

those countries in which the regulations are implemented has risen'. Wagner believes that another important component in reducing animal use is the sharing of results and animals between scientists. This will help to reduce duplication, but it is, perhaps, a little utopian given the commercial concerns—patent and intellectual property rights—often associated with such research.

So animals in biomedical research are here to stay. Their justification, however, hangs in the balance of ethical arguments, the education of the younger generation of scientists, and better formulation of scientific and medical aims. As Wagner concluded: 'I'm not in principle against scientists thinking a bit more before they do an experiment.'

Andrew Moore

Email: andrew.moore@embo.org

DOI: 10.1093/embo-reports/kve147

Proteomics: biology in the post-genomic era

Companies all over the world rush to lead the way in the new post-genomics race

In June 2001, when the Human Genome Project and Celera completed the first maps of the human genome, Francis Collins, head of the government-sponsored HGP, warned that only then would the real race begin. This was a prophetic insight indeed. No sooner was the human genome decoded than we found ourselves in the 'post-genomic era'—where the name of the game is proteomics. Proteomics is not only the systematic separation, cataloguing and study of all of the proteins produced in an organism, it is also the study of how proteins change structure, interact with other proteins, and ultimately give rise to disease or health in an organism. Since its application in drug discovery promises huge economic returns, it comes as no surprise that biotechnology, computer and software companies around the world are rushing to pour capital and resources into this new research field.

Proteomics is more complex by several orders of magnitude than genomics, with no one company, laboratory or consortium

No one technology will be able to fulfil proteomics' numerous tasks and new developments are sorely needed

remotely able to run the race alone, acknowledged Brian D. Chait, head of the mass spectrometry laboratory at Rockefeller University in New York City. Moreover, no one technology will be able to fulfil proteomics' numerous tasks, and new

developments are sorely needed, added Roy Whitfield, CEO of Incyte Genomics (Palo Alto, CA).

With the publication of the draft of the human genome in the February 16, 2001 issues of *Science* and *Nature* came what many had already suspected: instead of the earlier estimate of about 100 000 human genes, the actual count reduced this figure by 75%. If humans have only 10 000–20 000 more genes than the fruitfly and the roundworm, then the big question according to Eric Lander, head of the Whitehead-MIT Genome Center in Cambridge, MA, is how do we manage to be so complex? The answer: proteins—not genes—are responsible for an organism's complexity. The interaction of proteins in a complex network adds up to how an organism functions, according to Denis Hochstrasser, Professor of Medical Biochemistry at the University of Geneva's (Switzerland) Faculty of Medicine. The key to understanding health and disease within an organism is, therefore, to understand how its proteins function. 'In a multicellular organism, one needs to be able to look at the entire system in an integrated way. Proteomics is the study of where each protein is located in a cell, when the protein is present and for how long, and with which other proteins it is interacting', said Brian Chait. 'Proteomics means looking at many events at the same time and connecting them,' he added. New tools are necessary to enable the study of this web of events—to create a 'movie', rather than a static snapshot of the activities taking place.

A further dimension was added to this complex picture when scientists from the University of Pennsylvania School of Medicine reported this May in *Nature* that proteins are more active and dynamic than they had imagined. The researchers

The concept of proteins as dynamic entities may ultimately help scientists target more accessible sites for drug development

used nuclear magnetic resonance imaging to track the activity of a calmodulin-peptide complex across 13 different temperature settings ranging from 15 to 73°C. The data showed that there is a much larger range of internal motion in calmodulin than crystallographic studies can show. 'The interior of a protein is much more liquid-like than scientists originally anticipated,' A. Joshua Wand, Professor of Biochemistry and Biophysics and principal author of the study, said. 'Everything is moving, and it's moving all the time, very fast. They [proteins] move so much that potentially it dramatically influences how they work,' he said. 'This is the beginning of a long new story that will have a lot to do with understanding protein function.' The concept of proteins as dynamic entities may ultimately help scientists target more accessible sites for drug development, said Wand.

