Innovations

In search of the proteome

Oxford GlycoSciences plc

Chemistry & Biology September 1999, **6**:R259–R260

Glycobiology is not a glamorous discipline. As a founding technology for Oxford GlycoSciences plc (OGS; Oxford, UK) it was solid but dull. Companies like Genentech Inc. (South San Francisco, California) needed the sugars on their protein drugs analyzed, and OGS were the people to do it. The sugar theme continued in the study of a few obscure genetic diseases involving defective sugar metabolism. But then there came two fateful events: OGS bought a division of Millipore Corporation (Bedford, Massachusetts) that sold two-dimensional (2D)-gel systems; and the Board of Directors of Genentech asked their CEO, G. Kirk Raab, to resign.

Although OGS was doing some work for Genentech, Raab had never heard of them. But a friend told Raab that OGS was looking for precisely the person that Raab was hoping to become: a benevolent and powerful business adviser who could remake companies and bring in plenty of cash. Raab became the non-executive Chairman of OGS. In short measure he replaced the majority of management, hired a new CEO, and raised over £13 million in venture capital. Less than four years later, OGS is poised to move into new quarters in Oxfordshire, UK. The new facility will allow OGS to run 100,000 2D gels per year, and identify the proteins on those gels at a rate of 1000 per day. Glycobiology and modesty are dead; long live proteomics and grand ambition.

This new science amazes me

"In 1996 people hardly knew what proteomics meant," says Raab. The same could be said for many people today. In proteomics, much has been claimed, but far less has been demonstrated.

Proteomics has the same theoretical basis as its more developed cousin, transcription profiling. If you can detect and quantify thousands of mRNAs (or proteins), then you can compare the profiles of any two states: diseased versus non-diseased, drug-treated versus non-drug-treated. If the amount of an mRNA (or protein) goes up or down, you have a potential disease marker, drug target, or mediator of drug action or toxicity.

Not content with finding every gene, OGS wants every protein.

Post-transcriptional controls can make mRNA levels misleading. And, thanks to post-translational modifications, "the human proteome may be five to ten times bigger than the genome," says Andrew Lyall, formerly the Head of Bioinformatics at Glaxo Wellcome and now chief information officer at OGS. Once drug researchers start experimenting, that vast resource just keeps expanding. "Every time you give a new drug," says Lyall, "you've got a new proteome."

Fishing expeditions

So what to do with all this information? Just as with transcription profiling, proteomics can be used to find disease targets, based on their increased expression levels in a disease state. But proteomics runs into similar problems: there are lots of changes in a disease, and only a few of the changes are relevant to disease causation or possible treatment. Besides, there is more money and time spent later in the drug discovery process. "There's a wide array of technology available for target discovery, so that market is reasonably well served," says Leigh Anderson, CEO of the proteomics company Large Scale Biology Corporation (Rockville, Maryland). "But there isn't much technology to accelerate development."

Most of the clinical samples from drug trials are fluids, such as serum, cerebrospinal fluid and urine, that do not contain DNA. But they do contain proteins, so the samples can be searched for surrogates of disease or toxicology. Disease markers can give early clues to a drug's efficacy, and toxicology profiles can quickly explain how a new drug is causing trouble. According to William Rich, CEO of the proteomics company Ciphergen Biosystems Inc. (Palo Alto, California), pathologists using proteomics have "never discovered so many markers in their life."

Concepts such as these are the focus of an OGS deal with the clinical-testing company Quintiles Transnational Corporation (Research Triangle Park, North Carolina), while the whole spectrum of clinical and diagnostic markers and possible drug targets are under investigation in an Alzheimer's disease program with Pfizer Inc. The same logic applies to plants: with Pioneer Hi-Bred International Inc. (Des Moines, Iowa), OGS is looking for the basis of desirable seed traits. Finally, OGS is getting into the business of selling information in a deal with Incyte Pharmaceuticals Inc. (Palo Alto, California), who are providing software and informatics to go with the data generated by OGS.

How to see all those proteins

The core of proteomics is the ability to detect tiny amounts of protein using mass spectrometry (MS). Methods for generating and detecting ions of biological molecules have evolved rapidly in the last 20 years, with sub-picomole detection now common, and low femtomole detection becoming routine.

