

Proteomics: The Next Frontier

by Arvind Ravi

The entire human genome has been sequenced. Now What?

On April 4, 2003, the International Human Genome Sequencing Consortium announced completion of its work in sequencing the human genome. While describing the ramifications of the project, Dr. Aristides Patrinos, Director of the Department of Energy's Office of Biological and Environmental Research in the Office of Science, said, "We have opened the door into a vast and complex new biological landscape. Exploring it will require even more creative thinking and new generations of technologies" (<http://www.doegenomes.org>). But precisely what are these new technologies and innovative approaches that will extend the work of the Human Genome Project? And in what ways can these new approaches work towards improving health care?

Proteomics vs. Genomics

In many ways, genomics research has been fruitful, employing novel technology such as DNA microarray analysis to monitor the expression of thousands of genes at once. Consequently, researchers are becoming more interested in proteomics. While genomics seeks to catalogue an organism's genome, or entire set of genes, proteomics strives to do the same for an organism's "proteome", the set of proteins encoded by the genome (e-proteomics.net). However, many labs have moved beyond such an ambitious approach, instead shifting their efforts towards finding those proteins most critical to human health. Speaking for many in the field, Dr. Richard Zare, Stanford University Professor of Chemistry, says, "I am now really looking for patterns, and seeing if they can be predictive." Such a view evolved from the fact that effective medical treatment does not require knowledge of every protein in every cell of the body, but rather just the recognition of specific proteins and molecular patterns that are indicators of health.

In many ways, such exploration of proteins as opposed to the molecules "upstream" of them, such as messenger RNA (mRNA) and DNA, is well justified. First, proteins are immediately responsible for several facets of cell

life, such as structure, movement, reproduction, energy conversion, and communication. In addition, there is little relationship between a given cell's amount of mRNA and the cellular abundance of the proteins for which it codes. Consequently, protein levels are much better indicators of cellular dysfunction. Chemistry Professor Richard Zare, Ph.D., explains that doctors must not only know "what is there, but also whether it is overexpressed or underexpressed."

Furthermore, the same gene in DNA often leads to multiple proteins, as posttranslational modifications are sensitive to varying cellular conditions. Thus, while the genome contains 30,000 genes, estimates for the size of the human proteome range from 300,000 to 1 million! For these reasons, Ian Humphrey-Smith, a founding member of the Human Proteome Organization, asserts, "...without

Background: The Language of DNA

Sequencing the genome has the potential for critically advancing our understanding of both the normal functioning of the human body and any deleterious deviations from it, the so-called disease states. However, only limited insight can be gained directly from the sequence. Though DNA is the hereditary unit of information, proteins are the actual functional units.

DNA stores information in nucleotide bases, the commonly seen letters A, T, C, G, which stand for Adenosine, Thymine, Cytosine, and Guanine. This deceptively simple language is then transcribed into aptly named messenger RNA and transported from the nucleus to the cytoplasm. Here ribosomes, or cellular organelles fluent in both the languages of nucleic acids and proteins, convert the messages of mRNA into sequences of amino acids, the constituents of proteins. Once this process of translation is underway, other proteins within the cell work to make posttranslational modifications, and the chain acquires its mature functional structure (See Figure 1).