Currently, drug developers are working with only about 400–500 targets, many of which are receptors. With the shift from

analysis

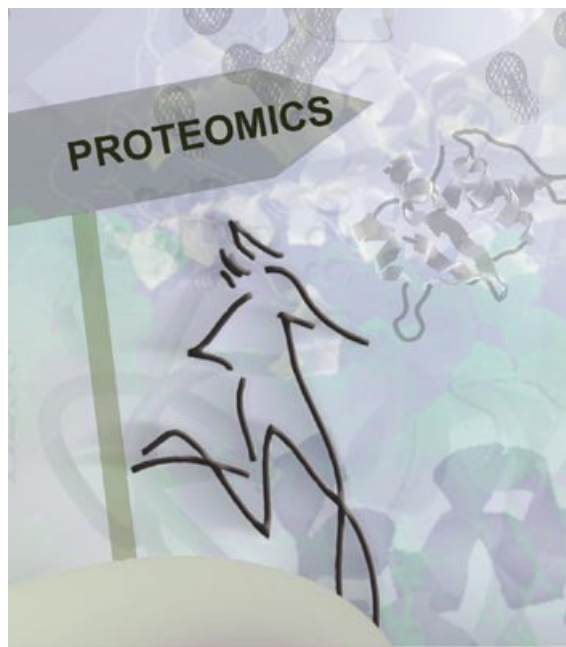
genomics to proteomics and the concomitant evolution of technology, many scientists expect the number of potential 'druggable' targets to expand many hundred-fold to between 10 000 and 20 000. With such numbers, it will become necessary to winnow through targets rapidly and accurately to determine which should be pursued. The marriage of business and science within the proteomics field indeed promises to achieve this.

A formidable task for proteomics is to develop new tools that can help scientists analyse cellular function with speed and accuracy. Proteins are too numerous, diverse and interactive to be studied by a single method. Proteomics, therefore, is comprised of a number of interrelated, overlapping disciplines: functional and structural genomics, functional and structural proteomics, and bioinformatics—a convergence of 'wet' and 'dry' laboratories.

One immediate challenge is the automation of the two major extant proteomics technologies, 2-D gel electrophoresis and nuclear magnetic resonance imaging. A leader in the proteomics field, Oxford GlycoSciences (Oxford, UK), recently teamed up with the Institute for Systems Biology (Seattle, WA) to develop an industrial high-throughput proteomic platform. At a recent symposium on proteomics held at the New York Academy of Science, Denis Hochstrasser described the need for faster and better ways to analyse proteins on a large scale. He and colleagues are working to develop a molecular scanner that would automate the separation and identification of thousands of protein types in a cell. This would combine the information gleaned from protein separation with protein databases, analysis and characterisation. The goal of Hochstrasser and other clinicians is to be able to send a specimen to the laboratory and determine what type of cancer it may be, at what stage, and to which drugs it might be susceptible.

A spate of start-up companies are springing up on both sides of the Atlantic which are homing in on specific aspects of proteomics: Structural GenomiX, Syrx (both in San Diego, CA) and Astex Technology (Cambridge, UK) have expertise in structural genomics; Cytos Biotechnology (Zurich, Switzerland) and Gemini Genomics Ltd (Cambridge, UK), excel in functional genomics. Caprion

Proteomics (Montreal, Canada) is concentrating on finding and virtually mapping proteins in organelles, which exist in low abundance but are highly significant in



disease. Large Scale Biology (Vacaville, CA), Proteome Inc. (Beverly, MA), Integrative Proteomics (Toronto, Canada) and Oxford GlycoSciences (Oxford, UK) specialise in protein expression profiling; and Axcell (Newton, MA) and Myriad Genetics (Salt Lake City, UT) are focusing on protein-protein interactions. On June 6, 2001, Large Scale Biology announced that using their database they have uncovered a number of markers for cardiovascular, psychiatric diseases and for liver toxicology in drug testing. They plan to use the liver genes in clinical trials in the near future.