Before detection comes separation. The standard technology here is 2D

gels. Sorting by charge in one dimension then size in the second dimension gives an array of spots numbering in the thousands, although many of those spots represent the same protein with different modifications. Proteomics companies lay claim to various improvements on the basic method, including larger gels, new matrices, and gels bonded to glass for storage, but, as Anderson says, "you have to wonder about a technique that was invented 24 or 25 years ago and is still in use."

And there is plenty to wonder about. Estimates of the number of genes expressed in a given cell type range from 10,000 to 30,000, so the few thousand spots on a standard 2D gel fall well short of identifying all proteins. Proteins that are small, low in abundance, or membrane-bound are systematically excluded. And the technology is expensive, cumbersome, and notoriously open to operator error. "It's very multifactorial," says Anderson. But, he says, making the process reproducible is "what the commercial sector can do. One person's hocus pocus is the other person's proprietary technology."

OGS is using other protein analysis methods such as immunoprecipitation of protein complexes, and labeling of membrane proteins followed by ligand-based purification. "Proteomics to us means a lot more than 2D gels," says Raj Parekh, the chief scientific officer at OGS. "The key is the mass spectrometry and the informatics." But OGS has made a huge investment in 2D gels, and for now they rely heavily on this imperfect technology. "The most important question," says Anderson, "is, what is the alternative? I don't see any."

The trouble with chips

Transcription profilers have put all their genes on chips, so perhaps a protein chip can rescue researchers stuck in a maze of 2D-gel dots. Ciphergen have made a modest start in this direction by creating chips with various surface chemistries. With combinatorial variations of chip surfaces and washing conditions, researchers can trap a few hundred proteins on a given chip. Mass spectrometry from the chip surface is then used to look for expression differences between two biological samples. The system is simple to use, can handle crude and minute samples, and can be tuned to focus on the proteins that 2D gels miss. "We're designing this to complement and not compete with technologies like 2D gels," says Ciphergen senior scientist Enrique Dalmasso.

Replacing 2D gels will take a more ambitious program. One possibility is an antibody array, where the binding of uniformly labeled proteins is detected by fluorescence. Generating the antibodies for such a chip "is a potentially tremendous hurdle because what you need is exquisitely specific antibodies," says Parekh. Alternatives such as aptamers or various chemical surfaces may turn out to be more reasonable, but any chip system will have to deal with the huge dynamic range of protein expression, and the difficulty of keeping a diverse group of proteins soluble, folded, and competent for binding.

"It's not the fault of the technology, it's the fault of proteins," says Anderson. "Nucleic acids are basically all the same but proteins are very heterogeneous, so it's always going to be more difficult to get proteins to behave."

Parekh says protein chips are "worth pursuing, but we're quite a way from them having discovery potential." Even a successful protein chip is unlikely to do away with the trusty 2D gels. "I don't see that there will be one all-encompassing technology that becomes the only way to do proteomics," says Parekh. "Because of the greater physicochemical diversity of proteins that is an unreasonable goal."

The difficulty of dealing with proteins on chips gives Raab some comfort. "That makes me confident," he says, "that what we are doing [with 2D gels] will remain useful for a very long period of time."

Who needs a gel?

Ruedi Aebersold at the University of Washington in Seattle is looking for another alternative to 2D gels. "We're going on the assumption that the technology that will ultimately be used for analysis of the proteome has not yet been developed," he says.

The key for Aebersold, a member of the OGS scientific advisory board, is quantitation. The problem is that successive runs in a mass spectrometer cannot be used for accurate comparisons. Aebersold's solution is to label two samples with a single reagent. One sample gets the isotopically heavy reagent, and the other the isotopically light reagent. The mixture is then put in a single liquid chromatography system followed by tandem mass spectrometry. Proteins from the two samples emerge as matched pairs, and can be quantified and identified in one fell swoop.

Growing fast

Technology such as Aebersold's will be vital for OGS, but Parekh says that aggressive financing is, if anything, more important. "Our approach has been to get big first," he says. "We've invested a lot internally in the hope that we will recoup it later."

"Cash is so important," he continues. "The successful technology companies that have grown required enormous investments in cash to get them going. The pack snapping at their heels behind them are still snapping at their heels because they have been cash starved."

For some at the company it is already time to declare victory. "People thought they'd get more mileage out of transcription profiling because it would take so much longer to get proteomics going," says Lyall. "OGS has sorted out proteomics so quickly that we've really caught people by surprise."

William A. Wells
1095 Market Street #516, San Francisco, CA 94103-1628, USA; wells@biotext.com.