A conversation with Thierry Rabilloud, Grenoble

"Stay Focused!"

Thierry Rabilloud has worked for more than 20 years in protein research. In this *Lab Times* interview he talks about what proteomics has really achieved so far, about the difficulties the field is facing due to overblown expectations and about the 'dirty' homework proteomics researchers still have to do in order to be successful.

Protein research boomed spectacularly when proteomics entered the arena in the late 90s. Millions of euros, dollars and yen were invested into research; new companies were founded; even big industry climbed onto the bandwagon. But what came out of these formidable efforts to map the proteins of whole cells and even organisms? Plenty of overblown headlines but hardly any excellent scientific publications. Proteomics was forced to face up to serious problems, including protein complexity and bad reproducibility of results.

No wonder that Myriad Proteomics, having announced only in 2001 its intention to analyse the complete human proteome within two years, with the help of Oracle and Hitachi, shut down its proteomics facilities in 2004. The Swiss company GeneProt, established in Geneva in 2001 as the first big industrial proteomics company, also closed. Though 50 mass spectrometers had combed human blood plasma in the search for disease markers, output was close to zero.

In its latest issue from June 2009 *Nature Methods* published an excoriating article, stating that, "27 laboratories, stuffed with proteomics specialists, analyzed a standard sample of 20 recombinant human proteins – and only 7 labs were able to report all 20 proteins correctly" (vol. 6, 423-30).

So what's going on in proteomics? We talked to Thierry Rabilloud, who worked for more than 20 years in protein research. He is now leading the Proteomic and Cell Biology group at the CEA Life Sciences Division in Grenoble.

LT: 'Proteomics: one small step for a digital computer, one giant leap for humankind' was the headline of one of the many articles that have promoted proteomics since the late 90s. Now the hype is over and industrial proteome research has been shut down almost completely.

Thierry Rabilloud: Right, the hype is over. The companies that had stepped into proteomics realised that, for years and years, proteomics was grossly oversold by many people, especially in the academic field, and that there's not enough bang for the bucks. And you know how companies are: they close everything and move to something else; that is metabolomics.

"In proteomics we are close to ten percent of the power of genomics or transcriptomics."

Why was proteomics oversold?

Thierry Rabilloud: There are psychological reasons that root very deeply in the history of proteomics, and there are some technical and practical reasons that scientists were not really aware of at the beginning of proteomics. Let's first talk about those practical reasons. People in proteomics didn't realise that

they are not comprehensive at all. However, genomics or transcriptomics are comprehensive; arrays comprising of whole genomes and genome wide transcripts have been produced. They are not perfect but they work quite nicely. In proteomics we are close to ten percent of the power of genomics or transcriptomics. And there's nothing in the foreseeable future that tells us that one day or another we can achieve our optimistic goals.

One of which was an atlas of the whole human proteome.

Thierry Rabilloud: For example. The protein atlas was brought up by Norman G. Anderson and his son Leigh already in the 70s, I think. But in my opinion we will have no protein atlases ever because an atlas needs to be comprehensive – and we know that we will not be comprehensive. Let me give you an example. Two years ago we worked on the bacteriophage T4. This little bug has 37 proteins. We set up a kind of little competition between a few groups: all tried to map the whole proteome of T4.

That doesn't sound so difficult.

Thierry Rabilloud: Not difficult? Well, none of us was able to see all of the 37 proteins, not by any method! Typically, participants identified 33 proteins. They lacked 10 percent. So did the group of John Yates from the Scripps Research Institute. On the other hand, Yates' group identified the 75 proteins of ribosomes and they got all proteins rather easily. What makes the difference? All proteins in ribosomes are in a 1 to 1 to 1 ratio. However, the phages' proteins range between 3 to 900 molecules per phage, meaning, the dynamic range of expression describing the range between the rarest and the most abundant proteins is 1 to 300. Now that we failed by two orders of magnitude in our work with the little phage how can we dare to think we could describe the proteome of a human, where the dynamic range of expression of proteins within the cell lies between one and one million?

What about the 'sub-proteome projects', namely the human brain proteome project, the plasma proteome project, the liver proteome project? Are they success stories?

Thierry Rabilloud: The neuron is the most difficult cell for making a proteome. So how can one expect to successfully describe a whole brain's proteome? Whether a project is a success or not depends on the people talking about it. When people are in the projects, they say it is a success. When people are out, they say it is not. The truth is in between. That's a matter of fact for all the organ specific proteome projects. In reality I'd say that from a general biology point of view these projects have not delivered a lot. Where are the breakthrough papers? They are not here – forever or not yet? I would be rather sceptical, personally speaking. Especially plasma is tremendously difficult because the dynamic range of expression goes up to ten to the ninth. There's no technology on earth that can handle a billion or more orders of magnitude. That was underestimated in the 90s, when proteomics was founded. This century, around 2002, various papers took up that point, for example one written by the Andersons. Nevertheless, people are still working on organ proteomes. Plasma proteomics probably was the most tragic part in whole proteomics in terms of people all doing all the same and keeping on and on and on. It's a nightmare.

