Featuring cutting edge research projects of the Netherlands Proteomics Centre
About

The Netherlands Proteomics Centre (NPC) is a strategic collaboration of research groups from six universities, three academic medical centres and several biotech companies. With a scientific program addressing key areas of proteomics in 56 projects, and specialised ‘research hotels’, the NPC performs high-quality research and knowledge transfer in an international context. The NPC is part of the Netherlands Genomics Initiative.

In NPC Highlights researchers present progress and results from NPC projects of the scientific program and the research hotels. NPC Highlights is published by the Netherlands Proteomics Centre.
Welcome

The year 2008 promises to be a busy but also exceptional year for the Netherlands Proteomics Centre (NPC).

A very good start has already been made, as on 19 February we were happy to welcome more than 180 people to the NPC progress meeting in the historical atmosphere of the Academy Building in Utrecht. They all attended the plenary program, with keynote lectures by Blagoy Blagoev (University of Southern Denmark) and Huib Ovaa (Dutch Cancer Institute, NKI), and enjoyed the poster discussions and the lively get together. As in previous years, the visitors were remarkably young, with participants from all NPC locations, but also from other universities and industry. The plenary lectures by the representatives of the NPC themes are presented in this issue.

Moreover, the NPC community maintains the production of high quality and quantity output. This emerges from the NPC progress report 2007 on behalf of NGI, as measured by the kinds and quantities of scientific publications, patents, presentations, new applications and collaborations with other academic and industrial researcher groups.

Another highlight this year will be the HUPO2008 Annual World Congress, to be held from 16-20 August in Amsterdam. The program has been nearly finalized, as indicated on www.hupo2008.com and a large and still growing number of participants and exhibitors has already registered. Please check the website for more news and be sure to submit your abstract and register yourself!

Last, but not least, the business plan for the next phase of the NPC has just been submitted to NGI. With much effort from many participants, the document outlines a focussed program for the years 2009-2012. The plan builds on the proven successes of the NPC with emphasis on the collaboration with selected NGI centres, an additional bioinformatics impulse and an extended valorisation program. We hope for a quick and positive evaluation of our plan, as we are fully confident that the NPC community is ready for the next phase of successful and rewarding collaboration in the coming years. In the next issue of NPC Highlights we hope to give more detail on the outcome and the continued NPC program.

Albert Heck, scientific director NPC
Before differentiated stem cells can be clinically applied they have to be cultured to homogenous populations. For optimizing the differentiation protocols, the authors searched for proteins involved in signalling pathway in the early steps of differentiation. Furthermore they identified membrane proteins that could be used as selectable markers for heart muscle cells. These applications become possible by combining quantitative proteomics techniques and biochemical purification schemes developed over the last few years at the NPC.

Embryonic stem cells (ESCs) are characterized by their combined ability to self-renew indefinitely, and to undergo extensive differentiation, eventually to all the about 200 cell types of the human body. Both properties — proliferation and pluripotency — have put ESCs at the centre stage of scientific and public interest, since they are considered not only to have enormous potential in regenerative medicine but also in providing insight into human development and disease. Efforts to control differentiation of ESCs in vitro have been successful in several cases, but these are only initial steps before hESC-derived differentiated cells can be applied clinically. For instance, such cells can only be successfully transplanted when they are available in sufficient numbers and as homogeneous populations to prevent formation of teratocarcinomas. This can be achieved either by optimizing differentiation protocols, or by isolating cells of interest from heterogeneous populations.

Our interest focuses on two aspects of ESC differentiation: the sequence of events in the early steps of differentiation and the molecular characterization of cells specifically differentiating to heart muscle cells (cardiomyocytes). For the first topic, we profile the dynamics of phosphorylation from the onset of differentiation, providing insight in the signalling cascades that are (in)activated during this process. For the second topic, we focus on plasma membrane proteins since these are likely to provide the targets that can be used for affinity-selection of cardiomyocytes to create homogeneous populations.

**Stem cell dynamics** Because of the unique properties of stem cells, they have been the subject of both targeted and genome-wide studies. This includes several proteomic efforts, including our own [1], to characterize stem cell lines in terms of protein composition. Although such datasets are valuable, they only provide a static picture of one cell type isolated in one specific condition. It would be more informative if this was combined with a measure of relative expression levels of these proteins. This
Stem cells are able to divide indefinitely and to differentiate into any adult cell type when the appropriate conditions are applied. This is promising for new therapeutic applications. “Before you are able to guide a stem cell into a specific cell type, you first have to understand how the process exactly works,” Jeroen Krijgsveld of the University of Utrecht explains. “We can mimic the natural process of stem cell differentiation to for instance heart muscle cells, but the process is not very efficient. Less than 25 percent of original stem cells differentiate in the desired type. We therefore searched for proteins and processes that play a crucial role in cell differentiation.”

The work presented in this article focuses on two aspects of cell differentiation that can contribute to obtaining homogeneously differentiated cells. The first aspect concerns the identification of proteins accessible at the outside of the cell membrane of heart muscle cells. Krijgsveld explains: “This part of the research connects directly to application. If we know which specific proteins are present on the outside of the membrane of developing heart muscle cells, we can use them as handles to selectively isolate these cells from a heterogeneous population of differentiating cells.”

The second line of research is to elucidate the proteins involved in the on-set of differentiation, focusing on protein phosphorylation. Therefore Krijgsveld compared the composition of phospho-proteins at four different moments in the early development stage of stem cells. “We have identified hundreds of proteins whose phosphorylation is modulated during differentiation. But it is still a complicated puzzle. The question now is how to put the pieces together and discover new patterns of regulatory mechanisms.”

Krijgsveld and co-workers aim to solve this puzzle in the near future with the cell biologist from the Hubrecht Institute.

What this research is about:

**Identifying new proteins to guide stem cells into heart muscle cells**

Stem cells are able to divide indefinitely and to differentiate into any adult cell type when the appropriate conditions are applied. This is promising for new therapeutic applications. “Before you are able to guide a stem cell into a specific cell type, you first have to understand how the process exactly works,” Jeroen Krijgsveld of the University of Utrecht explains. “We can mimic the natural process of stem cell differentiation to for instance heart muscle cells, but the process is not very efficient. Less than 25 percent of original stem cells differentiate in the desired type. We therefore searched for proteins and processes that play a crucial role in cell differentiation.”

The work presented in this article focuses on two aspects of cell differentiation that can contribute to obtaining homogeneously differentiated cells. The first aspect concerns the identification of proteins accessible at the outside of the cell membrane of heart muscle cells. Krijgsveld explains: “This part of the research connects directly to application. If we know which specific proteins are present on the outside of the membrane of developing heart muscle cells, we can use them as handles to selectively isolate these cells from a heterogeneous population of differentiating cells.”

The second line of research is to elucidate the proteins involved in the on-set of differentiation, focusing on protein phosphorylation. Therefore Krijgsveld compared the composition of phospho-proteins at four different moments in the early development stage of stem cells. “We have identified hundreds of proteins whose phosphorylation is modulated during differentiation. But it is still a complicated puzzle. The question now is how to put the pieces together and discover new patterns of regulatory mechanisms.”

Krijgsveld and co-workers aim to solve this puzzle in the near future with the cell biologist from the Hubrecht Institute.

Research Theme NPC1: Accurate comparative proteomics and protein quantitation

would be particularly useful in time-course studies or for the comparison of different cell lines or populations. Since both of these applications are highly relevant for the questions we wished to address, we have developed techniques to label human ESC with stable isotope-tagged amino acids [2] (as discussed in NPC Highlight 6, 2007; p8), permitting large-scale quantitative analyses commonly known as SILAC-based proteomics [3]. We have used this to assess the dynamics of phosphorylation during early differentiation of ESC, and in the comparison of cell lines searching for cardiomyocyte-specific membrane proteins.

Plasma membrane proteins are key players in cellular communication because they are exposed on the outer surface of cells, playing an essential role in cellular communication. They serve as sensors for extracellular factors and can activate intracellular activation pathways upon their binding. In this way, cells are capable of effectively responding to their environment. This is also true for ESC, which responds to growth factors released by the stem cell niche (or added to the culture media) to either remain pluripotent, or to start differentiation pathways to a specific cell fate. The direction
of differentiation depends largely on the pathway (or combination of pathways) that is activated, and is accompanied by a change in protein expression pattern that will gradually determine the ultimate stem cell fate. Therefore, many cell types can be characterized by a specific protein expression pattern. In many cases, these include plasma membrane proteins which are often used for cell sorting by affinity selection.

