GILDER
BIOTECH
REPORT

February 2002
Vol. 2, No. 2

By Scott Gottlieb, M.D.

Published by Gilder Publishing, LLC

Plumbing the Proteome

Genome is only half the story; MDS Proteomics invents the tools to digitize a whole new industry

Meet my new best friend, *alpha fetoprotein.*

In the last few months, after admitting two-dozen liver transplant patients, I've developed a deep appreciation for this particular protein, both for its own sake and as a harbinger of even better things to come. Take Michael, who like many such patients, ended up on the transplant waiting list. Michael's own liver was ravaged by hepatitis C, and he was at high risk for developing a kind of cancer called hepatocellular carcinoma.

Between the hepatitis and the prospect of liver cancer, Michael desperately needed a new liver. The problem is that a patient with widespread liver cancer is not likely to do well even with a new liver. Until we can cure the cancer, doctors must save scarce organs for those who are most likely to benefit. But it's not always easy to tell who has liver cancer and who does not: x-rays and even biopsies sometimes miss tiny tumors. How to identify early, easy-to-miss cancers?

Enter alpha fetoprotein, which is secreted by tumor cells into the blood and measured with an ordinary blood test. Michael's level was off the charts, and we concluded he probably had liver cancer. Little liver spots on the CAT scan confirmed the diagnosis. Bad news for Michael. Good news for another patient, whose life (unlike Michael's) could probably be saved by a new liver.

Yet alpha fetoprotein is just a first-generation protein marker: it can't pick up the tiniest tumors, and it can be elevated in patients with testicular or ovarian cancer or with other serious liver problems, making diagnosis confusing. This protein marker can only tell you so much; doctors use it because it's the best we have right now.

But doctors and investors alike are salivating over the medical innovations just around the corner. Imagine the day, soon to come, when researchers have discovered and catalogued proteins accompanying diseases. Imagine simple blood tests that accurately diagnose hundreds of illnesses.

In the near future, new diagnostic tools will rely not on spying crude symptoms, but on detecting the underlying molecular processes that trigger disease—weeks, months, or years before the doctors see a difference. For many diseases from cancer to diabetes to arteriosclerosis, the best time to diagnose would be before the patient has developed any external symptoms at all.

Protein markers are the key to spotting and tracking diseases already active, long before their presence in the body can be spotted by doctors or by crude tests like x-rays. Genes tell doctors what is going to happen; proteins tell us what is already happening in the body.

How will the new protein diagnostics change medicine? Consider an ongoing National Cancer Institute study looking at a new drug called imatinib mesylate. This drug works by blocking a protein receptor called *tyrosine kinase*. Tyrosine kinase works as a microscopic on/off switch inside complicated signaling cascades, sending subtle instructions to turn on or shut down other regulatory nodes. Tyrosine kinase inhibitors jam the signal.

Why would a doctor want to jam that signal? To turn off the molecular signals that instruct a patient's cancer cells to proliferate. Imatinib mesylate is being tested for its potential to block ovarian tumors. But researchers are also testing a different possibility: whether

INSIDE:

four weeks into treatment the activation of a protein marker in the tyrosine kinase pathway correlates with standard measures of drug efficacy, such as reduction in tumor size on an x-ray. Thus the result of these clinical trials may not only be a new drug, but also a new diagnostic tool. A simple blood test may replace the x-ray or surgical biopsy as a way to gauge the spread or retreat of hard-to-detect ovarian cancer. Sound amazing? It's just one example of how new molecular information is transforming medical practice, moving doctors from surface to cause, from symptoms to underlying processes.

Patients will not be the only beneficiaries. Last year, the average approval time for new drugs rose to 18.5 months, up from 12.5 months in 1999. Much of that six-month lag is spent satisfying increasingly onerous, old-tech requirements like looking for subtle differences in crude x-rays. That's one of the things that doomed ImClone (See box, page 5)—the company did not collect enough x-rays to prove that patients who were started on its proposed new drug Erbitux had already failed to benefit from the existing standard cancer drug. (Meaning the company experimented with human beings it could not prove had received the best available treatment—a major FDA no-no). Better diagnostics translates not only into good marketing product—it also makes it easier, cheaper, and faster for companies to collect critical data on efficacy in clinical trials. Even saving a few months' time translates into a huge increase in profitability. Each year that can be cut from a clinical trial not only saves an average $100 million, but also gets products to the market faster. Being first to sell a new drug translates into gained market share and is often the difference between financial success and failure. Drug companies able to find protein markers to track the progress or retreat of diseases will have a huge competitive advantage.