"People in proteomics claimed a lot and delivered little."

Still?

Thierry Rabilloud: There are still people doing plasma proteomics in the classical way. They take plasma from a diseased patient and from a healthy patient, make a comparison of the plasma by classical proteomics techniques and then they look at differences. That's all terrible. Poor clinical design, poor proteomics – so the results are poor as well. They claim their findings to be specific but they are not! They try to publish it. What's new is that they get rejected. The example of clinical proteomics explains a lot of the failures that really upset large companies. Because every time one of the identified and highly advertised markers was retested, it failed. There is no single marker that came out of proteomics.

"Those people who had a hypothesis are the ones that applied the power of proteomics really successfully."

Not one single marker?

Thierry Rabilloud: No, zero! The only ones who came up with interesting candidates were the people who made assumptions at the beginning. When you compare plasma you make no assumptions except that plasma A is from a healthy person, plasma B from a diseased one. The point is, if you're studying for example liver disease, you would assume that liver proteins would be released to the plasma in the diseased condition. That's exactly what we do when we go to the doctor: he measures transaminases in the blood. If you go for that assumption and check which proteins are released by dying liver cells compared to any other tissue or condition, then you may end up with some candidates that you then test in the plasma.

Well, formulating a hypothesis is exactly what omics-sciences are not doing.

Thierry Rabilloud: But those people who had a hypothesis and made assumptions are the ones that applied the power of proteomics really successfully. I'd like to cite four of them:

First, the analysis of the nuclear pore by Brian Chait and Michael Rout. That was absolutely gorgeous and showed the power of proteomics.

Second, Michel Dejardins figured out how a phagosome works. There was a big controversy about how highly phagocytic cells manage to get the amount of membrane they need for their phagosomes. Careful proteomic studies suggested to Michel Desjardins that this membrane came from the endoplasmic reticulum, other authors claimed that this was not the case. Subsequent studies proved that Michel Desjardins was right.

Third, Jean Bergeron did a proteome analysis of the Golgi apparatus and helped to solve some biological controversies. The cell biology approaches were not very conclusive in order to solve how vesicles trafficked back and forth in the Golgi. Here again, proteomics provided the decisive evidence.

Fourth, Tony Pawson, working in Toronto, presented yet unpublished data about the dialogue between two cells on the HUPO congress in Amsterdam last year. He restricted his work to tyrosine kinases, so a very small subset of a cell's proteins. He labelled two cell types with different isotopes and co-cultured them. The cells act on each other, they send signals forth and back and via the two isotopes you can see this crosstalk happening. It's wonderful science. In fact, what the study shows is that all that people do by stimulation of isolated receptors with soluble stimulators etc. is not true, compared to what happens between the two cells.

Those studies are wrong?

Thierry Rabilloud: Previous studies are only partially right, just for the reason that when you stimulate with a soluble factor, you miss two points. First you miss the geometry of stimulation. What is done by a freely soluble stimulator is not equivalent to what happens in real life with stimulators which are grouped in a patch of membrane. Second, you miss of course the effects of the back stimulation of the target cell on the stimulating cell.

You mentioned HUPO, the human proteome organisation. Are their members now suffering from hangovers or do they see any future for their ambitious projects?

Thierry Rabilloud: People are still discussing what the human proteome project is or should be. I would say that the core of the human proteome project is to measure independently the 20,000 proteins produced from the human genome, the 20,000 gene products to be honest. That's not trivial. Many genes are only putative, we don't know whether a gene really exists and how it is expressed. Proteomics actually can reveal at large whether a gene is expressed at the protein level under these or those conditions. Maybe that can be achieved. However it has not been proven yet. In papers, people claim they can detect 5,000 proteins in a cell. That's true but for each protein they have one or two peptides on average.

You mean they detect one or two peptides of each protein by mass spectrometry.

Thierry Rabilloud: Right. Assume that these peptides are in the core of the protein. You don't know where the protein extends and whether it is modified on both ends. And that leads us to the second point: complexity. Gene products are modified in the cell to a great extent. Will we be able to handle that? We don't know. We do not agree upon which level, to which complexity we want to describe a human proteome. Let's go back to the example of signal transduction: in that case we not only need to investigate which protein is present but which is modified, for example, phosphorylated or dephosphorylated, and at which site. I think for the major part of kinases we do not know how to measure that. We do not know how to monitor such modifications on a proteome-wide scale.

"In many people's mind it was time for revenge, when they could say, 'You cloning people need us to do something with the nasty proteins'."

So what's your conclusion?

Thierry Rabilloud: The gist is: our capacity is only a few percent of the complexity of the sample on the cell's scale. If you want to analyse into great detail you have to reduce numbers. If you want to have high numbers, you will pay on the details. That's some kind of trade off between the two and that's true for every field of proteomics

Proteome scientists have been very successful in attracting funds. How are the funding agencies rating proteomics proposals now?