**Plasma membrane proteins** In our research we are interested in identifying the plasma membrane proteins that are characteristic for cardiomyocytes. Using antibodies directed against these proteins in live-cell FACS analysis, highly homogeneous cell populations could then be obtained which would mark an important milestone towards clinical application of these cells. To identify these proteins, several hurdles need to be taken: first, membrane proteins are notoriously difficult to obtain because of their hydrophobic nature, and second, their supply is limited since stem cells can typically be grown in relatively small numbers.

We have therefore developed and optimized a protocol to address both of these aspects. By systematically testing and verifying a range of subsequent experimental steps, a protocol was designed which allowed us to purify and identify plasma membrane proteins from sample amounts as small as 500,000 cells. This protocol combines centrifugation steps with separate plasma membranes from most other membranous cellular compartments, wash procedures at high pH to remove non-integral membrane proteins, a deglycosilation step and subsequent proteolytic digestions in water and organic phase giving the highest yield of peptides. These were then separated by SCX and RP chromatography and identified by mass spectrometry. We have applied this to the comparison of mouse embryonic stem cells and carcinoma cells. Among the hundreds of identified plasma membrane proteins, several were distinctive for either cell line, which could thus explain the biological difference between them. Currently we are in the process of biologically verifying these data. We have used the same protocol in a larger quantitative analysis where we have analyzed cells that had differentiated to cardiomyocytes. Since the differentiation is rather inefficient and leads to approximately 25% cardiomyocytes, the identification of cardiomyocyte-specific proteins is not straightforward. Therefore we have labelled these cells with SILAC for the quantitative analysis of plasma membrane proteins to be compared both to stem cells that had not differentiated at all (0% cardiomyocytes), and to foetal hearts estimated to contain 80% cardiomyocytes. Based on the expected increase in the cells of our interest in each of these samples, we have identified several candidate membrane proteins, which are currently tested for their specificity for cardiomyocytes.

**Growth factors** ESC are rapidly proliferating cells primarily because of a relatively short G1-phase during the cell cycle [4]. Self renewal can be interrupted, thereby inducing commitment to differentiate, by growth factors controlled by the micro-environment of ESC. In vitro, this can be exploited by adding purified growth factors (BMP, activin, FGF) to the culture media, or by using feeder cells that excrete (largely unknown) differentiation-inducing factors. Subsequently, signalling cascades are activated that trigger downstream events ultimately leading to terminal differentiation of ESC. The major challenge is to identify those growth factors that are best in inducing differentiation to a particular cell fate. Only few of these combinations have been found, aggravated by the fact that mouse ESC, where most research has focused on, respond quite different when compared to human ESC [5].

We have studied the early events in signalling when inducing human ESC with bone morphogenetic protein (BMP). In order to follow phosphorylation levels on individual proteins, we have approached this by metabolic labelling of undifferentiated cells. These were then mixed with unlabelled cells 30, 60 or 240 minutes after the addition of BMP. From each of these combined samples, we have extracted and digested proteins, followed by SCX-separation of the resulting peptides.

---

**Figure 2** Some examples of sequence logos of kinase motifs, deduced from protein sequences surrounding phosphorylated serine residues. In some cases the kinase targeting this sequence is known (top), in other cases it is not (bottom). Generated by the program MotifX.

**Figure 3** A protein network around casein kinase 2 (Ck2) and NFκB. Symbols in red and green indicated proteins with increased and decreased phosphorylation levels, respectively, in embryonic stem cells 30 min after the induction of differentiation. Generated by the program Ingenuity.
A crucial step was the enrichment of phosphopeptides in each of the SCX fractions by on-line TiO2 chromatography [6]. In total, 144 LC-MS runs were performed using LTQ-FT and LTQ-orbitrap MS, collecting over 800,000 spectra. Using rigorous filtering, this has resulted in the identification of about 2000 phosphopeptides in each of the 3 datasets, totalling over 3600 unique phosphopeptides across all samples.

Phosphorylation profiling — More importantly, phosphorylation levels can now be quantitatively studied across the three time points. This can be done at various levels. First, dynamics of phosphorylation can be followed over time at the peptide level, since we have mapped the phosphorylation site for most peptides. This is important, since signalling can be mediated by one or few specific phosphorylation events. Indeed we have found many different profiles, which can be different for individual phosphorylation sites in the same protein (see Figure 1). Second, profiles can be deduced based on conserved regions around phosphorylation sites in the entire dataset (using software tools like MotifX). Such motifs are indicative of activity of specific kinases which usually target their substrates in a sequence-specific manner. Figure 2 shows some examples of identified profiles, some of which indicate the activity of specific kinases (cdc2, MAPK), others could not be correlated to a specific kinase so far. Third, (differential) phosphorylation can be mapped to signalling pathways by using gene ontologies and prior knowledge of protein-protein and biochemical interactions, and from literature. Using the program Ingenuity, networks can be constructed and displayed reflecting activation of specific pathways (see Figure 3). Finally, in our study we have identified 26 proteins that are widely used as stem cell markers because they are associated with the undifferentiated state of ESC. In 12 of these proteins, 30 phosphorylation sites were identified, of which 25 were unknown so far. This could indicate that (de)phosphorylation of these proteins could be associated with the induction of differentiation. Currently we are integrating these combined data in a model that describes the events in the initial phase of ESC differentiation.

References

Acknowledgements
I am very grateful to Dennis van Hoof, Stefan Braam, Robert Passier and Christine Mummery (Hubrecht Institute), Wilma Dormeyer, Javier Muñoz, Martijn Pinkse, and Albert Heck (Utrecht University) who all have made major contributions to the work described here.

Author
Contact
Jeroen Krijgsveld, PhD
Utrecht University
Biomolecular Mass spectrometry and Proteomics group
T +31 30 253 9974
j.krijgsveld@uu.nl
www.mass-spectrometry.eu

Summary
Embryonic stem cells (ESC) are unique cells that can propagate as primitive cells, but can be driven to specific cell types by extracellular signals. These properties make ESC highly interesting for regenerative medicine, aiming to replace damaged tissue in the adult by ESC-derived cells. However, before this can be applied clinically a better understanding needs to be gained on the molecular processes governing pluripotency as well as differentiation behaviour of ESC.

We are approaching this from a proteomics perspective, where we focus on two aspects. First, by monitoring phosphorylation events shortly after the induction of differentiation, clues can be found to what pathways are involved in stem cells starting to progress into differentiation. Second, we have developed protocols to enrich for plasma membrane proteins, which are a potential source of proteins reflecting lineage-specific differentiation. We are particularly interested in identifying membrane proteins characterizing heart muscle cells. In both approaches, quantitative proteomics has revealed several candidate proteins that are currently included in follow-up studies.
Overproduction of membrane proteins is often problematic but can be alleviated by using different production hosts. We have developed a platform for the simultaneous and high-throughput analysis of the expression potential of (membrane) proteins in multiple hosts. To increase the capacity of this platform even more, we aim at identifying and ultimately overcoming the bottlenecks limiting the functional overexpression of complex proteins.

Membrane proteins, in particular channels, transporters and G-protein coupled receptors, represent the major targets of the pharmaceutical industry for the development of drugs or antibiotics. The rational design of new drugs or therapeutic strategies is severely hindered by the lack of high resolution structures of membrane proteins and a basic understanding of function-structure relationships in these systems. One of the major bottlenecks in the structural analysis of integral membrane proteins is the production of these proteins in a functional state. This mostly results from the fact that the biogenesis of polytopic membrane proteins is far more complicated than that of soluble proteins. It involves several steps such as targeting of the nascent polypeptide chain to the membrane, insertion into and assembly in the membrane and each of these steps requires distinct components [1-3]. Upon membrane protein overproduction, the capacity of this pipeline is easily exceeded resulting in the production of misfolded protein, unsuitable for structural analysis. In Theme 2 of the NPC program, the groups of Driessen and Poolman use a proteomic approach to analyze expression hurdles and to improve the synthesis, membrane insertion and assembly of polytopic integral membrane proteins.

The heterologous expression of large multidomain assemblies and membrane proteins in a functionally competent state is often problematic in the well-established expression hosts E. coli and yeast. Lactococcus lactis has proven to be an excellent alternative and complementary host, but subsets of proteins fail to be produced in large quantities in this organism either. In collaboration with the group of Jan Kok (Molecular Genetics, University of Groningen), we have set out (i) to identify the bottlenecks in the production of membrane proteins; (ii) to develop methods for the rapid screening of their functional state; and (iii) to improve host strains (L. lactis and E. coli; see Box) for enhanced protein synthesis and functional assembly. An outline of our strategy is presented in Figure 1.