GILDER BIOTECH REPORT

Editor	Scott Gottlieb, M.D.
Publisher	Lauryn Franzoni
Managing Editor	Maggie Gallagher
Designer	Julie Ward
Subscriber Services	Melissa McNally
Editorial Director	Richard Vigilante
President	Mark T. Ziebarth
Chairman	George Gilder

Gilder Biotech Report is published monthly by Gilder Publishing, LLC. Editorial and Business address: 291A Main Street, Great Barrington, MA 01230. Copyright 2002, Gilder Publishing, LLC. Editorial inquiries can be sent to: scott@gilder.com

To subscribe call 800.720.1112, e-mail us at biotech@gilder.com, or visit our website at www.gilderbiotech.com

Protein Messages

Why protein markers? Proteins carry out the nuts and bolts of managing bodily processes. Hormones are proteins and so are antibodies and enzymes. Proteins are the regulatory messages that cells send and receive, turning metabolism on and off, keeping our hearts beating, and transmitting instructions from our brains to our muscles and organs. In human physiology, the right protein is everything.

Normally, for example, to produce a protein like insulin, our body first scans for the gene that contains the code for manufacturing insulin and then copies it out from the DNA into an intermediate set of instructions, called *messenger RNA* (mRNA). The process of copying the gene is called *transcription*. Another set of molecules, *ribosomes*, uses the mRNA as templates to manufacture proteins.

Proteins are the "business end" of genes, the final products that carry out all the DNA instructions. So many pharmacogenomic researchers are not just mapping the genome, they are plumbing the proteome—concentrating on identifying modifications made to proteins manufactured from mRNA templates. DNA alone does a lousy job of predicting how much of a particular protein is produced. Even looking at mRNA is not enough: the correlation between mRNA and the amount of protein that is actually produced has been estimated to be less than 1 percent. Clearly, to get a picture of precisely what is happening in the body, researchers need a way to read the proteins themselves.

In the November 2001 issue of the *Gilder Biotech Report*, we outlined the emerging importance of cell signaling in new drug development and identified Human Genome Sciences [HGSI] as one of the outstanding companies that's bringing bioinformatics to the challenge of cracking these complex signaling cascades. Human diseases such as cancer, autoimmune disorders, and viral infections occur because of aberrations in the way proteins send signals to one another along these complicated cascades. Sometimes proteins fail to send a signal at all; other times they send too many signals. In many cancers, they seem to forget to shut themselves off after a brief message is sent. In this issue, we're introducing the wet lab and in silico tools that are used to identify these protein networks and the companies that best understand how these new tools will revolutionize the development of whole new classes of protein-based drugs and diagnostics.

Plumbing the Proteome

The field is called *proteomics*. Regulatory proteins do not exist in isolation. Much like the connection of transistors, resistors, and capacitors on a printed circuit board, cellular protein networks consist of protein interactions and pathways in which information is passed along as precise protein-to-protein interactions. Proteins coalesce into networks and circuits on commands from a stimulus. As stimuli fluc-

tuate and as feedback loops return information, newly formed protein networks are constantly changing.

The scientific goal of proteomics is to characterize this highly dynamic information flow. In the past, this is what we have called the "physics" of biology, and it represents the single largest information bottleneck limiting the capacity of in silico tools. Understanding the physics of cells—how proteins interact to mediate cellular functions—unlocks many of the roadblocks now frustrating the full application of computers to drug discovery.

In a diseased cell, a protein network is disrupted, deranged, or hyperactive. The cause is often in the genes. Genetic defects cause disease because the proteins are constructed according to faulty DNA codes and unable to maintain normal cellular functions. Many cancers, for example, occur when mutations activate so-called oncogenes that cause uncontrolled cell growth by signaling cell proliferation. Once we understand these protein-to-protein pathways and the genes that trigger them, we can design drugs that target highly specific points in a particular cell signaling cascade, along with new diagnostic tests that can be used to screen for the presence or absence of that particular abnormal protein.

One of the first examples of the new drug-diagnostic tango is the breast cancer drug Herceptin. The drug is effective only in the 25 percent of breast cancer patients who have a Her2/Neu protein receptor, which a new diagnostic test spots. The drug and the diagnostic market each other, as the new test identifies the subset of patients for whom the new drug will work. The next generation of protein-marker tests will monitor proteins to tell doctors whether or not a patient is responding to a new medication, when to change the dose, and when to try a new drug.

