Proteomics for biomedicine: a half-completed journey

Although I studied physics and mathematics until the master’s level, I always knew that I wanted to get into biology, whose golden age I was sure was happening in my lifetime. When the opportunity came knocking – in the shape of John Fenn, who was spending a sabbatical at the University of Göttingen – I did not hesitate to take up his offer to come with him to Yale for my PhD. John was working on a far-out idea: that one could ionize biomolecules by applying an electric field to the outlet of a needle through which a liquid was flowing. At first there was little interest in ‘electrospray’, but as soon as it became clear that large, intact proteins survive this process (Mann et al, 1989; Meng et al, 1988), the community was hooked. This included the Nobel Prize committee, who gave John a share of their prize for chemistry in 2002.

I took one of the most ambitious and stressful decisions of my professional life: I bet everything on the success of mass spectrometry-based protein characterization. Electro spray has been my mainstay ever since. The only interruption was as a post-doc, which I did in Denmark – following my Danish wife, whom I had met at Yale. Peter Roepstorff’s group taught me a lot about protein chemistry but they had not acquired electrospray yet, so I missed the initial wave of discoveries done with our new technology! This turned out to be a blessing in disguise, however, because it motivated me to develop software instead. Specifically I attacked the problem of how to identify peptides in sequence databases given a minimal amount of mass spectral information. When in 1992 I was offered a group leader position at the European Molecular Biology Laboratory (EMBL), I took one of the most ambitious and stressful decisions of my professional life: I bet everything on the success of mass spectrometry (MS)-based protein characterization. At the time, despite much talk, MS had not become a serious tool for high sensitivity protein detection in biology, which was dominated by chemical methods such as Edman degradation instead. Fortunately, we soon made two crucial breakthroughs: Matthias Wilm devised nanoelectrospray, a miniaturized and highly sensitive form of electrospray, and Andrej Shevchenko developed protocols for liberating proteins from gels so that they could be analysed by MS. Together with our novel bioinformatic algorithms (Mann & Wilm, 1994), and the beginning genome sequencing efforts, this enabled identification of minute amounts of protein – far exceeding the capabilities of Edman sequencing and firmly establishing mass spectrometry in molecular biology research (Wilm et al, 1996). Among the prize catches of these years was the identification of caspase-8 (with Peter Krammer and Vishua Dixit) and the catalytic subunit of telomerase (with Tom Check). I am sure that these successes would not have happened without the high pressure environment at EMBL. Together with complementary work by the Yates and the Aebersold groups, these were the first steps of mass spectrometry in biological research, which eventually led to its establishment as the standard tool for protein identification.

After EMBL, my move back to Denmark (and reunion with my family) was made possible by the Danish Natural Research Council. The next 7 years were very fruitful and our group developed many of the technologies that are now in general use – including the method of stable isotope labelling by amino acids in cell culture (SILAC; Ong et al, 2002), which has become a gold standard in quantitative accuracy for proteomics. I continued to be involved in programming too, which gave us quite a competitive edge in devising new strategies to retrieve biologically meaningful data from proteomics. During these years we started to apply proteomics to define the members of protein complexes. First examples included the U1 subunit of the yeast spliceosome (with Reinhard Luehrmann), the entire human spliceosome (with Angus Lamond) and the centrosome (with Erich Nigg). Interaction and organellar proteomics continue to be some of the most promising areas for proteomics and together with Tony Hyman in Dresden we are currently pursuing a large-scale effort to produce a high quality human interactome. We also used proteomics to study mechanisms of growth factor stimulation of SILAC labelled cells. By precipitating and then quantifying tyrosine-phosphorylated proteins at different time points, we could follow the signal as it spread to many and diverse substrates (Blagoev et al, 2004). Some of the proteins that we cloned at the time have become quite important in the growth factor signalling field. Several years later, we extended these experiments to generate a first global picture of ‘early information processing’ by the cell in response to growth stimulus (Olsen et al, 2006).

In 2005, I moved to the Max-Planck Institute of biochemistry in Martinsried (Munich) where we diversified into a...
variety of biological and medical directions while keeping our core emphasis on instrumentation and bioinformatics. The Max-Planck institutes are fantastic environments for science in general and for our work in particular. Here, my group is able to work on the entire chain of proteomics – from technologies of proteomic sample preparation, through chromatographic and mass spectrometric innovations to methods in computational proteomics (Fig 1). Bioinformatics development connected to the analysis of proteomics data sets has always been a major focus in my group but in Martinsried it was kicked into high gear by Jürgen Cox. He developed the MaxQuant suite of proteomics tools that are now widely used by the community to analysed proteomics datasets (Cox & Mann, 2008). We then apply all these technologies to a wide range of biological problems; partly to show that proteomics can be a powerful tool in these different fields.