**Protein production platform** A prerequisite for any protein production platform is the efficient cloning and expression
What this research is about:

**Improving membrane protein production for the design of new drugs**

“The pharmaceutical industry is very interested in membrane proteins,” says Professor Bert Poolman from the department of Biochemistry at the University of Groningen. “About 70% of present medication is mainly directed against membrane proteins.” Membrane proteins therefore form an important pool for the generation of new drugs. However to design new drugs it is important to have detailed knowledge about the structure and function of membrane proteins.

To date relatively little is known about their structure-activity relationships as opposed to similar knowledge of soluble proteins due to several bottlenecks. Less than 1% of the high resolution structures available are from membrane proteins. One of the major bottlenecks is reproduction of functional membrane proteins which is far more complicated than the reproduction of soluble proteins. “There are several steps where it can go wrong,” explains Poolman. “The nascent protein needs to be directed to the right place in the membrane, it needs to be inserted into and assembled in the membrane and it has to have the proper folding.”

To tackle these problems Poolman and colleagues set up a protein production platform in which they focused on several aspects. They devised an efficient method to clone large numbers of genes and screen the expressed products (the proteins). Furthermore they set up a fast method to check the quality of the proteins, for instance to check whether they are properly folded or whether they are active. Also they composed, on the basis of proteomic data, a chemically-defined growth medium to improve expression of complex membrane proteins in the bacterium *L. lactis*.

The methods together led to a better expression of membrane proteins and, subsequently, isolation of correctly folded proteins. Poolman and colleagues’ research now concentrates on identification of other factors that might limit membrane production yield. The results will be used to design new *L. lactis* bacterium strains to improve protein production.

**Research Theme NPC2: Membrane Proteomics**

![Figure 1](image-url) **Figure 1** Outline of the NPC protein production platform. Blue lines indicate the approach of identifying expression bottlenecks through analysis of the proteome (and transcriptome), which is then followed by rational redesign of hosts and/or expression vectors. The red branch indicates a random screening approach from which key factors for protein production are subsequently identified by proteomics-based methods. Both approaches are complementary and used iteratively to improve the production of recombinant membrane proteins.
screening of large numbers of genes. We have developed a
generic method for high-throughput cloning in micro-organisms
that are less amenable to conventional DNA manipulations,
lke L. lactis [4].
The method makes use of ligation-independent cloning in
an intermediary Escherichia coli vector, which is rapidly
converted via vector backbone exchange (VBEx) into a bona
fide, organism-specific plasmid ready for high-efficiency
transformation.
Proof-of-principle for VBEx has been demonstrated for L. lactis,
but the method may be applicable to all organisms for which
plasmids are available. Moreover, we have developed methods
for the rapid screening of the folding state of membrane
proteins [5] and assays of their activities [6] and lipid depen-
dencies [7]. The quality control and lipid requirements of het-
 erologously produced membrane proteins are critical factors
for success but often ignored. Our quality control involves the
simultaneous quantification of folded and aggregated mem-
brane protein using GFP-fusions [5]. The method requires only
standard electrophoresis equipment, small culture samples,
is not labour-intensive and can greatly facilitate the optimi-
ization of the overexpression of both membrane and soluble
proteins.

Growth medium To identify bottlenecks in the recombinant
expression of membrane proteins, we have used proteomics
approaches to identify differentially expressed proteins in
L. lactis. So far, we have focussed on the differences in the

The overexpression levels of a subset of membrane proteins
are approximately 10 times higher when cells are grown in
M17 as compared to CDM; the latter medium is preferred
for the isotopic labelling of proteins or the incorporation
of amino acid analogues. Thus, the initial goals were to improve
the protein production in CDM grown cells and to identify the
biogenesis bottlenecks for lowly expressing proteins.

For the analyses, cells were grown in triplicate in pH-con-
trolled bioreactors under well defined conditions. Both for the
analysis of the soluble and membrane proteome, cell fraction-
ation was followed by two-dimensional separation to reduce
the complexity of the samples. Peptides were identified by
MS/MS. For quantification of the soluble and membrane prote-
omees, Cy dye labelling together with differential gel electro-
phoresis (DIGE) technology and isotopic labelling (iTRAQ™)
in combination with multidimensional liquid chromatography
were used, respectively.

Nitrogen metabolism The main differences in the proteomes
of M17 and CDM-grown cells were confined to the differential
expression of glycolytic enzymes and proteins involved in
nitrogen metabolism and polypeptide synthesis. Despite the
fact that CDM-grown cells displayed a reduced capacity to syn-
thesize complex multidomain membrane proteins, the levels
of the metabolic and ribosomal proteins were highest in these
cells. We thus anticipated that the maximal rates of glycolysis
and the energy status would be highest in CDM-grown cells but

proteomes of cells grown in rich M17 broth or chemically-
derined medium (CDM), and the consequences of membrane
protein production on the soluble and membrane proteomes.

The choice for L. lactis or other
alternative hosts
Although the tool box for gene cloning and expression studies in
E. coli is far greater than for any other organism, the screening of
the expression space becomes much larger when alternative hosts
are used. E. coli and L. lactis belong to the γ-proteobacteria and
Firmicutes branches of the phylogenetic tree of life, respectively.
Contrary to E. coli, L. lactis is a Gram-positive bacterium with a
single membrane and a relatively simple metabolism. The lack of
an outer membrane offers easy access to and manipulation of the
activities of channels, transporters and receptors in the plasma
membrane. Moreover, the redundancy of membrane proteins in
L. lactis is far smaller than in E. coli, which has advantages in
complementation and screening studies.
observed no significant differences for the two media. Because a higher level of ribosomal components does not easily explain the down-regulation of the expression of a subset of membrane proteins, we then focussed on the nitrogen metabolism components.

In fact, each of these proteins turned out to be under the control of the general transcriptional regulator CodY, which implied changes in amino acid and peptide uptake and peptide degradation. The differently-expressed proteins suggested specific nitrogen limitations in the CDM medium as compared to M17. Indeed, upon specific adjustment of the amino acid/peptide contribution in CDM, the expression levels approached those of M17-grown cells (see Figure 2). The newly designed medium, CDM+, allowed for similar levels of expression of (complex) multidomain membrane proteins as observed in M17 and, importantly, allowed for incorporation of amino acid analogues and isotopic labels. For two proteins, it was subsequently shown that CDM-grown L. lactis cells incorporate selenomethionine with near 100% efficiency.

**Expression bottlenecks** For the identification of expression bottlenecks, we have compared the proteomes of cells expressing membrane proteins to a high level (in some cases with little growth defect, in other cases the growth inhibition was complete) with those producing little if any protein (again with very different growth defects). Surprisingly, the number of soluble proteins regulated up or down was relatively small (<35) despite large differences in growth rates upon induction of gene expression. Our work indicates that the response of L. lactis to membrane protein overexpression involves a set of general stress proteins, which could be explained by the presence of a fraction of incorrectly folded protein. The formation of some misfolded protein might be a consequence of overloading of the membrane biogenesis machinery. In parallel with the proteomics analyses, the group of Kok analyzed the transcript levels of the same cell samples and the combined dataset is currently used to redesign L. lactis strains for improved protein production.

**Summary**

Membrane proteins, in particular channels, transporters and G-protein coupled receptors, represent the major targets of the pharmaceutical industry for the development of drugs or antibiotics. However the design of new drugs is hindered by a lack of high resolution structures of these membrane proteins and a basic understanding of function-structure relationships in these systems. One of the major bottlenecks in the structural analysis of integral membrane proteins is the production of these proteins in a functional state. In this article the authors describe the use of a comprehensive proteomics and cell physiology-based approach to improve the production of membrane proteins in selected media. By taking advantage of green fluorescent protein (GFP) as a folding indicator and using the differential mobility of folded and unfolded membrane proteins, they have been able to optimize the expression (and subsequent isolation) of correctly folded protein.

**References**


**Authors**

From left: Bert Poolman, Ravi Marredy en Eric Geertsma

**Contact**

Prof. Bert Poolman
Department of Biochemistry
University of Groningen
Nijenborgh 4
9747 AG Groningen
The Netherlands
T +31 50 3634190 or 4209
b.poolman@rug.nl
Disregulation of zinc-dependent metalloproteases function can lead to various disease states like inflammation but also tumour progression. The researchers in this article have found a method to selectively label the active forms of the zinc-dependent metalloproteases. This will allow them to study the role of these metalloproteases in diseases.