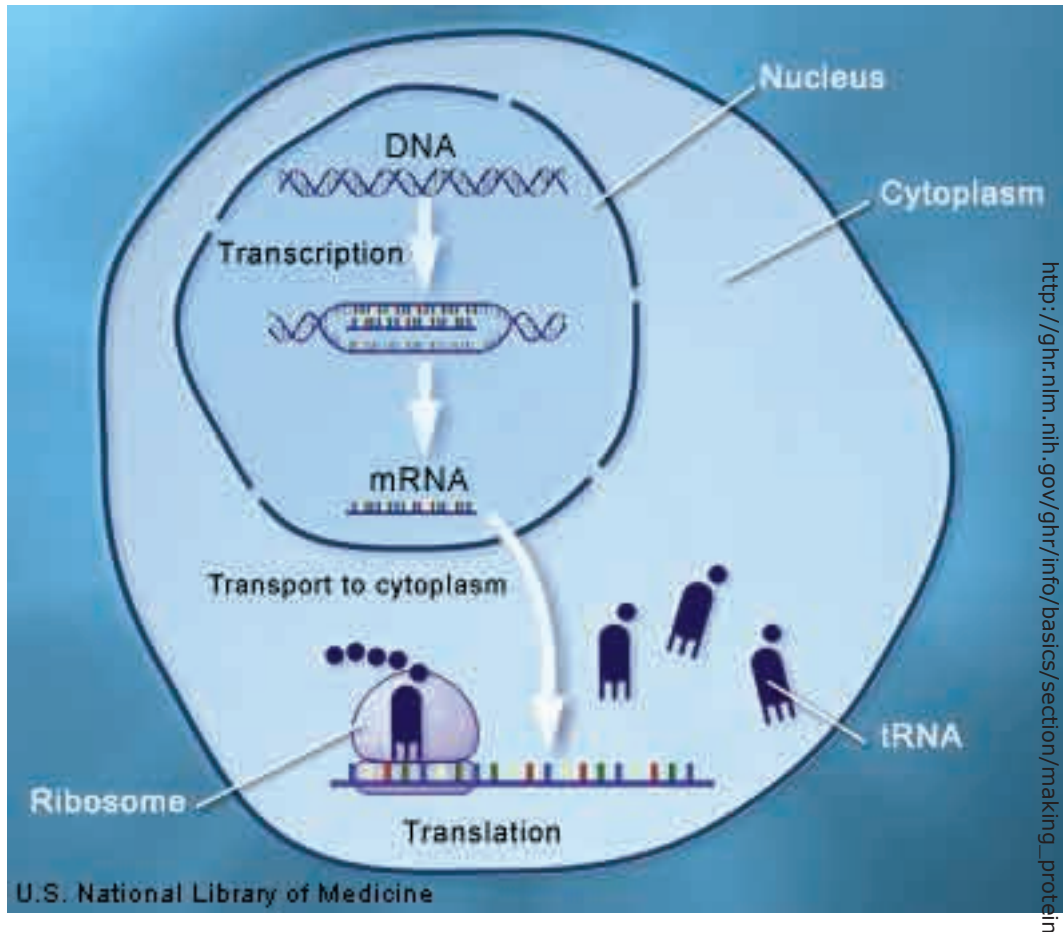


Figure 1: A summary of protein synthesis from DNA

a concerted effort in proteomics, the fruits of genomics will go unrealized” (<http://www.doegenomes.org>).

The Stanford Proteomics Center

Recognizing the importance of proteomics research, the National Heart, Lung, and Blood Institute (NHLBI), part of the National Institutes of Health, awarded Stanford University \$14.6 million in 2002 to fund the creation of a Stanford Proteomics Center. Though over 500 institutions applied, only ten were selected across the nation by the NHLBI initiative. As NHLBI Director, Dr. Claude Lenfant states, “Research at the level of the gene cannot provide a full picture of what’s going on within a cell. These state-of-the-art centers will help supply that missing information and so advance biomedical research and clinical care.” In order to make these innovations possible, the initiative will contribute a total of \$157 million dollars to proteomics research over a period of seven years.

Each site around the country has a different objective, with the goal of the Stanford Center being “Proteomic Analysis of Blood Components in Autoimmune Disease.”

Those who expect the Stanford Proteomics Center to be housed under a single roof will be disappointed, though, for as Center Director Dr. Garry Nolan explains, the center is really just a collaboration between various labs that share the goal of shedding light on the role of proteins in autoimmune disease: “There are really no [localized] centers anymore; science is far too interdisciplinary. We long ago gave up the mask that we only do one kind of science” (http://deansnewsletter.stanford.edu/archive/10_27_03.html).

