by Jeff Mortimer

Studying Human Proteins Opens New Worlds of Diagnostic and Therapeutic Possibilities

BEYOND THE

Photo: Martin Vloet

Sam Hanash and his U-M Colleagues Face Daunting Challenges in Shaping Tomorrow's Advances in Medicine

Samir M. Hanash, M.D., (Ph.D. 1976), used to knock on other people's doors. Now they knock on his. They also call, e-mail, fax and send letters to him. Once a voice crying in the wilderness, there are times when Hanash almost feels like crying out for the voices coveting his time to stop.

Hanash is a professor of pediatrics and communicable diseases at the University of Michigan Medical School, an attending physician in pediatric hematology/oncology at University Hospitals, and director of several major research programs in proteomics and related fields that have bloomed in the biological sciences' so-called "post-genome era." He is, in fact, one of only a handful of people in the world with sufficient credentials in those areas to be considered an authority on them, and the advances in medicine that they will shape.

Although still in its relatively primitive stages — the word itself was coined only seven years ago — proteomics research has already led to better detection methods and more effective treatments for several kinds of cancer. Not surprisingly, "Proteomics is attracting substantial interest, not only on the part of the NIH but also on the part of the private sector," says Hanash. "Currently, proteomics can be viewed as a multi-billion dollar industry."

Proteomics is the study of an organism's proteins as a whole in order to better understand not only how they work (or not) individually but also what they do; the functions of most of them are still a mystery. Proteins perform the tasks determined in DNA and transmitted by RNA. "They are the functional component encoded for in the genome," says Hanash. A frequently used analogy likens the genes to blueprints and the proteins to buildings, but the imagery is far more static than the reality. Compared to exploring the body's proteins, the mapping of the human genome was a snap, he says.

"Deciphering the proteins in the human body is vastly more complicated than sequencing the human genome, because you can start at the beginning of

the human genome, chromosome one, and go to the end," he says. "It's a linear process. There's no linearity in the deciphering of protein information. The proteins can change, and that's what makes their study even more daunting. Each gene, in principle, makes an RNA, but that RNA can be translated into proteins which get modified so that you have potentially tens of different proteins being made out of it. It adds a lot of complexity."

When Hanash first became interested in studying proteins en masse, proteomics not only didn't have a name but barely existed as a concept. His primary focus, then as now, was childhood cancer, and he earned an M.D. from the American University of Beirut, Lebanon, in 1972 and a Ph.D. in human genetics from U-M four years later.

"I wasn't satisfied with the amount of science behind clinical practice," Hanash says. "It was clear that there was much more that needed to be discovered, and at that time I was very attracted to human genetics. In the early '70s, I attended a congress on human genetics in Paris, and in talking to different people it was clear that Michigan had one of the best programs, so I decided to come here to get a Ph.D. One thing led to another, and I ended up staying here. It amazes me. I never thought I would be here for more than just finishing my Ph.D., but the opportunities basically were there at every turning point."

One, of course, was research. "I made a decision to move from the study of one molecule, which in my case was hemoglobin, to the analysis of protein expression programs in the cell, which meant having to analyze hundreds of proteins simultaneously," he recalls. "What got into me is that there is a need to understand programs of gene expression that you cannot understand by looking at one molecule. You need batteries of molecules to see how a trend in gene expression might emerge. More importantly, most of the proteins were unknown, and so by analyzing hundreds of proteins, the hope was that we would identify new proteins associated with many different types of cell functions, from cell proliferation to differentiation. In my particular case, I was very much interested in identifying proteins involved in cancer, such as novel cancer markers and proteins that are messed up in different types of cancer.

"At that time, this was going against the grain," he says. "Scientists were interested in going in depth in their studies of any given molecule, and to propose to simultaneously analyze hundreds of molecules was just not fashionable whatsoever."

Nonetheless, he kept knocking on doors, painstakingly raising money to develop the tools he needed — "which were just not there" — to answer his questions. "This represented quite a struggle," he says. "Most of the interest in the early 1980s was in genomics and developing DNA technologies, as well as in doing the sequencing of the human genome. Although it was very difficult to obtain the necessary funding, we did manage during that difficult period to convince the NIH to provide some of the needed resources."