Matrix Metalloproteases (MMPs) and A Disintegrin And Metalloproteases (ADAMs) are the two major classes of zinc dependent endopeptidases belonging to the metzincin superfamily (see Figure 1) [1]). The main physiological function of these proteases lies in the modulation of extracellular matrix (ECM), either by direct proteolytic cleavage of ECM proteins (e.g. fibronectin and collagen), or by liberation of membrane-anchored pro-forms of bioactive proteins such as growth factors, cytokines and chemokines [2]. Disregulation of MMP and ADAM function can lead to various disease states involving inflammation (such as arthritis or atherosclerosis) or pathological remodelling of tissue (such as in heart failure), but also tumour progression, such as metastasis and angiogenesis [3].

Metalloprotease catalytic activity is highly regulated in vivo. The enzymes are expressed as inactive proforms, or zymogens, that require proteolytic activation for the mature catalytic enzyme to become available. After the catalytic activity is no longer needed the enzymes can be efficiently inhibited by a family of endogenous inhibitors called TIMPs (Tissue Inhibitor of Metalloproteases).

Metalloprotease analysis The high level of functional regulation leads to problems when using conventional proteomics techniques to assess involvement of metalloproteases in disease, since it can be challenging to discriminate between inactive isoforms and the actual active enzyme. This lack of insight into the actual presence or absence of the biologically active protease in clinical samples may be one of the causes of the often inconsistent results reported in literature. Development of analytical methodology to investigate the role of protease activity in a given disease is therefore necessary to understand their roles.

Traditional methods to determine protease activity all involve some kind of monitoring of substrate (either physiological or synthetic) conversion by the active protease, in either gel-based systems (for instance zymography), assays involving fluorogenic substrates or mass spectrometric measurement of cleaved substrate fragments. Although each method has its advantages, none is particularly suitable for family-wide profiling of active metalloproteases since substrate-specificity is, although quite conserved, not the same for all metzincins.
Metalloproteases are important enzymes in our body. Their main function is the modulation of proteins at the outside of cells. That is, they can cleave proteins, such as collagen, or they can liberate membrane-anchored inactive precursors of proteins, such as growth factors and cytokines, in order to turn them into active proteins. However, when the metalloproteases are not properly regulated, different kinds of diseases can be the consequence such as inflammation (arthritis) and tumour progression (metastasis and angiogenesis).

The enzymes are expressed in their inactive form. In order for the enzyme to become active, a piece of the enzyme has to be removed and this process is highly regulated. If activity of the enzyme is no longer needed, an inhibitor stops it. “Sometimes activation of a certain enzyme of the metalloprotease family can cause damage but it can give protection to another enzyme,” says Theo Klein from the department of Analytical Biochemistry at the University of Groningen. “In order to study the role of the metalloproteases in disease, we are developing a method which with we can detect and extract specifically the active forms out of clinical samples. This way one can investigate whether there is a connection between the presence of the active form and the disease.”

Klein and his colleagues have designed a probe of which the structure is based on inhibitors of the metalloproteases and selectively binds to the active form. The probe has a biotin-label with which its presence can be detected. “Although our first experiments on cultured lung carcinoma cells are promising there are still some hurdles to take,” says Klein. “There are two main types of metalloproteases and our probe does not label one of the two types very well. Furthermore, biotin also occurs naturally in some proteins and therefore interferes with the detection of the labelled metalloproteases. We are now designing a new probe with better labelling qualities which hopefully will help in analyzing clinical samples of diseases like COPD or cancer.”

Chemical proteomics Chemical proteomics is a relatively new approach to the analysis of protein functionality by employing small molecules as probes to monitor interactions with and between proteins. This methodology can also be used for the determination of the functional status of proteases in a given sample by use of activity-specific site-directed molecules as probes, and has been successfully demonstrated for many classes of proteases, such as cysteine proteases [4], serine proteases [5], proteasomes [6] and recently also metalloproteases [1, 7].

Probes for activity-based proteomics (ABP) are typically based on small-molecule synthetic inhibitors of the corresponding protease. Catalytic site-directed probes selectively interact with the active form of the metalloprotease, since the catalytic centre is inaccessible in the inactive zymogen and the TIMP-inhibited form.

**ABP probes** Figure 2 shows an example of a novel probe developed within this project. The probe contains a peptide-like backbone displaying high affinity towards a large range of
MMPs and ADAMs and a hydroxamate moiety to bind the active Zn^{2+} ion in the catalytic centre of the metalloprotease. The backbone structure can be optimized by parallel synthesis of libraries of inhibitors that are subsequently screened against a panel of recombinant metalloproteases to ensure the probe has a sufficiently high affinity towards the enzymes of interest [8].

The probe further contains a UV-reactive cross-linking group for covalent labelling of the catalytic site of the active protease, and a biotin group attached to a spacer arm to enable visualization of the labelled proteases through biotin blotting. The control probe in Figure 2 is structurally identical, except for the missing biotin group. This makes the compound useful as a control compound to exclude non-specific labelling of non-metalloprotease proteins, a phenomenon that may occur in complex samples due to the inherent non-specificity of the UV-activatable photo-crosslinker group.

Labelling of samples with the ABP probe is a straightforward procedure: the probe is added to the aqueous sample at a concentration of 1 µM and the sample is irradiated under a UV lamp at 366 nm for 30 minutes. The sample can then be analyzed by gel-electrophoresis followed by Western blot, where biotinylated proteins are visualized by incubation with conjugated streptavidin. Figure 3a shows covalent labelling of three recombinant active ADAM proteases, demonstrating the proof-of-principle.

Pre-incubation of the ADAMs with their natural inhibitor TIMP-3 negates the labelling in a dose-dependent manner, as does pre-incubation (including UV irradiation) with the control probe, demonstrating that labelling is selective for the active form, and site-specific. Experiments to determine whether the labelling site is actually in the zinc-binding region of the catalytic pocket by mass spectrometry are underway. One interesting observation from the labelling experiments with recombinant metalloproteases is the fact that although ADAMs are labelled, the probe doesn’t appear to react as efficiently with MMPs. This is surprising since the inhibition efficacy of the probe is similar or even higher for MMPs. The next step in optimizing this probe for family-wide labelling is to modify the position of the photo-crosslinker moiety to improve labelling of MMPs.

Figure 3b shows the result of a labelling experiment on a lysate of cultured A549 lung carcinoma cells. Earlier experiments showed that these cells express a large range of MMP and ADAM proteases and could therefore be an interesting biological sample to test our methodology on. Incubation of the sample with the ABP probe results in several additional bands appearing on the anti-biotin blot, indicating that multiple proteins are labelled. The blot further shows three strong bands in the non-incubated control sample, these correspond to endogenously biotinylated proteins that also interact with the conjugated streptavidin used in the staining procedure.

**Alternative detection** The presence of a biotin group on the probe also enables labelled protein to be pulled-down by incubation with immobilized (strept)avidin. This enrichment of biotinylated proteins allows alternative methods of detection of labelled proteases, by for instance Western blotting with immunostaining, or by identification through mass spectrometric techniques. Western blot analysis of the pulled-down fraction in a labelled lysate showed that endogenous ADAM-10 (known from other experiments to be present in the active form in A549) is labelled.

Preliminary experiments with ChipLC-nanoESI-MS/MS identification of pulled-down proteins from ABP-labelled A549 lysates show large interference of endogenously biotinylated proteins that are also pulled down onto the streptavidin-beads. These proteins (mainly involved in energy metabolism) are highly abundant in the A549 carcinoma cells and make identification of actual probe-labelled proteins difficult. The problem of endogenous biotinylated may be overcome by pre-clearing of the samples by immobilized streptavidin, but care has to be taken not to lose the active metalloproteases in the process.
Another approach that is currently being pursued is the use of fluorescent dyes as reporter molecules instead of biotin. Two zinc-dependant proteases have been identified in a pull-down experiment; Meprin β, a disintegrin-metalloprotease and a close relative to the ADAM proteases, and ADAMTS-16, a thrombospondin motif containing soluble form of the ADAM family.

**Future outlook**  The activity-based probe that was developed in this project is capable of covalently tagging active ADAM proteases. Pre-incubation or competition experiments with the natural inhibitor TIMP-3 and a control ABP without reporter molecule show that this labelling is selective for the active form, and site specific. Interestingly, MMPs are labelled to a far lesser extent while the inhibition constant, used as a measure for affinity, is equal or even lower for many MMPs than for the tested ADAMs. We are presently investigating a new generation of probes that has the photo-crosslinker group on a different part of the probe in order to improve labelling.

The next challenge that this project faces is to find suitable biological or clinical models to validate the applicability of the chemical proteomics approach for the analysis of metalloprotease activity. Stimulation experiments on cultured cells are ongoing to provide well-controlled biological samples. This approach may be invaluable in shedding light on the often complex mechanisms behind activation of certain metalloproteases in biological systems. A greater challenge lies in obtaining high quality clinical samples to study the role of metzincin activity in diseases such as COPD or cancer.

