

## **Proteomics Backgrounder: Advancing Proteomic Technologies to Combat Cancer**

### **The Potential of Proteomics for Detecting Cancer**

While genes are the “recipes” of the cell, containing all of the instructions for assembly, proteins are the products of these recipes and function as the cellular “engines,” the molecules that do much of the actual work to keep the cell and the body functioning. Some proteins are enzymes, catalyzing chemical reactions. Others play structural or mechanical roles, such as helping to maintain cell shape. Still other proteins participate in sending and receiving chemical signals, both within and between cells. And some proteins can perform several of these functions.

The term “proteome” refers to all of the proteins in a cell, tissue, or organism, while “clinical proteomics” refers to the study of proteomes in health and disease. Of importance to cancer research and treatment was the finding that tumors “leak” proteins and other molecules into blood, urine, and other accessible bodily fluids. This insight has led to the possibility of diagnosing cancer at an early stage simply by collecting such fluids and testing them for the presence of cancer-related molecules, or “biomarkers.” Such an approach could be clinically important because the earlier a patient’s cancer is diagnosed, the more treatable it is by surgery, radiation or chemotherapy, or drug therapy. Biomarkers found in blood and other fluids might also be valuable for monitoring the response to cancer during treatment or detecting the recurrence of tumors after treatment.

Certain blood proteins are already being used as cancer biomarkers. For example, elevated levels of prostate specific antigen (PSA) suggest the presence of prostate cancer, while elevated levels of cancer antigen 125 (CA-125) may indicate cancer of the ovary or other organs. Unfortunately, both tests may result in “false negatives” – failing to detect cancer in those who have it (poor sensitivity), or “false positives” – testing positive for the presence of cancer in people who are actually cancer-free (poor specificity).

Scientists have proposed that in order to develop more sensitive and specific cancer diagnostic tests, one should measure many biomarkers in a fluid simultaneously. It is estimated that there are on the order of 100,000 to a million different proteins in the human proteome, many of which may be found in the bloodstream. It is thought that patterns revealed in a panel of proteins (known as a “protein signature”) associated with a form of cancer might have better diagnostic and predictive capabilities than the current

single-marker approach. There is also a greater need to more rigorously correlate the proteins and their levels with the presence of disease – a process called “validation.”

### Key Developments in Proteomics

Several scientific developments have propelled proteomics research forward during the past decade, such that attention has now turned toward proteomics as a way to identify protein signatures and to validate their association with cancer. These developments include:

- **The mapping of the human genome.** Because genes are the “blueprints” for the production of proteins, an understanding of the human genome is necessary to accurately decipher the human proteome. The completion of the Human Genome Project in 2003 has been a major catalyst for proteomics research.
- **Advances in mass spectrometry.** New developments in mass spectrometry (MS) allow scientists to detect and identify ever-smaller amounts of proteins from increasingly complex samples. The method is extremely precise, enabling the detection of molecules that differ in composition by only one hydrogen atom (the smallest atom).
- **Development of protein microarrays.** Although still an emerging technology, protein microarrays are already powerful new tools for capturing and measuring proteins from blood and other biological samples. A protein microarray typically consists of a microscope slide-sized piece of glass or plastic, with up to several thousand different capture reagents – molecules that can “grab” specific proteins out of a solution – arranged on its surface in a grid (array). The method allows scientists to isolate and analyze multiple proteins from blood and other sources that may prove to be important as cancer biomarkers.
- **Development of mouse models.** Mice have been genetically engineered to carry mutations leading to cancers similar to those found in humans. Because the genetics of the mice, as well as their environment, can be well controlled, these mouse models are essential tools to cancer research and biomarker discovery. Mouse models also enable researchers to develop and test new proteomic technologies without having to use up the precious resource of human tumor samples, which can then be applied at later stages to identify and validate putative biomarkers of clinical relevance.

### The Challenges

Proteomics has advanced significantly with these developments, but continued progress is challenged by several obstacles. First, the proteome is highly complex, with large protein numbers and variations. Second, cells are continually modifying proteins once they are produced. As a result, the types of proteins measured can vary considerably from one person to another, under different environmental conditions, or even within the same person at different ages or states of health.