Taking the philosophy behind Bio-X to heart, the Proteomics Center involves groups with diverse research backgrounds, with each group taking a slightly different approach toward autoimmune disease. Dr. P.J. Utz, an assistant professor in the Department of Immunology and Rheumatology, and Dr. Juan Santiago, professor in the Department of Mechanical Engineering and Director of the Stanford Microfluidics Laboratory, have joined forces in developing high-throughput methods of detecting ultra-small levels of cytokines, the proteins that coordinate immune response at the cellular level. One of the techniques they are investigating involves on-

chip liquid-phase microchannel networks that can perform various physical operations such as separation and mixing for small sample sizes (utx-microfluidics). Part of the power of some of the technologies they are investigating is that instead of requiring expensive machinery, they can be put together with “just the things lying around the lab.”

Professors Bill Robinson, M.D., and Larry Steinman, M.D., are working on antibody and antigen arrays. The idea behind this technology is that if antigens, or substances that trigger immune response, can be ordered on slides and then spotted with patient blood, the reaction of the spotting can be used to determine which antibodies are present, or which antigens the patient’s blood recognizes. These tests can then be used to investigate patients’ immune response at different stages of illness, or pre- and post- treatment. As Dr. Nolan explains, “If some [responses] disappear in the presence of certain drugs, then you know that drug is working.”

Rob Tibshirani, courtesy professor of Statistics, and Gill Chu, M.D. Ph.D. of Oncology and Biochemistry, work on the computational end of the projects with so-called

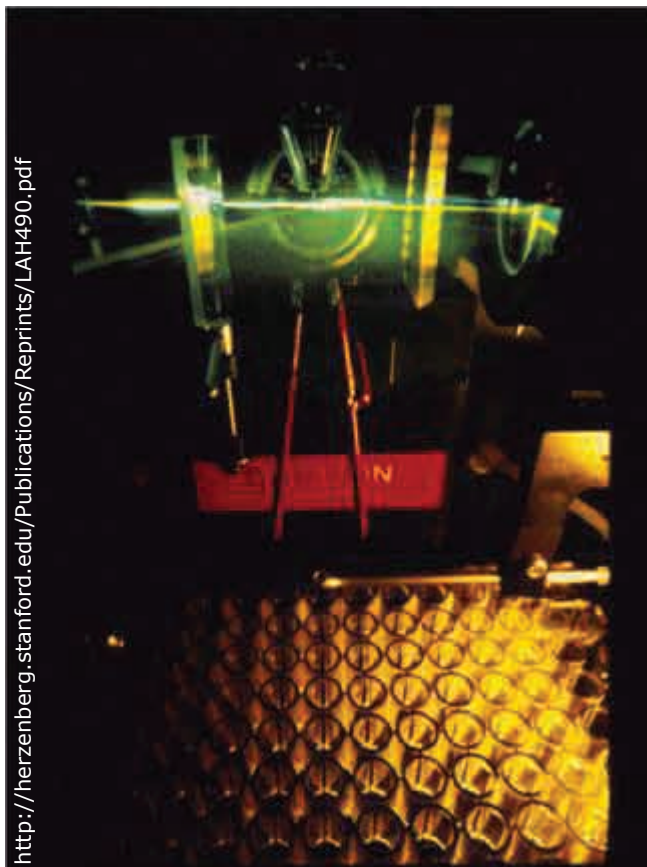


A robotic DNA microarrayer: As with many other cases, proteomic technologies evolved from this originally genomic device: researchers like Robinson and Steinman can now make antigen microarrays by spotting antigen solutions onto glass slides. These new applications are a critical step towards patient-specific therapy for autoimmune disease.

“mega-datasets.” They are developing the computational abilities to process the large amounts of data acquired in proteomics work, so that one may see “the signal amongst the noise.” Tibshirani and Chu’s previous collaboration led to theories that culminated in the program SAM, or Significance Analysis of Microarrays, which allows investigators to discern underlying patterns of expression from arrays of thousands of genes.