Thierry Rabilloud: HUPO is talking with funding agencies. Last year in Amsterdam I saw that the funding agencies were very, very cautious. They were not negative, they just say, 'let's talk, let's discuss, show us details and then, maybe we go'. They were not as enthusiastic as they were for the human genome project.

To be honest, that's fair.

Thierry Rabilloud: Sure. People in proteomics claimed a lot and delivered little. I would like to mention a very remarkable paper written by a group from Prague last year which is called the 'Déjà vu in Proteomics'. Those guys compiled quite a number of proteome analyses. Their survey showed that some proteins appear on all lists of differentially expressed proteins regardless of the experiment, tissue or species. Whatever you do to the cell, the cell reacts. And what you see when you explore the proteome of the cell is a general reaction of the cell to stress. If you are aware of that you get a wide picture of how a cell reacts to stress. But don't go beyond that with your interpretations, unless you go for more focused studies. Indeed, only focused studies so far lead to excellent results. On the other hand, we haven't yet seen any biomarker from clinical proteomics; we haven't got any general comprehension of biology from broad proteomic studies.

"We haven't yet seen any biomarker from clinical proteomics; we haven't got any general comprehension of biology from broad proteomic studies."

There is still some homework to be done in proteomics

No biomarkers? Though clinical proteomics was so generously supported?

Thierry Rabilloud: Not a single one. 14-3-3 proteins that are medium markers for Creutzfeld-Jacob disease have been identified from spinal fluid – again from a more focused approach.

One argument often heard is that proteomic studies are not consistent because there are no standards for probe sampling and probe preparation, or standards are not applied. If that's a problem, why isn't it possible to solve it?

Thierry Rabilloud: I've been in proteomics for more than twenty years and I have heard that standardisation issue throughout all that time. Workflow is always a matter of debate because no workflow is perfect. So for ages people can debate pros and cons for every workflow. And that's terrible because you discuss a lot and don't get much. Clinical routine tests tell us a lesson: all those tests now applied in medicine are those that have survived the screening process, the robustness process, meaning the enzyme to be tested can survive the coffee break or standing on the bench for one hour. Only robust tests will survive. That will be true for proteome-driven discovery of indicators, or in other words: markers.

So is that why genome research has been so successful – because DNA is robust?

Thierry Rabilloud: In essence, yes. DNA is awfully robust and it is homogenous. The genomics people have been lucky. They could develop "one-size-fits-all" experiments. You can clone and

sequence whatever you want by essentially the same methods. You're not bothered by membranes or solubility problems or very wide quantitative ranges.

DNA is stable for centuries, as we just heard from the people who have sequenced the Neanderthal genome.

Thierry Rabilloud: And the proteome is extremely instable and variable. Even in a lab under very precisely standardised procedures it will be difficult to reproduce proteomes on and on. We can see that with microbes. Working with these bugs means you need to standardise extremely culturing conditions. Aeration, cell densitiy of inoculum, composition of medium – you can detect every tiny change in their environment via changes of the proteome. Here, the largest difference is not between condition A and condition B in the experiment but between the putative identical experiments. A proteome obtained under condition A sometimes differs more to another experiment under condition A than to one under condition B. That's what I experienced in my collaborations. It is important to realise that the way you make an experiment from the biological point of view reflects how robust your setup is.

Nevertheless microbe proteomics has celebrated some success.

Thierry Rabilloud: Indeed. In a microbe you can easily detect 70 to 80 percent of the proteome without any problem. That's enormous and that gave us a major picture of what the microbe is doing, provided that your microbe culture is under top level control. Take as an example the work on *Bacillus subtilis* done by Michael Hecker from Greifswald, Germany. Roughly, we see all of the microbes' proteins except inner membrane proteins and DNA-binding proteins. For example, we have never seen the lac-repressor of *E. coli*.

"We have never seen the lac-repressor of* E. coli.*"

To sum up: you think that proteome researchers should go back to the lab and do their homework instead of talking big?

Thierry Rabilloud: Maybe they haven't even realised that. Proteomics is no more than analytical chemistry. Speaking about low limits of detection, quantification, dynamic range, reproducibility issues – that's the dirty homework to do. But many scientists in proteomics behaved and still behave as children in a toy shop: they want everything that's on the shelf. They haven't done their homework because they were so hyped. And here we come to the psychological dimension of proteomics, which I mentioned at the very beginning of our talk. Since the end of the 70s, cloning and sequencing were on their way. And people working on proteins were smiled at because they were the 'no-cloning' people. Then the time of revenge came…

Revenge?

Thierry Rabilloud: Yes, in many people's mind it was time for revenge, when they could say, 'You cloning people need us to do something with the nasty proteins'. That's when they went for the headlines in the newspapers. Really frankly speaking, that was amazing.

To conclude, do you have any optimistic views for our readers?

Thierry Rabilloud: (Laughs), well, I'm convinced that at the end of the day we will get amazing results with proteomics, provided we do our homework properly.