Collaborations Lead to Industry Standards

The standard tool for studying proteins "was not enough to do large-scale analysis," he says. "What was needed was the development, on the one hand, of computer tools to analyze such complex images and, on the other hand, tools for protein identification, because a lot of the proteins are present in the tiniest amount. Then we needed to develop procedures that allowed us to extract that tiny amount of protein and to identify the nature of each one of those hundreds of proteins."

> PROTEOMICS IS THE STUDY OF AN ORGANISM'S PROTEINS AS A WHOLE IN ORDER TO BETTER UNDERSTAND NOT ONLY HOW THEY WORK (OR NOT) INDIVIDUALLY BUT ALSO WHAT THEY DO; THE FUNCTIONS OF MOST OF THEM ARE STILL A MYSTERY.

Working with computer scientists, Hanash and his colleagues created software for image analysis. Working with LKB Sciences (later LKB Pharmacia), a private company, they developed the next generation of two-dimensional separation technology, called IPG (for "immobilized pH gradients"), which became the worldwide standard. Working with scientists at Michigan State University, they linked protein separation with protein identification through the use of mass spectrometry, which has also become the " industry standard" in proteomics.

Now, he says, "the struggles of the past are bearing fruit. Attention is shifting to the functional analysis of the genome, and the proteins are the most functional component encoded for. It used to be perceived in a negative fashion, that Sam Hanash is studying hundreds of proteins and where is that going? All of a sudden, it has become somewhat of a requirement, in the post-genome era, to be able to utilize technologies that can capture thousands of genes and thousands of proteins simultaneously." A cancer proteomics symposium he participated in at the annual meeting of the American Association for Cancer Research this April in New Orleans attracted some 2000 people.

Once captured, the secrets they so grudgingly yield open up new worlds of diagnostic and therapeutic possibilities. "A lot of our activities now have ➤

to do with integrating information from the genomic level to the RNA level to the protein level," he says. "This type of approach is currently being funded to allow the identification of new markers for cancer, which would lead to earlier diagnosis as well as screening on a large scale, and the development of novel classification schemes for cancer, with the idea that if we can characterize a cancer through the analysis of the thousands of proteins and the thousands of genes expressed in it, we can much better define how that cancer could respond to different therapies that are currently available, make the most appropriate choice of therapy, and at the same time be able to identify novel pathways and molecules involved in cancer, which may lead to new therapies." Hanash and his team have obtained substantial funding from the National Cancer Institute to implement this approach for lung, colon and ovarian cancer.

One example is the work of David Beer, Ph.D., a professor of surgery and radiation oncology who is in charge of the lung tumor component, funded by the National Cancer Institute, of the molecular classification of tumors headed by Hanash. "He has been able to identify, with the application of this technology, two subgroups of a certain type of lung cancer that have very different clinical behaviors and outcomes," says Hanash. "This is incredibly interesting because you could use this technology to tailor the treatment based on the profiles that you have seen and say this person needs more aggressive treatment and that person may require less. We have also seen similar kinds of differences in profile between patients with brain tumors. We're very excited about that."

> "WE'VE BEEN ABLE TO IDENTIFY A LARGE NUMBER OF GENES AND ALSO SOME SPECIFIC PROTEINS WHICH SEEM TO DISTINGUISH TUMORS WHICH HAVE A BAD PROGNOSIS FROM ONES THAT HAVE A GOOD PROGNOSIS."
>
> — DAVID BEER

Hanash describes this process as "mining the genes that tell us something about cancer," but the same picks and shovels can be used to attack other diseases. "I'm helping set up a consortium of investigators who want to use the very same technologies we are using for cancer to shed light on cystic fibrosis, a children's disease," he says. "I've also worked with a group headed by Ron Koenig, professor of internal medicine at U-M on a proposal to the NIH to use a similar strategy for diabetes, kidney disease and digestive disorders which got funded."

Photo: D.C. Goings

He and his colleagues are not alone in their excitement. "Funding in the area of proteomics is likely to skyrocket in the next few years," he says. "Right now, we are inundated with offers of different kinds, ranging from private sector groups wanting to simply fund part of the operation in the lab through collaborative agreements, to others who want to spin off biotechnology companies."