The most ambitious project we undertook at our new location was to crack a complete proteome. One of the limitations of proteomics had always been that usually only a small number of proteins were actually identified and quantified. Pre-MS techniques such as 2D gel electrophoresis usually identified a few dozens of proteins, a far cry from the thousands of probes that the microarray community was putting on their chips. To show that the full proteome was amenable to MS analysis, we collaborated with Tobias Walther (now at Yale), to quantify haploid versus diploid yeast (de Godoy et al, 2008). Very recently, we have shrunk the analysis time needed to cover nearly the entire yeast proteome to just a few hours (Nagaraj et al, 2011a). The human proteome is not quite complete yet, but we are getting very close (Nagaraj et al, 2011b). However, these studies are just first steps; in the future the task for proteomics will include detection and quantification of protein isoforms, a nearly complete set of protein modifications and doing all this as a function of time and cellular localization – a full program for at least one more generation of proteomics researchers!

Using a variant of the SILAC technology called ‘spike-in’ or ‘super’ SILAC it is now possible to measure the proteome of tumour biopsies at great depth and with extremely high precision. Recently, proteomics technologies have become sufficiently evolved that it is now realistic to measure clinical material. Using a variant of the SILAC technology
called ‘spike-in’ or ‘super’ SILAC it is now possible to measure the proteome of tumour biopsies at great depth and with extremely high precision (Geiger et al., 2010). This is one of the areas that my group will focus on in the coming years and we already have proof of principle in defining protein expression signatures in difficult to diagnose lymphoma subtypes. It is clear that there is now an opportunity to make a great clinical impact using high resolution and robust proteomics technologies (Fig 2).

In conclusion, whenever one really wants to understand biological function, one has to deal with proteins and mass spectrometry is the method of choice to do this. Clearly, the fun is just beginning as we now get to apply the tools developed by the community over the last decades. So, what does the future hold for proteomics and where will it make unique contributions? Areas to watch include the analysis of complete mammalian proteomes, including absolute quantification of cellular proteins, increasingly sophisticated studies of the function of thousands of post-translational modifications as well as protein interaction studies. Proteomics is starting to be used to probe the effects of genome variation between humans at the functional level and I predict that this will be an expanding area. In a slightly longer perspective, proteomics will become an important basis on which systems biological modelling of the cell will be built (Cox & Mann, 2011).

The author declares that he has no conflict of interest.

References
Blagoev B, Ong SE, Kratchmarova I, Mann M (2004) Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nat Biotechnol 22: 1139-1145
Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26: 1367-1372
Cox J, Mann M (2011) Quantitative, high-resolution proteomics for data-driven systems biology. Ann Rev Biochem 80: 273-299
de Godoy LM, Olsen JV, Cox J, Nielsen ML, Hubner NC, Frohlich F, Walther TC, Mann M (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. Nature 455: 1251-1254
Geiger T, Juergen C, Pawel O, Wisniewski JR, Mann M (2010) Super-SILAC mix for quantitative proteomics of human tumor tissue. Nat Methods 7: 383-385
Mann M, Meng CK, Fenn JB (1989) Interpreting mass-spectra of multiply charged ions. Anal Chem 61: 1702-1708
Mann M, Wilm M (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Anal Chem 66: 4390-4399
Meng CK, Mann M, Fenn JB (1988) Of proteins or proteins – a beams a beam for a that (Burns, O.S.) Z Phys D Atom Mol Cl 10: 361-368
Nagaraj N, Kulak NA, Cox J, Neuhaus N, Mayr K, Hoerning O, Vorm O, Mann M (2011a) Systems-wide perturbation analysis with near complete coverage of the yeast proteome by single-shot UHPLC runs on a bench-top Orbitrap. Mol Cell Proteomics DOI: 10.1074/mcp.M111.013722
Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Paabo S, Mann M (2011b) Deep proteome and transcriptome mapping of a human cancer cell line. Mol Sys Biol 7: 548
Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell 127: 635-648
Org SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1: 376-386
Wilm M, Shevchenko A, Housthaeve T, Breit S, Schweigerer L, Fotsis T, Mann M (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 379: 466-469

DOI 10.1002/emmm.201100198

Figure 2. Proteomic analysis of breast cancer samples (Tamar Geiger, Jacek Wisniewski and Matthias Mann).