IOFFE, A. MOVSISYAN, J. YU, C. JAMES, AND O. BOGLER

Since the term "proteome" was coined in 1994, an integrated field of study has emerged that involves the systematic large-scale identification, characterization, and quantification of all proteins in biological pathways, together with their interactions. Generating such comprehensive data requires an integrated, multitechnology strategy encompassing cellular isolation, antigen identification, protein fractionation, protein characterization, data evaluation, and, ultimately, disease diagnosis and treatment. Although advanced analytical tools are being developed to streamline and automate the proteomic process, it is still clear that one tool alone will not meet the challenge.

Mass spectrometry has emerged as a key technology for identifying proteins, either intact (top down) or in fragments (bottom up), via digestion followed by peptide analysis. However, even with the most sophisticated spectrometer, analysis of whole-cell lysates is not practical without prior simplification of the sample.

By defining a proteomic process, it is possible to identify stages at which simplifications can be made and processes improved to remove bottlenecks.

The proteomic process

The primary objective of proteomics is identifying diagnostic biomarkers or therapeutic targets. By comprehensively analyzing differences between diseased and nondiseased, or treated and untreated, states, researchers aim to discover proteins that uniquely indicate disease or response to therapy. The proteomic process can be represented by the following series of activities.

Identification. The first step in the proteomic process is at the cellular level and focuses on identifying differences between normal and abnormal states.

Isolation. Next, the cells and/or subcellular organelles can be isolated for further study, allowing the differences to be compartmentalized.

Fractionation. Once unique traits have been identified and the required source of proteins isolated, the next step is to extract the proteins and fractionate them into manageable liquid fractions.

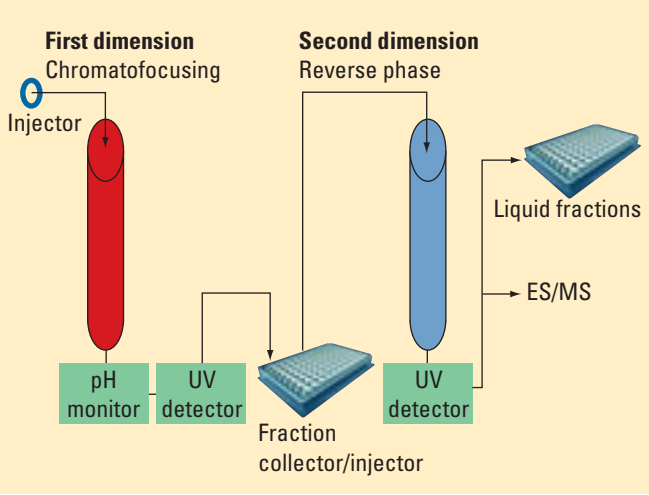

Figure 1. Automated protein fractionation system. (Image courtesy of Beckman Coulter.)

Characterization. Comparing profiles from the fractionation stage highlights differences between samples and indicates potential biomarkers, which are further studied to determine identity and function.

Evaluation. At this point, information about structure and function is used to evaluate a particular protein or group of proteins as potential markers for disease diagnosis or as targets for drug therapy.

Diagnosis. If all goes well, the end result is a targeted assay for diagnosing disease and, hopefully, a recommended therapeutic treatment.

In this article, we will discuss automated high-resolution fractionation of complex mixtures of intact proteins that results in simplified liquid fractions for mass spectral or other forms of analysis.

Protein fractionation

Once a particular cell species has been targeted and examples of normal and modified states isolated, lysates, which can contain as many as 10,000 different proteins, are prepared either from whole cells or from subcellular components such as mitochondria, nuclei, and other organelles. In the search for biomarkers, the fractionation protocol can be used for biological fluids such as plasma, serum, urine, or cerebrospinal fluid. Fractionation is usually performed by two-dimensional polyacrylamide gel electrophoresis (2D PAGE). 2D PAGE offers fairly high resolution but has considerable shortcomings, namely difficulties with highly acidic, highly basic, hydrophobic, and low-molecular-weight proteins; a narrow dynamic range; poor quantitation; and relatively high labor intensity (*1*).

Recently, other multidimensional techniques, such as chromatography, have offered promising results (*2*) as a complement to, if not a replacement for, 2D PAGE. The advantages of multidimensional chromatography include relatively high loading capacity (up to 5 mg of total protein) without significant band distortion, which improves the detection of low-abundance proteins; an enclosed liquid flow path, which reduces the risk of sample transfer loss and contamination; the use of solvents, which increases the range of proteins accessible for analysis; and liquid fractions, which obviate the need for gel extraction of proteins, facilitate storage for future analysis, and simplify automation.