Will Michigan Maintain Its Pioneering Lead?

While U-M had a head start, thanks in part to Hanash's work, it is by no means clear that it will be at or near the top of the heap when the dust settles. "There's a lot of jockeying going on," he says. "In the Midwest, just about all the Big Ten institutions have plans to develop life science initiatives and initiatives in the post-genome era in general, from proteomics to bioengineering to bioinformatics. It would be fair to say that who the leaders are going to be is still up for grabs. It's going to depend on institutions putting in a lot of resources, developing a lot of talent and developing structures that allow much more integration between the physical sciences and the medical sciences than has ever been done before."

It's also, in his view, going to depend on stronger ties between the academic and corporate arenas. "There isn't a strong tradition at Midwest institutions of substantial partnerships with the private sector," says Hanash. "Those types of dealing seem to gravitate to the East Coast and the West Coast. It's been very difficult to try to do that here, and it's going to take more than getting NIH funding. It's going to take building the infrastructure that can deal not just with the traditional academic venues but also with the corporate world, and developing effective strategies to move the research from the discovery phase to the translation phase."

"...it's going to depend on... developing structures that allow much more integration between the physical sciences and the medical sciences than has ever been done before." — Sam Hanash

The ferment in Hanash's life mirrors the ferment in the field. On one recent trip, he traveled from Ann Arbor to Boston to meet with a company there for two hours, flew from there to Tampa for a meeting organized by the National Cancer Institute, from there to Japan to give a talk at an international meeting, from there to San Diego to give a talk at the Experimental Biology 2000 meeting, from there to Paris to meet with collaborating investigators on a project, from there to Washington for another NCI workshop, and then back to Ann Arbor, all in the space of 10 days.

The irony of his newfound status is not lost on him. "You bang on so many doors and there's no response and now all you have to do is sit back and wait for all the offers to come," he says. "It's quite a dramatic change. This is something I would only have dreamed of 10 years ago, but I'm realizing that with success, you have new challenges. How to manage the success is very important. How to maintain your sanity through all of this is very important. And how to please everybody is incredibly challenging, because all of a sudden you find yourself having to say 'no' to a lot of people."

Always "Yes" to Children's Research

He always says "yes" to children, however. Maintaining both a practice and several major research programs, in addition to all the other activities that keep him in the air so much of the time, is more than a bit of a stretch, but Hanash still manages to keep his eyes on the prize. It's not just his relatively sudden cachet that's ironic, but the fact that it's come to a specialist in pediatrics, not usually regarded as the locus of high-profile research.

"I've been doing research on childhood diseases for 20-some years now," he says. "The patients that we take care of are not great advocates for the types of disorders that they have, so they are dependent on others being the advocates for them. They don't have the equivalent of a Norman Schwarzkopf having prostate cancer and then going on TV and saying, 'I want you to support prostate cancer research.' From a public health point of view, childhood diseases have not amounted to a whole lot."

Being hitched to Hanash's star might not change that, but it can't hurt. It has been both his great gift and his good fortune to focus on work that "travels" well and is thus likely to attract resources. As he says, "When you apply for funding to do research on a rare childhood disease, as opposed to a common disease that represents a major public health problem, the science could be the same but the impact is different."

Ah, yes, the science. Through it all, the lean times and the prosperous, the triumphs and disappointments, his passion for the science has sustained him, and still does.

"There's nothing that's more exciting than the discovery process itself," he says. "The most exquisite aspect of it is the eureka part, being in front of the machine or talking to the post-doc or getting the results printed and having it hit you in the face, 'Aha, it looks like I may have found a marker for lung cancer.' This is something between you and yourself. Nobody else is passing judgment on it, nobody else is elevating it to some other level. It's that one short moment when the neurons get fired up and you see something."