The first fruits of proteomics are already on the market, and others are in clinical testing. Current examples of new drugs targeting protein circuitry include the anti-cancer drugs Gleevac, developed by Novartis [NVS] for certain leukemias; Iressa, being developed by AstraZeneca [AZN] for head and neck cancer; and Tarceva, being developed by OSI Pharmaceuticals [OSIP] for lung cancer.

How much better will these new protein drugs be? Consider cancer, which accounts for one of every four deaths. Each year, 550,000 Americans—more than 1,500 a day—will die from it. Traditional chemotherapy poisons unhealthy and healthy cells alike, but these new classes of drugs target only cancer cells. Unlike conventional chemotherapy, which causes horrendous and deadly side effects, highly targeted drugs have little or no unintended consequences. No surprise, Gleevac racked up $36 million in sales in its first year on the market and is estimated soon to top $300 million.

Not just drugs, but new diagnostic tests will emerge as highly marketable products in themselves. Even today's crude protein markers (like my beloved alpha fetoprotein) form a lucrative market because of the inherent advantages for doctors and patients of simple blood tests over expensive, invasive biopsies—or worse, no diagnostic tests at all.

Our new knowledge of protein circuitry is giving rise to a new generation of much more specific, powerful, and cost-effective protein markers aimed at fresh molecular target areas. Today, the entire industry operates on about 500 protein drug targets. There are probably between 10,000 and 20,000 "druggable" protein targets just waiting to be uncovered. The proteome is rich, ripe territory, waiting to be mined for new drugs and diagnostics.

Fig. 1. From Genes for Proteins

DNA: Encodes the basic information to synthesize all proteins. Its information tells what could happen.	**RNA:** Is synthesized from DNA and is the intermediary in protein synthesis. Its information tells what might happen.	**Proteins:** Are synthesized from RNA and modified by other proteins in their maturation. They are direct mediators of most biology and drug effects. Their information tells what is happening.

Conceptually, the gap between companies that bill themselves as "genomic" outfits and the new "proteomic" companies is narrowing as everyone chases the same goal: find a protein target that can be manipulated to produce a medical benefit. Promising genomics companies like Human Genome Sciences and Millennium Pharmaceuticals [MLNM] are turning their efforts to detecting, identifying, and manufacturing proteins. A whole new proteomic tools industry is arising to serve their needs.

The trouble is that proteomic technology platforms aren't as well developed as genomic ones. Most processes for identifying regulatory proteins are still rooted in time-consuming and error-prone wet lab experiments. One reason? DNA is far simpler to digitize than proteomics. Gene sequences are a binary code, whereas protein data is still mostly of an analog sort, depending on the physics of the organisms (things like dissociation constants and kinetics), which are all still hard to measure and convert into digital data than can be analyzed and modeled in silico.

The genome is relatively static and essentially identical in every cell of an organism, but protein expression (the proteome) is in a state of dynamic flux, constantly changing and responding to both internal and external stimuli. Its dynamic flux makes it much more difficult to study, catalog, and analyze. For example, no one yet knows how proteins convert from long strings of amino acids into three-dimensional objects. Since the three-dimensional shape of a protein is central to its function, understanding how it takes shape (folds) is the key to understanding how it works.

MDS Proteomics

Enter MDS Proteomics, a technology leader in innovations that digitizes proteomic information. Proteomics, for example, requires uncovering subtle differences in the pro-

Table 1. Moving Medicine into the Proteomic Future.
Using proteins as diagnostic markers to predict disease and toxicity.

Field of Medicine	Therapeutic Application
Oncology molecular targets	Profiling of tumors for identification of molecular targets to determine treatment combinations
Monitoring for cardiac ischemia	Profiles of protein pathways for myocardial infarction diagnostics and prognostics and determination of drug-related cardiac toxicity
Diabetes	Development of proteomic profiles of insulin signaling circuitry to guide and monitor therapy
Infectious diseases	High-throughput, rapid antibiotic sensitivity profiling to see which drugs work and which don't
Immunology	Development of proteomic profiles of immune cell activation patterns for rapid evaluation of immune response to vaccines or to illness or injury
Toxicity monitoring	Development of proteomic "stress pathway" profiles from circulating immune cells as sentinels for impending toxicity

teins found in normal and diseased tissue. The first step, and one of the biggest challenges, is separating out proteins from blood and tissue samples. No conventional technique reliably accomplishes this separation, although many scientists say the closest is *two-dimensional gel electrophoresis*, which separates proteins based on their size and electrical charge. But on a mechanical level, gel electrophoresis is labor intensive, slow, and sloppy. It's the poster child for the drawbacks of wet lab science.