An integrated tool

The ProteomeLab PF 2D system (shown schematically in Figure 1) provides an auto-

mated, multidimensional chromatographic solution for proteomic research. It includes proprietary chemistry, optimized methodology, and simple-to-interpret visualization software specifically designed for the high resolution of complex protein mixtures. A unique combination of chromatofocusing (CF), in which proteins are separated according to their isoelectric point (pI), followed by nonporous reverse-phase chromatography provides ultrahigh resolution of proteins and delivers data that can be related to those obtained from 2D PAGE.

Fraction collection in a 96-well plate is directed, from the first dimension, by an in-line pH monitor to ensure that proteins within the same pI range are always collected in the same fraction. The system can be programmed to vary the pH increments of successive fractions. Once collection from the first dimension is complete, the dual-mode fraction collector–injector module automatically injects the fractions into the second (reverse-phase) dimension. The eluent from the second dimension can be collected as liquid fractions for further analysis, such as by MALDI MS of intact proteins or their digests, or connected directly to electrospray ionization-mass spectrometry.

First dimension. CF has been around for many years but has not been applied to complex protein expression analysis until recently. Used primarily as a preparative purification tool, adapting it for reproducible, relatively high resolution fractionation of complex protein mixtures has proved challenging. Both a CF stationary phase and a protein-solubilizing buffer system have been developed that work over a pH range suitable for use with whole-cell lysates and biological fluids. The range of pH coverage is from 8.5 to 4.0, and fractions are typically collected at every 0.2 or 0.3 pH increment. Fractions can also be collected, based on time, before and after the gradient period to ensure complete recovery of all proteins eluted from the system. These fractions will contain the highly basic and highly acidic proteins, respectively.

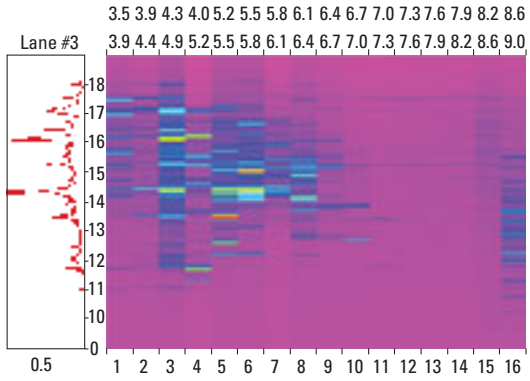

Figure 2. UV/pI map of *E. coli* lysate. (Image courtesy of Beckman Coulter.)

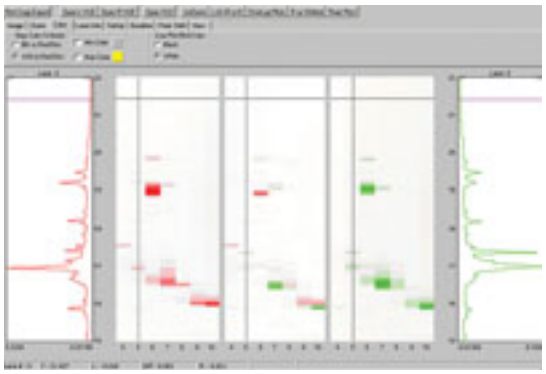

Figure 3. Comparison of fractions in the pI range 4.5–6.6 from fasting (left) and nonfasting (right) human plasma. Fraction 5 peaks are shown, illustrating quantitative and qualitative differences between the two proteomes. (Image courtesy of Beckman Coulter.)

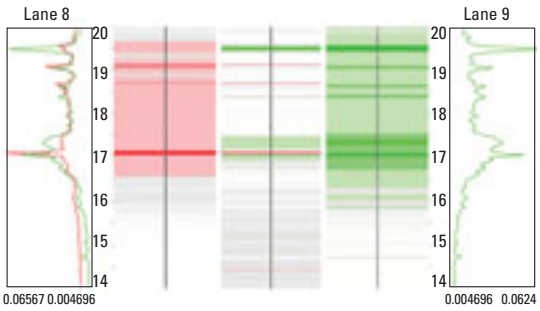

Figure 4. Cancer comparison. Close-up comparison of UV/pI maps of glioblastoma xenograft proteomes. Glioblastoma no. 6 is on the left in red, no. 12 is on the right in green, and the center panel displays the difference. An overlay of the chromatograms further enhances the detection of differences in the profiles. (Image courtesy of Beckman Coulter.)