It's a little like infatuation, and it doesn't last much longer than that condition, either. "Once you make this initial discovery, however modest or however exciting it is, turning it into something that is practically useful can be an extremely elusive goal," he says. "You have to apply for funds, you have to compete with others, you have to demonstrate it in the form of something that goes into the literature, then you have to design other kinds of experiments. All of a sudden, what seemed like a neat idea becomes a bureaucratic process. You may be very excited about it, but somebody else who doesn't have your line of work dear to their heart may be totally indifferent to it, and you have to convince those who are indifferent what you're excited about and try to get them excited. The downside is the drudgery of it all."

Although still in its relatively primitive stages — the word itself was coined only seven years ago — proteomics research has already led to better detection methods and more effective treatments for several kinds of cancer.

But another upside is how much remains to be learned, more than enough to keep Hanash supplied with "eureka moments" for the rest of his career. All those proteins that he tracks down, identifies and analyzes "are not proceeding either independently of each other or dependently in just one direction," he says. "Information does go from DNA to RNA to protein, but information goes back from protein to DNA that affects RNA. It becomes a very complicated circuit. By the end of the 21st century, we'll still have a lot to learn about how the circuit is connected." ➤

Cancer Biomarkers, Molecular Profiles May Improve Chances for Early Diagnosis and Treatment

Nana Lee is a researcher with a mission. Like a gunslinger in an old Western movie who's tracking the man who shot his pa, Lee is after the disease that claimed her mother's life in 1997.

"It's not just science but also a personal project," says Lee, a post-doctoral fellow in internal medicine who is part of the colon cancer profiling research team in Eric Fearon's laboratory. "It's something that drives me with my work."

"They took out her whole colon but they didn't know, they couldn't tell me, if it might have metastasized," she recalls. "It was just a big question mark. They said it might have spread to the lymph nodes, but they weren't sure; it was tricky to diagnose. Eleven months later, they said it had spread to her liver. She died a couple of months after that. I knew that if we had had the tools to diagnose if it had metastasized sooner, things would have been a lot different. Now one of my goals in life is to come up with something like that."

As Lee was finishing her Ph.D. in biochemistry at the University of Toronto, she learned about the colon molecular profiling project at Michigan, headed by Eric Fearon, M.D., Ph.D., and Stephen Gruber, M.D, Ph.D., "and I just jumped on that. I came here last September because they had a great program and I wanted to study genomics and proteomics of colon cancer and help find a way to use this information to improve clinical practice."

LEE FEARON CHO GRUBER

> "Instead of 1,000 samples and five variables, we have 100 samples and 7,000 variables. This changes the game in fundamental ways. Old methodologies don't apply."
>
> —George Michailidis

While not as personal, similar passions inflame other researchers working on the cancer molecular profiling project. Kathleen R. Cho, M.D., the principal investigator of the ovarian profiling cancer project and an associate professor of pathology and internal medicine, is a practicing surgical pathologist with a subspecialty expertise in gynecological cancer diagnosis. "I spend much of my life dealing with these kinds of tumors," she says. "It's exciting for me to get beyond the microscopic appearance of these tumors and begin to evaluate their molecular profiles. I'm optimistic it's going to take us way beyond what we've been able to do with just microscopic appearance alone."

For Cho, too, early diagnosis is the holy grail. "It's been a huge problem [with ovarian cancer]," she says. "Tumors tend to present very late in the clinical course, with high-stage disease. If they present with low-stage disease, they can often be cured with surgery alone or surgery and chemotherapy, but if they present with high-stage, they're very difficult to cure with the treatment modalities we have currently.

"The job now is to ferret out, among the hundreds of proteins and thousands of genes we're looking at, which of those individual molecules may be the most predictive in determining a particular tumor's biological behavior," says Cho. "If we can identify genes that are very highly expressed specifically in ovarian cancer that might be, for example, secreted into the blood, we could develop a simple blood test to enhance early diagnosis. That would be the long-term goal, to develop screening tools that would allow you to identify patients with low-stage ovarian cancer by doing some sort of simple, minimally invasive test, like a pap smear for uterine cancer, that's convenient and not frightening or costly."

Good Prognosis, Bad Prognosis: Genes and Proteins Tell the Story

The lung cancer profiling team is headed by David Beer, Ph.D., associate professor of surgery and radiation oncology.

