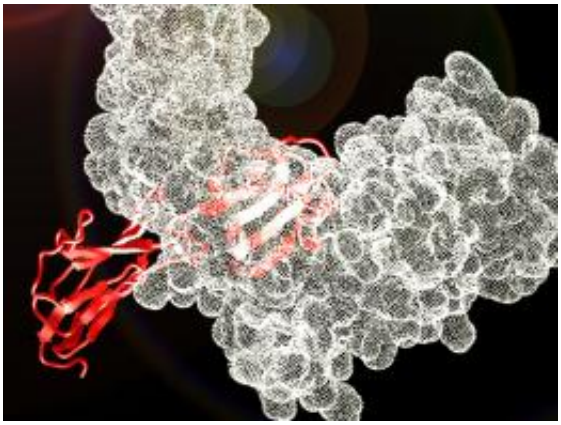
Research team fully maps human proteome

September 24 2010, By Peter Rueegg



The big wide world of proteins: Systems biologists at ETH Zurich and the IBS have drawn a complete map of the human proteome. Image: e-pics/ETH Zurich

A Swiss research team from ETH Zurich, led by Professor Ruedi Aebersold, and from the Institute for Systems Biology, Seattle, has used mass spectroscopy methods to fully map the human proteome for the first time. The data is being made available to all researchers.

A number not to be forgotten by anyone intending to work with the human genome on the one hand, and henceforth with the human proteome on the other. The reason is that this is the number of proteins which researchers at the Institute for Systems Biology (IBS) in Seattle and at ETH Zurich from the group led by Professor Ruedi Aebersold have recorded and identified using various mass spectrometry methods, and whose data they have entered into a database as a reference. The

researchers describe this as the “Gold Standard Reference for the human proteome”.

“This reference database will now be made available to all biologists for further research,” said Ruedi Aebersold in an interview with ETH Life, “just as the human genome data from the ‘Human Genome Project’ was published openly at that time for all interested experts.” The researchers are making their data accessible via the ISB/ETH SRMAtlas. This database is part of the PeptideAtlas developed by the two institutions over the past few years. The researchers have just announced this at the Ninth Annual World Conference of the Human Proteome Organization (HUPO) in Sydney.

Any samples are fully determinable

The proteome data enables researchers to determine the number and type of proteins in any kind of biological sample using various mass spectrometric measurements. This is an important development that will significantly improve the reliability and reproducibility of proteomics. The researchers hope it will greatly speed up fundamental and applied research.

Scientists working with Ruedi Aebersold at the Institute for Molecular Systems Biology at ETH Zurich and various representatives of the ISB, which Aebersold co-founded, took part in the project. The Proteome Atlas project was financed on the one hand by funds received by the ISB from the US National Health Institute, and on the other Ruedi Aebersold obtained an ERC Advanced Grant in Proteomics, amounting to 2.4 million Euro paid out over five years.

Interview with Ruedi Aebersold on decoding the human proteome:

Professor Ruedi Aebersold, is it correct to say that you now have spectra for every protein coded and expressed by the 20,300 human genes, thus enabling all the protein species present in a human cell at any given time to be identified using these spectra?

Ruedi Aebersold: Yes, that is correct. Biologists currently assume that the human genome has about 20,300 loci, i.e. gene locations, that code for proteins. Now we have actually succeeded in recording reference data for all the proteins encoded by these loci.

Many proteins are modified in the cell only afterwards, which increases the number of potential protein species even further. Do spectra also exist for such proteins?

Each of these loci can produce several different polypeptides with chemical modifications or by a process called “alternative splicing”. At present we cannot generally distinguish between the various forms of proteins produced by each locus. That is the next stage in our project, which is already underway.

How long did this mapping work take?

The actual measurements took about one year. Of course the development of the technology we used to generate this data, as well as the entire computer infrastructure including programming and building up the databases, took much longer: we have been working on it since 2003.

What were the biggest difficulties?

The greatest difficulty and challenge was to track down the rare protein species and to obtain their spectra. We overcame this difficulty by synthesising fragments of such proteins which we “predicted” by using a

computer, and by carrying out the measurements on these synthetic products afterwards. These spectra now act as the “Gold Standard” for discovering the corresponding proteins in any kind of biological sample.

Initially the decoding of the human genome was celebrated as an important step, and the decoding of the human proteome has now taken place. What is the significance of this new achievement?

I think it is an important step. However, the project will gain its real significance through the extent to which scientists use the resources that have now been made available, and how useful these prove to be. What we have actually created is a “map” allowing other people to research the proteome much more efficiently in relation to their own studies. The difference compared to the genome project is that we have constructed not just a catalogue but a map that helps other researchers to navigate through the proteome in their studies.

What does the fact that we now know the human proteome mean for research?

Proteins control or catalyse all the processes in the human body. We now have the instruments to enable every research laboratory, with certain restrictions, to discover and quantify every human protein. This will bring about great changes in biology and medicine, since we can now reliably and reproducibly measure all the protein constituents of a biological system.

Provided by ETH Zurich