Why? Unlike DNA, proteins vary wildly in abundance in any given cell—by five or more orders of magnitude. Researchers must often hunt down a single molecule in an entire vial of blood. And when they find it, they still have only one molecule with which to experiment. Gene research relies on "polymerase change reaction," or PCR, to make millions of copies of DNA to facilitate lab work. But right now, there's no PCR for proteins—no way to amplify the proteins by making millions of copies. So proteins are not only hard-to-find once; they are a scarce resource even when identified. Gel electrophoresis can't detect anything but a small fraction of the critical, hard-to-find proteins involved in cell-to-cell signaling.

MDS Proteomics's solution? Use microfluidics to bring protein separation out of the crude, visible world into the nanoscale, where it can be automated and integrated into flexible platforms, digitized, and fed directly into machines for analysis. MDS Proteomics's technology platform has one critical advantage: it is based on searching for protein clusters rather than individual proteins.

When proteins send signals to one another, they tend to cluster close together in our bodies. So regulatory proteins involved in the complex signaling cascades are most likely to be found clustered near other proteins. A technology platform that can only detect single proteins is therefore much less likely to zero-in quickly on the key protein markers. Given how many proteins our bodies produce and how little we know about their complex functions, any technique that can narrow the arduous search for relevant regulatory proteins constitutes a huge competitive advantage.

Compare MDS Proteomics with two of the proteomic companies Wall Street likes: Oxford Glycosciences [OGSI] and Large Scale Biology [LGSC]. Both of the latter continue to rely heavily on 2D gel technology to identify regulatory proteins. Time will tell whether Large Scale and Oxford Glycosciences are able to discover the kinds of proteins that make good drug targets: we're not convinced 2D gel tech will even work at all for this purpose, much less work as efficiently and effectively as MDS Proteomics's cluster-based system.

After isolating the protein from the sample, the next challenge is figuring out what you have. Proteins come in all kinds of shapes and sizes, many of which haven't been iden-

VERTEX REDUX

When Vertex Pharmaceuticals [VRTX] announced that its experimental hepatitis C drug VX-497 yielded mixed results in a phase 2 trial, its share price dropped 13 percent in a single day. Strange, because VX-497 was never likely to be a blockbuster new drug—merely an adjunct therapy to the current drug of choice, alpha interferon. What would be a blockbuster advance for hepatitis C? Enter Vertex's newest drug candidate, which like the lucrative anti-HIV drugs, works by inhibiting a protease enzyme the virus needs to replicate. Now in preclinical stage, it has proven nontoxic in animal tests and could soon move to human trials. Wall Street values biotech companies based on the latest clinical trial. But companies like Vertex, which exploit in silico tools, produce continually bulging pipelines of promising new drug compounds. Vertex recently added a caspase inhibitor for sepsis (systematic infection). The company expects to bring at least four other drug candidates into development in 2002. Forget that one lousy trial that blew up earlier this year. And with Big Pharma on the prowl for new talent, Vertex could be worth bundles in a takeover. The company is a steal at $20.

BYE-BYE VENTER

When Celera Genomics's [CRA] CEO Craig Venter announced he was stepping down, the market beat down the company's stock 6 percent. Investors should have cheered. Venter was always ambivalent about making money for his investors, swayed by academicians who argued it was unnatural to lay claim to genes. But the company still sits on a mountain of intellectual property that could, with the right management, translate into real value. Check out Celera Diagnostics, a joint venture between Celera and its sister company.

tified yet. It's not as easy comparing the structures on the gel to those in your database. Researchers need an accurate way to sequence (or describe) the individual proteins found in each of these complex protein mixtures. The current answer is *high-throughput mass spectrometry* linked to powerful interpretation software.

Mass spec, as it's called, uses electricity to burn protein molecules, causing them to disintegrate and ionize. The resulting fragments are then analyzed based on the ratio of mass to charge to produce a molecular fingerprint. While the standard machines work, they are often slow and can sometimes confuse proteins that look similar.