**Acknowledgement**
This work was carried out in the Department of Analytical biochemistry lead by Prof. Rainer Bischoff at the University of Groningen with co-supervision of Professor Henk Kauffman, and the Bio-organic Synthesis group lead by Professor Hermen Overkleeft at the University of Leiden. This project is funded by Dutch technology foundation STW (project GPC 6150).

**Summary**
Zinc-dependent metalloproteases are highly regulated catalytically active endopeptidases that play an important role in the modulation and regulation of the extracellular matrix. Because of the high level of regulation, analysis of the functionality of these proteins and their involvement in disease poses problems. In this project we aim to use chemical proteomics methodology to elucidate the activity status of metalloproteases in biological and clinical samples to gain better insight in the role of MMP and ADAM catalytic activity in primarily pulmonary diseases.

To achieve this goal novel active-site selective probes based on synthetic small molecule inhibitors of MMPs and ADAMs have been developed. These probes contain a photo-reactive crosslinker for covalent tagging of the protease, and a biotin reporter for visualization and affinity pull-down of labelled proteins. Proof-of-principle experiments show that labelling of ADAM proteases can be achieved, and is selective for the active form of the enzyme. Labelling of proteins in biological samples is also possible, but as with every novel technique several practical problems have to be overcome before this technique can be taken to the clinic or used to elucidate the complex roles of metzincins in biological processes.

**References**

**Authors**
Theo Klein            Paul Geurink

**Contact**
Theo Klein, M.Sc.
University Center for Pharmacy
Department of Analytical Biochemistry
P.O. Box 196
9700 AD Groningen, The Netherlands
T +31 50 363 3345
t.klein@rug.nl
www.biomac.nl

*group for visualization and affinity pull-down of labelled proteins. Proof-of-principle experiments show that labelling of ADAM proteases can be achieved, and is selective for the active form of the enzyme. Labelling of proteins in biological samples is also possible, but as with every novel technique several practical problems have to be overcome before this technique can be taken to the clinic or used to elucidate the complex roles of metzincins in biological processes.*
Transcriptional regulation is a mechanism of unprecedented importance in the development and growth of living organisms. The 'molecular motors' that drive transcription are built from various protein components which are interacting with each other in an ingenious way. The importance of protein-protein interactions for protein functioning was recognized a long time ago and numerous reliable in vitro and yeast-based methods have been developed for their analyses. However, studies in a native situation such as a plant are rare.

Our lab is interested in the MADS box transcription factor family consisting of 107 members in the plant model system Arabidopsis thaliana. Plant MADS box proteins are involved in a plethora of developmental processes. Biochemical experiments showed that these proteins are active in a combinatorial manner. In the first part of our NPC project, binary and ternary interactions have been elucidated for all Arabidopsis MADS box proteins by yeast two- and three-hybrid assays, and almost 300 dimers and over 100 multimeric complexes could be identified [1,2].

In these yeast-based systems, two or three proteins are screened directly against each other and hence, only the potential to interact becomes apparent. Which of the identified dimers and complexes will be formed in a native situation — where proteins are post-transcriptionally modified and where often a whole set of potential interactors is available possibly competing for binding — is one of the most important and challenging questions to answer. Furthermore, we wonder if it would be possible to make use of the generated protein interaction data-set to create a bioinformatics tool that predicts interactions and to understand how the exquisite interaction specificity is generated based on sequences that are quite similar across the whole family.

Competition FRET In a living organism various proteins that have the potential to interact and are co-expressed, may compete with each other in protein dimer formation. Comparing the results from our high-throughput yeast-based screenings [1] with the outcome of MADS box protein interaction studies in living plant cells based on FRET (Fluorescence Resonance Energy Transfer) [3], revealed clear differences in interaction affinity for MADS box protein dimerization. This suggests that probably only part of all possible interactions will take place at a particular moment during development in vivo. Thus, it is essential to obtain information about interaction affinities for the MADS box proteins.
What this research is about:

**Predicting protein interactions for a better understanding of the development of plants**

Transcription factors are proteins operating at the start of regulation pathways. They do so by binding to DNA and switching genes on or off. They may tell some genes to start creating a wing on a fruitfly, or a certain number of flowers in a plant. “We are interested in the development of plants,” explains Richard Immink, researcher at Plant Research International in Wageningen. “Why, for example, does a plant start to flower? Why does it form five leaves and not ten? The MADS box transcription factors that we investigate play important roles in these processes.”

Immink’s group investigates the interactions of the transcription factors on a molecular level. Complex formation between these proteins is essential because one protein alone can not reach the DNA in the cells nucleus to activate genes. It needs to form complexes with other proteins, such as dimers, to be able to bind to DNA. Immink has investigated these interactions in yeast and found about 300 different complexes. In this study, the researchers wanted to know which combinations are relevant in real plants. Using a novel method called Competition FRET they found that certain complexes are more stable than others. Sometimes homodimers consisting of two identical proteins were stronger, in other cases two different proteins form a tight complex. Concentration also plays a role.

Using these results, Immink went on to predict complex formation using bioinformatics. A computer program identifies stretches of a transcription factor - a certain order of four to twelve amino acids - which may bind to other stretches of another protein. These motifs are then tested using the data from the wet experiments to predict whether the two proteins with the motif are likely to form a complex. Whether these complexes are actually formed is difficult to predict. “Complex formation is a dynamic process,” says Immink. “Often it depends on coincidence whether the proteins are in each others’ vicinity long enough to form a dimer and bind to DNA.” Accurately predicting all of the interactions between the transcription factor proteins will be possible in the near future, according to the researchers.

Unfortunately, no quantitative information about protein-protein interactions can be obtained from standard FRET measurements. The semi quantitative data that are derived from yeast assays does not seem to accurately reflect the differences in interaction affinity for the proteins under native conditions. Therefore, a collaboration was initiated with Prof. Dorus Gadella from the University of Amsterdam (UvA), aiming to develop a novel method for the analysis of protein-protein interaction competition in living plant cells.

The method named ‘Competition FRET’ is based on the FRET principle that energy can be transferred from one fluorescent molecule (FRET donor) to another fluorescent molecule (FRET acceptor), when the two molecules are in very close proximity and when there is an overlap in their spectra. In standard FRET-based protein interaction studies two proteins of interest are labelled with two different fluorophores (FRET donor and acceptor), followed by transient expression in living cells and measuring the FRET signal [3,4]. A protein-protein interaction event will bring the two fluorophores into close proximity and as a consequence FRET will occur.
In Competition FRET, the proteins under study are each labelled with a different fluorophore, in our case Cyan Fluorescent Protein (CFP), Yellow Fluorescent Protein (YFP), or Red Fluorescent Protein (RFP). Subsequently, the labelled proteins X and Y are transiently co-expressed in living plant cells and FRET is measured between the YFP (FRET donor) and RFP (FRET acceptor) label on the proteins. Subsequently, the third protein labelled with CFP is added (see Figure 1). If this protein does not bind with protein X (see Figure 1a), the FRET signal remains. If it does compete with X for interaction with Y the FRET signal reduces (see Figure 1b). Due to the physical characteristics of these fluorophores, no energy can be transferred from YFP or RFP to CFP, therefore the possible interaction between the YFP labelled protein or the RFP labelled protein with the third protein is not measured. Besides the measurement of the FRET signal the relative concentration of each protein is determined based on the fluorescent intensities, which allows obtaining semi-quantitative data for the interaction strengths.

**Combinations**  The Arabidopsis MADS box transcription factors AGAMOUS (AG) and SEPALLATA3 (SEP3) were selected to test the system, because standard FRET analyses revealed that these two proteins can form a heterodimer and are also both able to homodimerize. Different combinations of labelled AG and SEP3 proteins were co-transfected and analyzed using Competition FRET to determine if either heterodimerization or homodimerization is preferred. The obtained results clearly showed that SEP3 cannot compete with AG homodimerization even at high SEP3 concentrations. In contrast, it is easy to interfere with SEP3 homodimerization by providing the AG heterodimerization partner (see Figure 2).

The developed method to monitor competition for interactions in living cells has shown the importance of quantifying interaction affinities in vivo. It was always thought that plant MADS box proteins prefer to heterodimerize instead of forming homodimers. Our results revealed that this is the case for SEP3, but clearly not for AG, which implies that interaction networks determined by binary screenings in an artificial high-throughput system (yeast) or by in vitro methods can differ from the interactome under native conditions.