Dr. Nolan himself concentrates on single cell analysis and cell signaling systems. Using a technique known as flow cytometry in addition to Fluorescence-Activated Cell Sorting (FACS), Nolan is able to “reach inside of cells and read out those key decisions in the cells that are being made at the level of phosphorylation.” Older technologies such as immunofluorescence, in which fluorescent antibodies are used to bind target phosphorylated molecules in the cell, are inefficient and not truly quantitative. Nolan’s innovative use of technologies, on the other hand, allows for the phosphorylation of over ten molecules at a time to be quantitated in a single cell! Monitoring phosphorylated compounds is important in studying autoimmune disease because certain critical proteins in the cell signaling pathways can essentially be turned on or off by phosphorylation. As colleague P.J. Utz says of Nolan’s work, “[It] has the potential to blow certain fields wide open; it is really exciting.”

As Dr. Nolan explains, the motivation behind the



An older Fluorescence Activated Cell Sorting (FACS) machine, in which cells in charged droplets are sorted into an array of culture wells.

technologies of the Stanford Proteomics Center is both finding tools that can better screen drugs and developing tools that will allow doctors to stratify patients into separate treatment classes, even if they appear to have similar symptoms: “We are still in an era where we haven’t been able to stratify patients because we don’t have the downstream tools. What the proteomics centers throughout the United States are charged with doing is to provide the new tools so that the next generation of doctors will be able to stratify patients in retrospect.”

The Future of Proteomics at Stanford and Beyond

In the one year since the inception of the Stanford

Proteomics Center, funding has opened the door to further collaboration and a new set of proteomics technologies. As Dr. P.J. Utz notes, the implications for Stanford are far reaching: “Once you have a core you can attract funding from other sources to look into transplantation, cancer, and essentially anything we want.” As far as where proteomics itself will be, Dr. Utz notes, “The area is in its infancy still,” but he is optimistic that “it will take off in the next five years.” Within ten years, Utz suspects that the field may be comparable to where genomics is now. As with any new field, “It is limited by the available reagents and better technology, but it will take years to develop, and some luck.” **S**

Project Focus: The Zare lab

“Protein fingerprinting” joins proteomics and medicine

In addition to groups directly associated with the Stanford Proteomics Center, other researchers have taken up the task of better understanding the relationship between proteins and health. Dr. Richard Zare, Professor of Chemistry at Stanford, and his group of researchers work in close collaboration with several other investigators at Stanford, including Dr. Gilbert Chu and Professor Robert Tibshirani, to make technological innovations that will allow scientists to efficiently analyze a vast array of proteins from patients. They hope to develop a method for extracting information from patient blood: “I believe there will be fingerprints, patterns that we will be able to recognize and associate with disease states,” says Zare.

However, such work will require a new generation of technology to see it to fruition. One example of the technology with which Dr. Zare is experimenting is time-of-flight mass spectrometry (TOFMS). This technique involves the shooting of high-energy electrons at the protein molecules to make various ions. The ion fragments are then accelerated by an electric field due to their charge, with more massive particles reaching the detector later than light ones. From these conditions, the relative abundance of various particles with

certain charge-to-mass ratios can be determined. This characteristic pattern of fragmentation is like a fingerprint that can be used to detect a specific protein. Researchers at the Zare lab, in collaboration with other Stanford research groups, have enhanced this technique by using Hadamard transform time of flight mass spectrometry (HT-TOFMS). This method sends multiple ion packets to a detector by modulating between on and off positions, and then demodulates the signal to produce the spectrum. As Zare lab researcher Ignacio Zuleta explains, “The main advantage is that we can take more spectra per unit time than with a conventional beam. What has been shown in this lab is that we could acquire 377 spectra per second, whereas normally you would get 10 per second.” This technology allows for an increased rate of protein sequence analysis.

While Dr. Zare continues to explore and improve aspects ranging from separation of proteins to mass spectrometric detection, he keeps track of the potential of this technology to improve medicine. As he explains, with this technology, projects can be undertaken to understand “whether people have certain risk factors for things ranging from disease to response to radiation treatment.” Simply put, he says, “We’re looking for markers.”