Proteomics researchers are now competing furiously with variations in mass spec machines designed to increase speed and accuracy. In *MALDI-TOF mass spectrometry*, for example, a laser burns and ionizes a protein, and the ions then fly onto the surface of an electrode that's held about a meter away. By measuring the time it takes to fly to the detector very precisely, software can calculate the mass of the original protein.

MDS Proteomics is using MALDI-TOF as well as an even more innovative tool called *Fourier transform mass spectrometry* or FTMS. The technique uses a superconducting magnet that creates an ultra-high-resolution detector. The trade-off in mass spectrometry is usually sensitivity for resolution. The more proteins you spot, the less clearly you see them. But FTMS accomplishes both.

If you get the sense that there's a mass spec arms race underway, you are right. As Esther Dyson famously observed, many paradigms are changed by the tools that enable them. It was true of the genomics breakthroughs and of Celera Genomics [CRA]—the rise of both enabled by the high-speed sequencers supplied by Applied Biosystems [ABI] (later renamed Applera). In mass spectrometry, MDS Proteomics has continuously been ahead of the technology curve, giving it a huge advantage in figuring out the complicated networks of proteins that make up the signaling cascades.

Yeast Two-Hybrid Screening

MDS Proteomics showed off its proteomics expertise in the January 10 issue of *Nature*, reporting its first successful attempt to make a complete map of the intricate ways in which proteins work together in a yeast cell. (Human cells are next.) Mapping such protein interactions is important to identifying regulatory proteins because if two proteins interact, they usually participate in the same (or related) cellular functions. Most companies are generating maps of how proteins are relying on something called *yeast two-hybrid screening*. Companies using this technology to develop protein-interaction maps and databases include Myriad Genetics' [MYGN] ProNet database, CuraGen's [CRGN] PathCalling database, and Hybrigenic's PRIMRider.

Although there are several variations, the basic principle behind yeast two-hybrid screening consists of fusing a "bait" protein to one part of a transcription factor and a "prey" protein to another. To detect any interactions, one protein is genetically fused to a spot on the yeast cell (called *the DNA-binding domain*), while the others are fused to another part of the gene called *the gene expression activator*. If the two proteins don't interact, then nothing happens—there's no expression of the reporter gene. If the two proteins interact, then the two fragments are united and the activator gets turned on, usually producing an easily monitored color change. If the entire process sounds like one big wet beaker, it is. Shortcoming are numerous: a lot of false positives (around 25 percent) and false negatives (around 50 percent). Plus you still you need experimental confirmation of any potential hit. Proteins that exist in large assemblies (to form signaling cascades) are also difficult to separate out and analyze with this system; this is perhaps the most significant problem of all, since these are the kinds of linked proteins most likely to be involved in important disease-causing (or curing) pathways.

Then there is the inherent drawback of studying cells in unnaturally yeasty environments. Who knows if proteins will

By July, the company looks to complete four to six studies linking genes to disease, and the company claims it can go from gene to diagnostic test in just six months. Diagnostics represents a $1 billion market, growing at more than 25 percent a year. Just one hit could justify the company's entire market cap.

BIOTECH'S LITTLE ENRON

The saddest part of the ImClone debacle? If the company can survive the lawsuits (a pretty big if—), it looks like the drug Erbitux actually works, at least according to leading researchers. The FDA warned the company about problems in the design of its clinical trials as early as 1999, even while ImClone's jet-setting chief executive (he reportedly dated Martha Stewart and her daughter) swore to investors all was cool with the FDA and dumped his own holdings. Why did the FDA keep investors in the dark? Bizarre FDA rules allow companies to hide clinical information practically in perpetuity, while waiting for FDA final rulings. Something needs to change. The FDA could and should release data contained in a company's filings at each stage in the process, from initial filing through safety and efficacy phases of clinical trials. Better yet, the FDA could release the progress reports that companies doing FDA-sanctioned clinical trials must file annually. Why shouldn't markets and patients know what bureaucrats and insiders do? Not to let Little Enron execs off the hook, but the ImClone debacle also points to an antiquated honor system protected by a paternalistic federal agency out of step with the realities of the new medical information age. When someone claims certain information is just too confusing for you to know, generally that is a guarantee you ought to know it. B

behave the same way inside the human body? It remains to be seen. We believe fewer of the interactions uncovered by yeast two-hybrid screening are directly related to disease.