In our experiments the competition for interaction was analyzed upon ectopic expression of the proteins in isolated plant cells. The challenge for the future is to improve the sensitivity of the techniques and to implement a method for real-time monitoring of interactions between proteins that are expressed under native conditions in intact tissues. In combination with the latest developments in the field of protein complex analysis using mass spectrometry [5] this will ultimately unravel the dynamics of protein complex formation and the biological relevance of the complexes for the control of developmental pathways.

**Prediction** Having a large data-set for specific MADS box protein-protein interactions on hand and some detailed information about differences in interaction affinities, we were wondering if it would be possible to predict protein interactions based on protein sequences alone. This is especially challenging for large protein families, like the plant MADS box transcription factor family, whose members have relatively high sequence similarity, and for which no information about the protein structures is available from wet experiments. In order to develop such a predictor, we collaborated with bioinformatics Dr. Aalt-Jan van Dijk and Dr. Roeland van Ham.
in the framework of our NPC project and a NBIC (Netherlands Bioinformatics Centre) project running at Plant Research International in Wageningen.

The first step in the approach was to search for short correlated motifs (6 to 10 amino acids) that are over-represented in pairs of interacting protein sequences (see Figure 3) [6]. In the second step all possible combinations of MADS box proteins were analyzed for the presence or absence of these motifs. Subsequently, the interaction data from the wet experiments were used to train a so called ‘Random Forest’ classification algorithm, in order to identify important motifs and to predict for each possible combination of MADS box proteins if there would be interaction or not. The resulting “predictor” appeared to have a good performance and interactions could be predicted with a reasonable specificity for MADS box proteins from other plant species. Furthermore, the approach was successfully applied to obtain interaction predictors for other transcription factor families in both plants and vertebrates [6].

**Novel interactions** Although huge protein interaction data sets have been generated over the last years, the currently available information represents only the tip of the iceberg. Continuously, novel interactions are discovered and reported for model species, showing that still large parts of their interactomes are unknown. From many other non-model species no protein interaction data are available at all, emphasizing the need for novel fast and reliable protein interaction identification tools.

The bioinformatics approach to predict interactions discussed here, is a powerful step towards a robust computational method that can predict complete interactomes based on sequence information only. Besides the possibility of deciphering interaction networks, this method also provides information about the specific parts of the proteins that are involved in the interaction and that are essential for interaction specificity. This is not only interesting from a scientific point of view, but can also be applied to create novel strong interaction domains for biotechnological applications.

### References


### Acknowledgements

Dr. Isabella Nougalli Tonaco, Dr. Aalt D.J. van Dijk and Prof. Dr. Gerco C. Angenent from Plant Research International, Bioscience. Bornsesteeg 65, 6708 PD Wageningen.

Prof. Dr. Theodorus W.J. Gadella Jr. from University of Amsterdam, Swammerdam Institute for Life Sciences, Molecular Cytology. Kruislaan 316, 1098 SM Amsterdam.

### Summary

Development is orchestrated by a large number of signalling proteins that are assembled into specific protein complexes. Transcription factors represent herein an essential class. For a complete understanding of developmental regulation it is of great importance to unravel the composition of the protein complexes and to monitor the spatial and temporal dynamics of the complexes *in vivo*. For example, it is frequently assumed that plant MADS box proteins prefer to heterodimerize instead of forming homodimers. We have successfully developed a method to study competition for protein-protein interactions *in vivo*, which gives semi-quantitative information about interaction affinities.

Implementation of this method for members of the plant MADS box transcription factor family revealed clear differences in interaction partner preferences. The SEP3 protein was shown to be unable to compete with AG homodimerization even at high SEP3 concentrations. In contrast, it is easy to interfere with SEP3 homodimerization by providing the AG heterodimerization partner. In addition, a computational tool has been generated which facilitates the prediction of transcription factor protein-protein interactions based solely on sequence information. Besides the possibility of predicting interactions, the procedure gives insight in the domains that determine interaction specificity. In the near future it should be possible to optimize this kind of algorithms and predict complete interactomes.
Erythrocyte proteome studies are currently hampered by the major abundance of haemoglobin and carbonic anhydrase, which obscures proper quantitative analysis of the remaining 1-2% of the protein mass. Therefore we sought a new method that efficiently removes the highly abundant proteins. We have developed a novel approach based on affinity and ion exchange chromatography that depletes the erythrocyte proteome for hemoglobin and carbonic anhydrase. This double depletion improves the number of identified proteins dramatically.

Erythrocytes have an important function in the gas transport of oxygen and carbon dioxide to and from tissues. For the passage through narrow capillaries, the erythrocyte has a discoid shape while it has also lost its nucleus and cytoplasmic organelles. Thus erythrocyte proteins that have lost their function cannot be newly synthesized due to lack of appropriate organelles. Many inherited and acquired anaemia occur [1], like thalassemias, hemochromatosis or erythrocyte enzymopathies, however many other forms of anaemia are not understood at the molecular level. In our collaboration with the Laboratory of Clinical Chemistry and Haematology (UMCU, Utrecht) we aim to unravel the protein molecular background of yet unknown erythrocyte disorders causing anaemia.

Two-dimensional electrophoresis (2DE), a well established method, has its strength in detection of protein isoforms and degradation products, as these protein species appear as separate spots in the 2D image. On the other hand, 2DE is unsuccessful for low abundant, extreme pI/MW proteins, and hydrophobic proteins. This problem can be solved by using alternatives such as the MudPIT method [2] that utilizes multidimensional chromatography on protein digests. However, protein isoforms and degradation products cannot be detected and quantified easily using a MudPIT or similar method. Therefore we have developed a novel approach, based on chromatographic separation of proteins rather than peptides, which circumvents the above described drawbacks of both the 2DE and MudPIT approaches [3].

Haemoglobin depletion To investigate the erythrocyte at the proteome level, we had to tackle two problems. Firstly, the erythrocyte proteome is heavily dominated by two proteins, haemoglobin and carbonic anhydrase. Secondly, the complex proteome should be sufficiently separated to enable accurate quantitative analysis of extreme pI/MW proteins, hydrophobic proteins and protein isoforms.

Thus our first challenge was to deplete the erythrocyte proteome for haemoglobin, while at the same time preventing
What this research is about:

**Efficiently removing hemoglobin from erythrocytes for a clearer proteome analysis**

Many patients are hospitalized each year with a deficiency in their red blood cells; an anaemia. This condition may be acquired or inherited, and in the latter case often involves children who fall seriously, and sometimes fatally, ill. “In about half of the cases we simply do not know what is going on,” says proteomics researcher Jeffrey Ringrose. “These children however are seriously ill. Our research may help to find a cause for each form of anaemia. This stimulates me to do this research.” Studying the proteins in red blood cells may shed new light on some forms of anaemia.

However, investigation of the many proteins in red blood cells is hampered by the abundance of the protein haemoglobin (97%) and carbonic anhydrase (1%) in the samples. This almost fully obscures the quantitative analysis of the hundreds of remaining proteins. Ringrose therefore devised a new method to completely remove these two proteins without losing the others. In this way he identified 700 different proteins instead of the 167 that were found using previous methods. Subsequently he altered an existing set of techniques to separate, identify and quantify all these proteins.

The research is explorative. “We are screening as many proteins as possible and hope to find a difference between the red blood cells of anaemic patients and those of healthy people. Before we were able to remove haemoglobin this was hardly possible.” Currently Ringrose is devising a way to label samples from patients and healthy persons. “Then we will start analysing samples of patients with anaemia of unknown origin. We hope to find a protein which is deviating, for example higher or lower expression level or in another isoform.”

Ringrose’s method for haemoglobin depletion, meanwhile, has even attracted commercial interest. Since the method may be applied to any biological sample that is contaminated by blood, such as lung tissue or even samples in malaria research, a number of companies showed their interest in developing a depletion kit, says Ringrose. “Therefore this result has potential for valorisation.”

Research Theme NPC5: High-throughput separation and array-based technologies
Protein fraction Subsequently we analyzed the haemoglobin depleted soluble protein fraction and the bound fraction with SDS-PAGE. Figure 1 depicts the SDS-PAGE pattern of the haemoglobin-containing fraction in lane 2. The protein pattern of the haemoglobin-depleted fraction was now dominated by the second most abundant protein, carbonic anhydrase-1 (not shown).