Other companies, including Celera and Proteome Inc., are developing proteomic databases. Celera's goal is to analyse up to 1 million proteins per day. Information derived ultimately from the body's proteins should produce more specific,

tissues, which it is using for diagnostics, drug and drug target discovery. 'We expect the protein discoveries for medical applications enabled by this database to drive a major shift in pharmaceutical R&D from genes to proteins, and the development of technology for personalised medicine,' said Leigh Anderson, President of the company's proteomics subsidiary. He estimates that its index covers the protein products of about 18 000 human genes.

Bioinformatics boutiques such as GeneFormatics and Structural Bioinformatics (both in San Diego, CA) help companies plot protein activity. GeneFormatics uses algorithms to help predict the function of proteins encoded by newly discovered genes by comparing the proteins in question to those of known structure, which generates a 'sketch' of what each protein looks like and then suggests what its functions could be. Structural Bioinformatics actually makes digital 'movies' of how proteins change shape when they interact with drugs.

This helps find small molecules that can arrest these molecules in action.

At the same time, large genomics companies like Incyte, Myriad and Celera have increased their investment in proteomics and are striking deals with other smaller companies, such as the former's collaboration with Genicon Sciences (San Diego, CA), to measure infinitesimal amounts of protein in biological samples. Celera recently contracted with Odyssey Pharmaceuticals (San Ramon, CA) to use Odyssey's functional assays with Incyte's gene transcripts to evaluate protein interactions in living human cells. Using gene transcript data, gene expression and bioinformatics, Incyte has discovered that over 100 genes appear to be involved in insulin signaling pathways; and it will validate these as drug targets using Odyssey's Protein Contact Assay technol-

With 10 000–20 000 potential drug targets, it will become necessary to winnow through targets rapidly and accurately to determine which should be pursued

powerful and individualised therapies. In January 2001, Large Scale Biology announced it had completed the first version of its human proteome database, the Human Proteome Index, an inventory of proteins found in all major human

organs with a view to improving the treatment of diabetes. Numerous companies, such as Large Scale Biology and Biosite Diagnostics, are teaming up to develop protein chip arrays as tools to measure large numbers of proteins in cells and tissues, while

analysis

Ciphergen (Fremont, CA) specialises in developing such chips. Myriad recently struck a deal with Hitachi and Oracle for US\$ 0.5 billion dollars to identify all human proteins and their interactions.

Instrumentation companies like Tecan (Zurich, Switzerland) are broadening their focus to include proteomics supplies. 'We are increasing our R&D in this area from less than 1% to about 20% this year,' said Gregory Porter, Tecan's global marketing manager for proteomics. Last year, Tecan had US\$ 3 million in proteomics sales; this year it had the same amount in the first quarter. The company plans to hire new scientists and hopes to have 30 working in the area by the end of 2001. Large computer firms have also jumped on the proteomics bandwagon; Sun Microsystems and IBM formed a

co-operation with Oxford GlycoSciences (Oxford, UK). Most recently, on May 30, 2001, IBM and MDS (Toronto, Canada) agreed to create a database for proteins of many organisms, including humans, each company contributing nearly US\$ 3 million. The new not-for-profit company, Blueprint Worldwide Inc., plans to consolidate data in the public domain, including information from 200 000 scientific papers, government research and biotechnology companies willing to make their work public, and supply the information free of charge.

Large pharmaceutical companies are also investing heavily in-house as well as licensing in technology and doing deals. 'I see proteomics as a group of technologies that may be able to leverage what we're really trying to accomplish at Pfizer—the

discovery and development of new medicines,' said B. Michael Silber, Director of Pharmacogenomics and Clinical Biochemical Measurements (New York, NY). Ultimately, 'these are the very early days of proteomics,' said Incyte's Whitfield. 'While we understand how powerful the technologies can become, we have a long way to go,' he added.

Vicki Brower

The author is a freelance science writer in New York, NY.

E-mail: vickibrower@cs.com

DOI: 10.1093/embo-reports/kve144