Worst of all, the most important cellular functions are less likely to be performed by individual proteins than by whole clusters of interacting proteins, and the yeast two-hybrid screening approach misses protein clusters entirely. But MDS Proteomics's PathMap process doesn't. One of the most distinct advantages of this proprietary PathMap process is that it identifies not only direct interactions between proteins, but also second- and third-generation links, allowing researchers to quickly interconnect entire cellular pathways.

Table 2. Companies Specializing in Experimental Protein Structure Determination

Company	Technology
Astex Technology	High throughput x-ray crystallography
MediChem Life Sciences	High throughput x-ray crystallography
Structural GenomiX	High throughput x-ray crystallography
Syrrx	High throughput x-ray crystallography
Integrative Proteomics	NMR
Structure Function Genomics	NMR
Signature Biosciences	Microwave and radiofrequency to probe protein structure

Bioinformatics Challenge

MDS Proteomics has also taken on the other lingering information challenge in proteomics: the development of the integrated computer architectures that can handle the massive data sets required to describe and analyze the very complex protein regulatory functions. To interpret the complex biological processes involved requires tremendous processing power. Right now it may be difficult or impossible, but as Moore's law marches on, these huge data sets can be compiled and crunched to establish the linkages between gene sequences, proteins, and body functions and dysfunctions. MDS Proteomics is investing heavily in processing power, envisioning the next generation of digital tools that will dominate proteomics.

MDS Proteomics has developed and industrialized new ways to look at human cells in action, gaining novel insights into disease processes and leading to faster drug development. This research capability will be shared with MDS Proteomics's partners and used internally by the company to identify 1,000 new drug targets over the next five years. In practical terms, this promises to help accelerate the drug discovery process, reduce failure rates of drugs in clinical development, and lead to improvements in the productivity of the pharmaceutical industry.

MDS is working to generate mathematical models that will predict biological processes, a monstrous bioinformatics problem. To begin with, all the data generated will need compiling. Celera Genomics has more than 3 terabytes of hard disk space filled, and that's just holding its DNA data. Protein data is going to be orders of magnitude greater in terms of quantity of data and heterogeneity. Software, too, is lacking. There aren't any commercially available information management systems for protein-based information, so MDS Proteomics is developing its own programs.

MDS Proteomics is also set apart from its competitors through its investment in the downstream technology that will turn proteomic hits into drugs. A protein's biochemical function is largely determined by its three-dimensional shape—its structure is stippled with pockets and grooves into which other molecules fit as precisely as a key into a lock. Researchers obtain information about the three-dimensional structure of a protein through two approaches: experimentally and computationally. In the lab, a protein's shape can be determined using nuclear magnetic resonance (NMR) spectroscopy or x-ray crystallography (the private San-Diego based biotechnology companies Syrrx and Structural GenomiX excel in these techniques). New computational techniques for identifying protein structures are a topic we'll be taking up in detail in an upcoming issue.

MDS Proteomics is developing expertise in experimental structure determination, but is also taking advantage of its computing capabilities and its growing trove of information to make structural predictions based on in silico calculations—a promising space occupied by only a handful of other mostly private biotechnology firms such as GeneFormatics, DeNovo Pharmaceuticals, Structural Bioinformatics [SBI], and Inpharmatica.

The rapid improvement of in silico methods, together with advances in computational speed, will not only accelerate the transition from newly determined proteins to drugs, but also allow biodigital advances already associated with DNA analyses (genomics) to be transferred to proteins (proteomics). Structural pictures of key proteins can be digitized, and software tools called docking programs can be used to test different chemical structures against their binding sites, looking for molecules that are snug fits. Computers linked to massive databases play a central role in this process, instructing chemists which small changes will turn test compounds into better drugs, diminishing toxicity or making them more easily absorbed. Libraries of millions of drug compounds can be rapidly screened against targets and refined in a highly iterative fashion.

Best of all, computational proteomics has a learning curve. The more protein structures you're able to identify and catalog using wet lab techniques, the smarter the computational predictive software becomes. No single experimental or computational approach is likely to result in accurate and complete models of all proteins, protein complexes, and pathways: the greatest advances will be made through the integration of computational methods with physical data obtained through x-ray crystallography, NMR, and other experimental results. MDS Proteomics is investing in each of these tools.