"We're trying to determine whether we can identify the genes and the proteins which are associated with a poor clinical outcome in early-stage lung cancer," says Beer. "Most of the patients who have stage-one lung cancer will do well, but about 25 percent will have a poor clinical outcome."

So far, so good, says Beer. "We've been able to identify a large number of genes and also some specific proteins which seem to distinguish tumors which have a bad prognosis from ones that have a good prognosis," he says. "The proteomics studies are aimed at trying to distinguish not only patient prognosis but also the types of proteins which are potentially unique or highly expressed in tumors with different clinical features, such as invasive characteristics, and also the proteins that are associated with the ability to metastasize to local lymph nodes."

by Jeff Mortimer

In an enterprise energized by connections and collaboration, the fields of computer technology, statistics, medicine and the physical sciences converge on research into several kinds of cancer

For example, he says, "You could look at two different tumors under the microscope and you can't tell the difference between them, but the genes that are expressed and the proteins that are encoded by those genes are expressed differently. Careful quantitation of the levels of those proteins, as well as their identification, gives us a tool to try to distinguish tumors in a way that you just can't determine by looking at the tumor histologically."

The day when such tools are clinically available may not be far off. "It will probably start showing up fairly soon," Beer says. "We've identified those genes and proteins which are of interest and may be useful. The next step will be to bring this to the clinic to test it in a prospective manner by taking many more early-stage tumor patients and repeating this study that we've just done with nearly a hundred and see if the markers that we've identified truly do define their prognosis."

KUICK TAYLOR KARDIA MICHAILIDIS

A New Way of Doing the Numbers

The "careful quantitation" to which Beer refers is founded on the work of a team of statisticians and image analysts headed by Jeremy Taylor, Ph.D., professor of biostatistics in the U-M School of Public Health, and including Sharon Kardia, Ph.D., assistant professor of epidemiology there; George Michailidis, Ph.D., assistant professor of statistics; Kerby Shedden, assistant professor of statistics; and Rork Kuick, M.A., a statistician who has worked on genomic and proteomic quantitative analysis for a number of years.

"We barely had machines to do this work back then, of course," says Kuick. "It was only at the very end of the 1980s that computer windowing systems permitted spot detection and quantification algorithms to allow a user to work with many, many images and get this work done. To make even marker discovery work, there's a lot of image analysis involved. The images have to be obtained and spots measured with algorithms and spots between different patterns matched with algorithms, so just obtaining the data is a bit of an engineering and computer science task."

The field of statistics, as well as medicine and science, has benefited from the challenges involved. "Some of these new technologies require new methodologies in order to look at the data and extract the most out of them," says Michailidis. "What you would like to do is make these comparisons and see which genes have changed significantly. One of the important statistical calculations is to quantify precisely what you mean by significant difference. Say a normal tumor sample is 1 and we find one whose expression level is 1.1. Is that a significant expression? If it went from 1 to 15, you'd say this is really large, but from 1 to 1.1, is it such a big difference or not? That's where statistics come in, to see if differential expression is really there."

Moreover, he adds, "The data are fairly noisy. It's a complicated technology and there are lots of sources of errors, so you need to correct for as many of these errors as possible to see if you really have a significant difference or not. If you have, then you go to the next stage and try to learn more things about these changes."

Project statisticians are also working with a relatively small number of samples, due to the cost and difficulty of obtaining them. Michailidis says this "leads to a totally new paradigm in statistics. Usually what we have are fairly large samples and a few variables; this is the other way around from what we are used to. Instead of 1,000 samples and five variables, we have 100 samples and 7,000 variables. This changes the game in fundamental ways. Old methodologies don't apply. It becomes fairly exciting because there is a lot of room for new ideas."

The entire biomarker enterprise is energized by connections and collaborations, not only among medical specialties but also between medicine and other disciplines. The project's individual actors are intrigued by the whole that their parts comprise. "To be honest, I don't understand the underlying science that well," says Michailidis, "but I'm brushing up on my biology."

Says Cho, "I'm absolutely turned on by this. It's been a great opportunity to collaborate with other investigators doing similar work with other tumor systems."

The mission that unites them is finding out more. As Beer says, "The more we learn, the more new targets we can potentially uncover for both diagnosis and therapy."