Second dimension. Using nonporous silica-based columns provides for rapid, efficient separation of fractions from the CF dimension, and eliminating the support media's pore structure enhances both efficiency of separation and recovery of the proteins. Data combined from the two dimensions enable construction of 2D protein "maps" for easy visual comparisons.

Graphic visualization. To facilitate visualization of subtle differences between protein expression levels in normal and modified proteomes, software has been developed to convert chromatographic data into 2D "maps". Plotting pI determined from the first dimension versus retention time calculated from the second dimension produces a pI–hydrophobicity map of the intact proteins in the lysates (Figure 2.) By comparing maps from normal and abnormal states, it is possible to detect expression differences, indicating which fractions to concentrate on for further analysis.

Applications

The technique is applicable to a wide range of sample types, including stem cells, tissue cells, microorganisms, bacterial cells, plasma, serum, urine, cerebrospinal fluid, amniotic fluid, and plant tissue. The following examples show its potential.

Plasma analysis. To demonstrate the technique's power, we compared human plasma samples taken under normal conditions and while fasting (*3*). Blood samples were drawn from a healthy male adult after 12 h of fasting and again 2 h after a balanced meal. The plasma was carefully removed from each sample, aliquoted, and frozen at –80 °C. The protein concentrations determined by biuret assay were 83.95 mg/mL for the fasting plasma and 87.21 mg/mL for the nonfasting sample. Following equilibration of the chromatofocusing column with pH 8.5 Start Buffer, sample equivalent to 2.5 mg of total protein was injected and the pH gradient initiated.

Fractions injected into the second dimension were separated using an acetonitrile/trifluoroacetyl/water gradient, and separations were visualized with the systems software.

To determine areas of difference, profiles were compared using the differential display mode of the software, and the result is

shown in Figure 3.

Analysis of genetically characterized human glioma xenografts. Dysregulation of receptor tyrosine kinase signaling is a major contributor to cancerous tumors, including glioma tumors (*4*). High-level expression of epidermal growth factor receptor (EGFR) is frequently observed in glioma, usually in combination with wild-type and/or mutant EGFR gene amplification that has been associated with survival (*5*). The most common rearrangement leads to the deletion of exons 2–7 in the EGFR mRNA, causing an in-frame deletion of 801 bp in the extracellular domain. The resulting protein, deleted-(2–7) EGFR (also known as EGFRvIII, EGFR*, and ΔEGFR—referred to as ΔEGFR here), is a potent, ligand-independent oncogene for glioblastoma (*6*).

The PF 2D system can be used to identify molecular differences between glioma cells bearing amplified EGFR and amplified ΔEGFR in human glioma xenografts (*7*).

Two glioblastoma xenografts were established in nude mice, one that overexpresses ΔEGFR and another that overexpresses wild-type EGFR. Proteome maps were generated and profiles compared using the software suite. A zoom image comparing the profiles obtained with the two proteomes is shown in Figure 4.

Automated, multidimensional chromatography clearly provides a "ready-to-go" solution for fractionating and profiling complex protein mixtures.

Automated, multidimensional chromatography clearly provides a "ready-to-go" solution for fractionating and profiling complex protein mixtures. It also addresses many issues that are problematic with classical techniques, such as 2D gels. This technique results in a stronger foundation for the next steps in the proteomic process.

References

(1) Issaq, H. J.; et al. *Electrophoresis* **2002,** *23,* 3048–3061.
(2) Zhu, K.; et al. *Anal. Chem.* **2003** *75,* 6209–6217.
(3) Betgovargez, E.; Simonian, M. H. Beckman Coulter Applications Information Bulletin A1963A; 2003.
(4) Pawson, T.; Hunter, T. *Curr. Opin. Genet. Dev.* **1994,** *4,* 1–4.
(5) Ekstrand, A. J.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1992,** *89,* 4309–4313.
(6) Nishikawa, R.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1994,** *91,* 7727–7731.
(7) Ioffe, V.; et al. Beckman Coulter Applications Information Bulletin A1971A; 2003.

J. S. Hobbs, **M. H. Simonian**, and **E. Betgovargez** are from Beckman Coulter. **V. Ioffe**, **A. Movsisyan**, **J. Yu**, and **O. Bogler** are from the Laboratory of Brain Tumor Biology at Henry Ford Hospital. **C. James** is from the department of laboratory medicine and pathology at the Mayo Clinic. ■