The next separation dimension consisted of ion exchange chromatography, and we showed that by choosing appropriate conditions, the carbonic anhydrase-1 could be collected in the flow-through fraction of this separation step. The column-bound proteins were eluted in one step under high-salt conditions to establish the fraction depleted for both haemoglobin and carbonic anhydrase. Figure 1, lane 3, shows the IEX flow-through with the high abundant carbonic anhydrase-1 at 38 kDa, and lane 4 shows the pattern of the lower level proteins in the erythrocyte soluble protein fraction that account for 1-2% of the start sample. Thus the depletion for haemoglobin and carbonic anhydrase-1 has dramatically improved the detection of many other erythrocyte proteins.

Next, we analyzed the proteins from this erythrocyte lysate in all steps of the procedure by LC-MS/MS analysis of the proteins tryptic digests. We established that the number of identified proteins increased from 167 in the non-depleted sample (see Figure 1, lane 1) to 677 in the doubly depleted sample (see Figure 1, lane 4). Furthermore, we detected 20 and 39 proteins in the samples of the haemoglobin and carbonic anhydrase-1 fractions shown in Figure 1, lane 2 and 3, respectively. These depleted proteins are obviously not lost for analysis. Instead, these fractions can be included in the differential analyses.

Combining all data, we detected in total 700 unique proteins, which is an important increase compared to the 167 proteins only in the undepleted erythrocyte sample.

Quantitative proteome analysis Our final goal is to perform quantitative differential analyses of erythrocyte samples of patients with unknown inherited or acquired anaemia versus healthy controls. Therefore, multidimensional chromatography was set up to separate the proteins in the haemoglobin-depleted erythrocyte, using steps based on variation in protein net charge in one dimension and protein hydrophobicity in the next dimension.

The first step consists of ion exchange chromatography, for which we equipped an Ettan LC system with two ion exchange columns connected in series, for strong cation and anion exchange, or SCX and SAX, respectively. The proteins were sequentially eluted using salt gradients, first from the SCX exchange column while bypassing the SAX column and then vice versa from the SAX column. As shown in the elution plots in Figure 3a, this step results in a nice separation of the erythrocyte proteins. The SDS-PAGE patterns of the collected protein fractions are depicted beneath the plots, showing again that carbonic anhydrase-1 is collected in the flow-through. Importantly, the complexity of the fractions is vastly reduced, when compared to the pattern of Figure 1 lane 4.

Analysis with LC-MS/MS on several bands showed, however, that many proteins could be identified per band. To further enhance the resolution, we used a reversed phase column as second separation dimension. The result of this reversed phase separation dimension is illustrated for two SAX fractions in Figure 3b, which shows again a very good resolution of the protein patterns.

In combination with quantification tools such as protein fluorescent labelling [5], this method is very well suited for differential analysis of erythrocyte proteomes. Two erythrocyte soluble

Figure 3 | Multidimensional separation of haemoglobin depleted erythrocyte lysate. A Haemoglobin depleted erythrocyte lysate was fractionated in series by cation- and anion exchange chromatography. B All IEX-fractions were separated in the second dimension by reversed phase chromatography, the result is illustrated for two fractions. The proteins in all these fractions were separated by SDS-PAGE, as shown directly under the plots.
protein samples can be labelled using Cy3 and Cy5, which permits mixing of the two samples prior to the chromatography steps. The advantage of this approach is that differences in protein levels due to experimental variation are eliminated, thus relative protein quantification is more accurate. An additional advantage is that expensive analysis time with mass spectrometry is restricted to the protein fluorescent bands that show significant changes in protein level.

Conclusion We aimed to develop a proteome approach suited for differential analysis of erythrocyte samples from patients with anaemia of unknown origin and from healthy controls, which may provide new diagnostic marker proteins.

For this in-depth proteome research, we were faced with the challenge to deplete the erythrocyte proteome for the high abundant haemoglobin and the second most abundant carbonic anhydrase. A successful depletion approach was developed, showing that comprehensive analysis was only possible through depletion of these two proteins, as detection of the low level proteins has improved more than four-fold.

The setup of the two-dimensional chromatographic separation method, using first in-series SCX and SAX, followed by reversed phase separation, shows well-separated protein patterns. We chose for separation of proteins rather than of their peptides, since analysis of the peptides enhances the complexity of the sample approximately 20-50 times. This two-dimensional chromatographic approach allows separate detection and quantification of protein isoforms, and proteins with extreme pl or MW, which is an advantage compared to 2DE and MUDPIT approaches.

We have shown in other experiments, e.g. on lung tissue, that this very efficient depletion technique can in fact be applied to any biological sample that is contaminated by blood. Also, the erythrocyte is an important cell in the malaria life cycle, while malaria annually causes the death of a few million people worldwide. Therefore, this approach is applicable for many related proteomic projects. This result has potential for valorisation, and indeed, a number of companies showed their interest in developing a depletion kit.

References

Authors
Jeffrey Ringrose, M.Sc
Utrecht University
Bijvoet Center for Biomolecular Research and UIPS
Sorbonnelaan 16
3584 CA Utrecht
The Netherlands.
T +31 30 253 9554
e-mail: j.h.ringrose@uu.nl

Contact
Jeffrey Ringrose and Monique Stijper

Summary
Many human erythrocyte disorders exist with to date unknown ethiology. Erythrocyte proteome research is capable of unraveling the background of these disorders at the molecular level and may provide new biomarkers for diagnosis. Erythrocyte proteome studies are, however, currently hampered by the major abundance of haemoglobin and carbonic anhydrase, which obscures proper quantitative analysis of the remaining 1-2% of the proteome [6]. Secondly, the complex proteome should be sufficiently separated. We developed a simple affinity chromatography approach that depletes the erythrocyte proteome for haemoglobin, which accounts for approximately 97% of the soluble erythrocyte proteome. The subsequent step, ion exchange chromatography, allows for depletion of the second highly abundant protein, carbonic anhydrase. This ion exchange chromatography step permits simultaneously the separation of the proteins that are expressed at lower levels. The last steps in this protein separation workflow consist of reversed phase chromatography and SDS-PAGE. Compared to SDS-PAGE separation as single step, the double depletion improves the number of identified proteins more than four times to approximately 700 proteins. In combination with quantification methods such as fluorescent- and/or stable isotope labelling we can now establish the protein differences that are related to erythrocyte disorders of yet unknown origin.
Primary brain tumours possess a dense network of blood vessels, which makes it an interesting target for drug development and therapy. With a special laser micro-dissection technique to obtain accurate samples the authors applied advanced proteomic techniques, as described in this article, to search for proteins present only or at an elevated level in tumour blood vessels. They were able to identify and validated four new proteins that are exclusively expressed in blood vessels of brain tumours.

Gliomas are the most frequently observed primary brain tumours in adults with a reported annual incidence of 5 to 10 per 100,000 in the Western world. The prognosis of patients with a glioma depends on the histology (class and grade), clinical parameters and genetic aberrations. Unfortunately gliomas cannot be cured and all patients eventually succumb to this disease. With current treatment modalities including surgery, radiotherapy and chemotherapy, the prognosis of glioma patients is dismal with an overall median survival of less than one year. Experimental therapies based on nanoparticles, gene therapy, brain specific drug delivery strategies or specifically targeted drugs are still in a developing phase.

A search for new treatment and drug targeting modalities is clearly warranted as is the search for serum biomarkers to monitor disease activity and drug response. Until now, almost no clinically relevant biomarkers or therapeutic targets have been detected in gliomas. One of the reasons is the technical problem of addressing the huge range in concentration of the various proteins in tissue and a blood samples. In addition, the concentration of relevant target proteins may be femtomolar or below.

For glioma specific delivery of drugs, knowledge is needed about the neo-vasculature of these tumour tissues. In this article we describe the detection and validation of a few proteins that are specific to glioma neo-vasculature by advanced mass spectrometry. We intend to identify more of these specific proteins for the specific targeting to glioma tissue in the near future. Our approach is multidisciplinary and involves neurosurgeons, neuro-oncologists, pathologists, chemists and biotech and pharmaceutical companies.

Laser micro-dissection Technical improvements in sample preparation make it more and more possible to find also low abundant proteins in a tissue. We reasoned that it would be advantageous to take specific cells of the tumour obtained during brain surgery. In this case specific proteins in the affected tissue are at the highest concentration and not diluted
There is still no real cure for a brain tumour, and the prognosis is extremely poor. On the other hand we see a difference in how patients respond to therapy resulting in a longer life expectancy,” explains Theo Luider of the Department of Neurology at the Erasmus Medical Center in Rotterdam. “Since this difference also involves a genetic component it gives us hope that we can improve treatment. When we are able to find the right knobs to turn we might be able to interfere and stop the process of tumour development. Therefore we first have to understand how a tumour develops in the brain, on DNA level and on a protein level.”