A third major complicating factor in the study of proteins is that they exist in a wide range of concentrations in the body. For example, the concentration of the protein albumin (a protein used for transporting fatty acids, hormones, and other substances) in blood is more than a billion times greater than that of the protein interleukin-6 (a protein which helps stimulate an immune response), making it extremely difficult to find the low-abundance proteins in a mixture. Unlike DNA, it is currently not possible to make copies of proteins that exist in very small amounts – and some scientists believe that the most important proteins for cancer may be those found in the smallest concentrations.

Analytical technology (such as mass spectrometry or protein microarrays) has advanced substantially in recent years, but further improvements are required to achieve the accuracy and reproducibility needed to make sense of the complex proteome.

Finally, beyond the technological hurdles that researchers must overcome to maximize the use of proteomics for cancer research and diagnosis, there are also procedural and organizational hurdles. Different treatment centers are likely to collect and store tumor samples in different ways, creating heterogeneity in the samples. Data are analyzed using software that differs from one laboratory to the next, and the results are stored in databases that cannot be easily shared among the research community.

### **Moving Proteomics Forward**

The research community requires improved proteomic technologies in order to:

1) increase its capability and reliability for understanding the underlying mechanisms of cancer and identifying biomarkers of disease, and 2) support the development of diagnostic tests that can be applied in a clinical setting. To achieve these objectives, several more advances will be required:

- **New technologies that can quantify proteins across the entire concentration range as well as detect modified versions of proteins.** Immense variations in protein concentrations and type are found within cells and fluids. Current technologies are often unable to identify and analyze all of the proteins in samples with a wide range of concentrations.
- **The standardization of proteomic technologies.** Current and emerging protein measurement technologies must be standardized and calibrated to produce comparable output between different laboratories.
- **Common bioinformatics resources, with shared algorithms and standards for processing, analyzing and storing proteomic data.** Proteomics informatics tools that permit data sharing and computation among laboratories are essential for rapid progress in the field.
- **A standardized procedure for processing and storing biological samples used in proteomics research.** Methods of biological sample preparation must be more

consistent to reduce variability in experimental results. Uniform sample quality, as well as access to large numbers of high-quality, clinically annotated samples, will lead to more reliable results.

- **Availability of high-quality reagents.** In particular, capture reagents that can be used in protein microarrays, as well as other techniques used to measure proteins, are essential resources.

Another development needed for the advancement of proteomics is an interdisciplinary team approach to science. No one laboratory working on its own could possibly examine all of the potential biomarkers, develop all of the necessary technologies for isolating and validating biomarkers for research or clinical use, or assemble all of the pieces of evidence required to understand the molecular mechanisms of cancer. It will require many laboratories working together to accomplish these goals.

In many ways, the challenges of developing clinical proteomic technologies and finding cancer biomarkers are comparable to those faced by The Human Genome Project, in which the work of sequencing the human genome was divided up among many laboratories throughout the world. All of the laboratories used standardized methods of collecting and analyzing their data so that the data could be assembled together at the end of the project. Similarly, standards are needed in cancer proteomics – standards for collecting and processing clinical samples, conducting experiments, and collecting and analyzing data – so that teams of scientists in different laboratories can collate and analyze their data to achieve meaningful results.

### **Clinical Proteomic Technologies Initiative for Cancer**

The National Cancer Institute (NCI) has launched the Clinical Proteomic Technologies Initiative for Cancer, a five-year, \$104 million program. The overall objective of this initiative is to build the foundation of technologies, data, reagents and reference materials, analysis systems, and infrastructure needed to systematically advance our understanding of protein biology in cancer and accelerate discovery research and clinical applications.

The specific objectives of the Initiative are to:

- Enhance technical abilities to identify and measure proteins accurately and reproducibly in biological systems.
- Advance proteomics as a reliable, quantitative field that can accelerate discovery and translational research.

For more information about the NCI's Clinical Proteomic Technologies Initiative for Cancer, please visit <http://proteomics.cancer.gov>.