To win the proteomic arms race, a company needs expertise

in both computational and experimental techniques. Expect many companies to broaden their current technology platforms through consolidation with companies that have expertise in these other aspects of functional genomics. San Diego-based GeneFormatics, a private company with expertise in homology modeling and ab initio structure determination, recently purchased another private company, Structure Function Genomics, to add to its computational capabilities. Structural GenomiX, a private company with expertise in x-ray crystallography, added the private company Prospect Genomics to gain access to that company's homology modeling and in silico Ligand docking technologies. MDS Proteomics is investing in both experimental and computational tools for structure determination, even as it remains poised for future dealmaking. MDS Proteomics hasn't advanced any of its drug candidates into clinical development, so assigning a value to its potential product pipeline is premature. But the potential payoff of a protein-based approach to drug discovery is the ability to identify entirely new drug targets that previous technology was unable to spot, much less take advantage of. Only in the past two years has the right mix of tools, technologies, and the critical mass of high quality genomic and proteomic data become available to break through barriers in the drug discovery.

Just in the nick of time, too. Over the next three years, Big Pharma will be faced with an average 1 percent to 8 percent revenue gap between the growth from pipeline products and the effect of patent expirations. Currently, a typical top-tier pharmaceutical company produces on average only 0.5 to 1.0 new chemical entities each year. Using traditional methods, it takes six to twelve years and up to $800 million to transition from target discovery to clinical development. In order to sustain the industry's historical earnings' growth rate of approximately 10 percent, pharmaceutical companies need to launch three to five new drug candidates annually, each with a sales potential of at least $300 million per year.

To launch three novel drugs annually, drug companies must discover more than two hundred new targets each year. The pharmaceutical companies cannot achieve that productivity without the new biodigital tools—tools like those developed by MDS Proteomics for rapidly identifying and characterizing novel drug targets.

While MDS Proteomics is using its proprietary protein mapping tools to develop its own drugs, we believe these information tools are also valuable assets in their own right. They can be used to develop partnerships with pharmaceutical companies or sold on a subscription basis as an independent drug discovery tool.

Right now, Big Pharma is flying blind. Ten years ago, the average number of literature references per target under consideration at a pharmaceutical company was more than one hundred; now it is eight. That means pharmaceutical companies may have four or five times as many targets under development as they did just five years ago, but most of those targets are new and poorly understood. The proteomic maps being generated by companies like MDS Proteomics are valuable tools to help plug those knowledge gaps. Evaluating hundreds of such protein targets for pharmaceutical companies (at say $1 million to $5 million a pop) is potentially a big business in itself, not to mention the potential for royalties in joint development deals. This kind of intellectual property will be even more valuable in proteomics than genomics, because regulatory proteins are a step closer to being targets for new drugs than DNA sequences: you can ask for more money by turning up information about them.

So how much is MDS Proteomics worth? CuraGen [CRGN] recently struck a deal with Bayer [BAY], in which the two parties agreed jointly to fund research, development, and commercialization activities up to $1.34 billion over a 15-year period. The agreement includes an $85 million equity investment in CuraGen by Bayer and $39 million in committed funding to CuraGen. MDS Proteomics, in comparison, is just now entering a phase of looking for serious partnerships to drive its development and could be in line for a similar deal (of course, after the fallout over ImClone settles down and Big Pharma gets back on the prowl for partnerships). In August 2001, MDS Proteomics struck a deal with antibody maker Abgenix, Inc. [ABGX], in which Abgenix agreed to make a $15 million investment in MDS Proteomics. Together the two parties will develop antibody drugs from among 150 novel targets that MDS Proteomics has helped provide. The company is positioned to develop additional novel targets that can be leveraged in new collaborations with pharmaceutical partners as well as build out its own internal pipeline.

Table 3. Companies Specializing in Computational Modeling of Protein Structure

Company	Technology
DeNovo Pharmaceuticals	Homology modeling
IBM (Blue Gene Project)	Computational protein folding
GeneFormatics	Ab initio modeling
Inpharmatica	Bioinformatics
Locus Discovery	Computational chemogenomics
Prospect Genomics	Homology modeling
Structural Bioinformatics	Homology modeling

MDS Proteomics recently received an equity investment from IBM that valued the company at $33 a share and put its total equity value in the neighborhood of $800 million. We suspect the special computing and analytical power needs of companies like MDS Proteomics will lead to more such strategic collaborations or consolidations within the biotechnology and computing companies.

For his part, MDS Proteomics's CEO Frank Gleeson told me he'd like to compare his company to Millennium Pharmaceuticals [MLNM], which has transitioned itself from a gene discovery outfit to a fully integrated pharmaceutical company and generated $5 billion in equity wealth in just five

BIOTECH COMPANIES

Company	Technology Leadership	Reference Date	Reference Price	1/31/02 Price	52-Week Range	Market Cap
Vertex (VRTX)	Rational Drug Design	9/17/01	28.60	19.74	15.50 - 74.75	1.48B
Human Genome Sciences (HGSI)	Cellular Signaling	10/26/01	43.97	28.13	26.41 - 77.00	3.59B
Nanogen (NGEN)	BioChips	10/2/01	4.95	5.33	3.00 - 13.44	114.9M
Gilead Sciences (GILD)	Rational Drug Design	12/05/01	67.72	65.42	26.88 - 73.67	6.28B
Quorex (none*)	Rational Drug Design	12/05/01				
Sequenom (SQNM)	Pharmacogenomics	1/09/02	9.00	6.84	5.65 - 21.25	256.0M
MDS Proteomics (none*)	Proteomics	2/05/02				

* Pre-IPO startup companies.

NOTE: This list of Gilder Biotech companies is not a model portfolio. It is a list of technologies in the Gilder biotech paradigm and of companies that lead in their applications. Companies appear on this list only for their technology leadership, without consideration of their current share price or the appropriate timing of an investment decision. The presence of a company on the list is not a recommendation to buy shares at the current price. Reference Price is the company's closing share price on the Reference Date, the day the company was added to the table, typically the last trading day of the month prior to publication. The author and other Gilder Publishing, LLC staff may hold positions in some or all of the companies listed or discussed in the issue.

years. Wouldn't every biotech CEO? But MDS Proteomics is making the same kind of transition: exploiting the power of in silico techniques to make the leap from a proteomics company to a full-range pharmaceutical concern. And MDS Proteomics has one significant advantage: plumbing the proteome is likely to be even more lucrative than mapping the genome, as the leap from protein to drug is much faster and easier than from gene to drug.

MDS Proteomics is currently a unit of MDS Inc. [MDZ] of Toronto, which is reported to have an 85 percent stake. The biotech unit shelved an IPO last year after the high-tech market meltdown, but it recently announced plans to launch an IPO to raise about $200 million in the second half of this year. Those tempted to buy MDS Inc. in the meantime—in order to get a piece of MDS Proteomics—will have to contend with the margin drag of MDS Inc.'s old economy, healthcare service businesses. You'd be better off waiting for the IPO.

Wall Street likes to compare MDS Proteomics to privately held GeneProt, Large Scale Biology [LSBC], and Oxford Glycosciences [OGSI], mainly because all three have developed automated systems for isolating and identifying proteins. Both Oxford Glycosciences and Large Scale Biology work with individual companies to build customized proteomic databases for solving particular problems. For example, Oxford Glycosciences is collaborating with Pfizer [PFE] to identify drug targets and disease markers for Alzheimer's disease and atherosclerosis. All three are pursuing targets in collaboration with partners as well as on their own, but their major thrust is marketing solutions to specific problems.

MDS Proteomics is in a class by itself with its investment in the tools of in silico drug design, using its growing trove of proteomic information to create one seamless platform for in silico drug discovery—from protein discovery to structure determination to computerized drug design. What sets MDS Proteomics apart is its investment in the next generation of in silico tools that we believe will solve the physics of protein interactions as well as in silico drug discovery technology like experimental and computational structure determination and in silico structure-based drug design. Even while using today's generation of technology to gather proteomic information, MDS Proteomics is plugging that information into systems capable of harnessing it on a computational platform that accelerates at the pace of Moore's law.

Scott Gottlieb
February 5, 2002

References

Anne-Claude, G., et al. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415:141-147.

Fields, S. 2001. Proteomics in genomeland. *Science* 291:1221-1224.

Hunter, T. 1995. Protein Kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* 80:225-236.

Schwikowski, B, et al. 2001. A network of protein-protein interactions in yeast. *Nature Biotechnology* 18:1257-1261.

Stewart, L. 2001. High-throughput crystallization and x-ray structure determination of protein-ligand complexes. *Drug Discovery by Design.* IBC.

Yoshida, M, et al. 2001. Proteomics as a tool in the pharmaceutical drug design process. *Current Opinions in Pharmaceutical Design* 7:291-310.