In this article Luider describes the search for these knobs to tackle the Achilles’ heel of brain tumours: the blood vessel. In order to grow a tumour needs - just like any other body tissue – a supply of nutrients, which are delivered by the blood vessels. “If we specifically can wipe out the tumours’ blood vessels we eliminate indirectly the tumour. Therefore we searched for proteins that are only present in the blood vessel cells of the tumour and absent in normal blood vessel cells.” To exactly pick out the cells of interest the researchers developed a laser micro dissection technique, by which they were able to cut out a very small piece of tumour blood vessel of about 1000 cells.

So far Luider and co-workers have already identified four proteins that are unique for the blood vessels in brain tumour. “With these results we went back to the original tumour tissue to see if our findings made sense. Whether the proteins we found are really present in the tumour’s blood vessel and not in unaffected tissue. This validation is a very important step.” The results look promising, but much work is still to be done. Luider has a whole list of proteins waiting for further identification.

**What this research is about:**

**Blood vessels targeted to tackle brain tumours**

“There is still no real cure for a brain tumour, and the prognosis is extremely poor. On the other hand we see a difference in how patients respond to therapy resulting in a longer life expectancy,” explains Theo Luider of the Department of Neurology at the Erasmus Medical Center in Rotterdam. “Since this difference also involves a genetic component it gives us hope that we can improve treatment. When we are able to find the right knobs to turn we might be able to interfere and stop the process of tumour development. Therefore we first have to understand how a tumour develops in the brain, on DNA level and on a protein level.”

In this article Luider describes the search for these knobs to tackle the Achilles’ heel of brain tumours: the blood vessel. In order to grow a tumour needs - just like any other body tissue – a supply of nutrients, which are delivered by the blood vessels. “If we specifically can wipe out the tumours’ blood vessels we eliminate indirectly the tumour. Therefore we searched for proteins that are only present in the blood vessel cells of the tumour and absent in normal blood vessel cells.” To exactly pick out the cells of interest the researchers developed a laser micro dissection technique, by which they were able to cut out a very small piece of tumour blood vessel of about 1000 cells.

So far Luider and co-workers have already identified four proteins that are unique for the blood vessels in brain tumour. “With these results we went back to the original tumour tissue to see if our findings made sense. Whether the proteins we found are really present in the tumour’s blood vessel and not in unaffected tissue. This validation is a very important step.” The results look promising, but much work is still to be done. Luider has a whole list of proteins waiting for further identification.

**Research Theme NPC6: Biomarker Discovery**

---

**Figure 1 | Process of laser micro-dissection.**

A Tissue sections are mounted on a specific polymer coated glass slide. By a laser specific cells can be cut and catapulted into a cap and subsequently process for proteomics analysis.

B Example of histology of a proliferated endothelial blood vessel surrounded by brain tumour cells.

C A laser micro-dissection device.

D Frozen tissue sections for and (E) after laser microdissection of tumour vessels.
as is the case in serum samples. We took cells of the neo-vasculature of the tumour by laser micro-dissection (see Figure 1). This microscopic technique gives the possibility to cut cells by a laser from a tissue section and harvest these cells directly into a small vial. By knowing exactly the area that is cut in the tissue section a quite precise estimate can be made how many cells are used for the analyses and prevent bias in the compared protein content in the various samples.

By a technology developed by our research group we can subsequently in situ digest these cells with an enzyme (trypsin) and measure these peptide mixture directly in the MALDI fourier transform mass spectrometer (Bruker Daltonics) (see Figure 2). In this case a relative small number of an equivalent of 1000 cells is needed to obtain an informative mass spectrum. The mass spectra obtained from the glioma neo-vasculature are compared to normal blood vessels and surrounding normal brain tissue and contained more than 2000 mono isotopic peaks. By using home-made software the MALDI-FT ICR spectra can be compared and peptides can be searched for by univariate and multivariate analyses. In these analyses we use the intensity of the observed peptides because the relative standard deviation of the intensities is in the range of 10%. In addition, the intensities show a good correlation coefficient and dynamic range over three orders of magnitude.

**Exclusive proteins** Although the more than 2000 mono isotopic peaks represent probably just a few hundred proteins, we found peptides that were almost exclusively present in the neo-vasculature of the brain tumours. We then laser micro-dissected a relatively large number of cells (n > 10,000) from corresponding tissue sections and analyzed the proteome by offline nano liquid chromatography combined with MALDI TOF/TOF (Bruker Daltonics) and nanoLC (C18) Orbitrap (Thermo Fisher). Using the excellent mass accuracies and superb sensitivity of the MALDI FTICR we were able to link the identified proteins to the differentially expressed peptides observed in the initial analysis. With this approach we identified amongst others fibronectin, colligin-2 and calponin.

Fibronectin has been previously described as a protein that is only found in the neo-vasculature of glioma. The other two proteins were not linked to endothelial cells in tumour blood vessels previously. With immuno-histochemistry we could validate that colligin-2 is exclusively present in the neo-vasculature of glioma. The validation of calponin with specific antibodies is hampered by the lack of calponin specific antibodies that can be used in immuno-histochemistry. Until now, differentially expressed peptides and their related proteins have been validated using immuno-histochemistry or immuno-assays. This approach requires the availability of commercial antibodies or the production of these specific antibodies.

**Femtomole detection** New developments in mass spectrometry that enable the quantification of peptides of interest directly in the mass spectrometer obviate the use of specific antibodies. Developments in MALDI/ ESI triple quad mass spectrometry allow the enrichment and detection of peptides of interest to multi reaction monitoring method without using antibodies. The sensitivity of this technology can be as good as 1-10 ng peptide/ml in the ESI mode. In the MALDI mode, femtomoles were observed in a prototype Maldi triple quad mass spectrometer. This technology is fast and many peptides of interest can be monitored in relatively short, below one minute, measuring times. A newly launched commercial Maldi triple quad (ABI) has been purchased and will also be used for this application.

We conclude from these experiments that it is possible to find proteins by advanced mass spectrometry that relate to the neo-vasculature of brain tumours. It is our impression that many more specific proteins can be found if the sensitivity of the system could be increased. In collaboration with Dionex, Amsterdam, we develop new chromatography materials to enrich low-abundant proteins.

**Increase sensitivity** To tackle the problem of validating large series of differentially expressed peptides we developed technology on a prototype MALDI triple quad mass spectrometer. This technology allows the rapid (less than one minute measurement time per sample) quantitation of known peptides and other compounds in the femtomolar range. Our ultimate goal is the identification of proteins that can be used to target drugs to gliomas.
The neo-vasculature is an obvious target to reach gliomas because these tumours have a dense neo-vasculature network. This neo-vasculature is probably one of the easiest ways to direct drugs to the tumour cells. Membrane proteins and intracellular enzymes that are specific for the neo-vasculature are the key proteins we are searching for. We expect that the described improvements in mass spectrometry will allow identification of such protein. We subsequently intend to translate this know-how into better treatment modalities for glioma.

References
7 Titulaer, M.K. et al. (2006) A database application for pre-processing, storage and comparison of mass spectra derived from patients and controls. BMC Bioinformatics 7, 403

Summary
The identification of proteins related to formation of blood vessels (angiogenesis) is important for the development of new therapies, and such proteins are potential new biomarkers for gliomas. Gliomas are the most common primary brain tumour. The aim of this study was to identify proteins that are exclusively present in newly formed blood vessels of the glioma and not in the vasculature of normal brain. By advanced proteomic techniques we compared the expression profiles of micro-dissected blood vessels from glioma with blood vessels of normal control brain tissue. We measured the enzymatic generated peptide profiles from these micro-dissected cells by matrix-assisted laser desorption/ionization Fourier transform mass spectrometry (MALDI-FTMS). Subsequently, the samples were fractionated by nano-LC prior to matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF/TOF) and online nano LC Orbitrap mass spectrometry. This combined approach enabled us to identify four proteins that appeared to be exclusively expressed in glioma blood vessels. Two of these proteins, fibronectin and colligin 2, were validated on tissue sections using specific antibodies. We found that both proteins are present in the active angiogenesis of glioma and reactive conditions in which neo-angiogenesis takes place.

The approach shows that gel-free mass spectrometric techniques can be used on relatively small numbers of cells (about 1000 cells) generated by micro-dissection procedures to successfully identify differentially expressed proteins.
In this NPC Highlights we provide a short list of papers that appeared in some of the top journals in between the summer of 2007 and the spring of 2008 to which NPC participants contributed. With the guarantee of being by far not comprehensive, this overview provides some very elegant ground-breaking research. To highlight a few: Veenhoff contributed to studies on the molecular architecture of the yeast nuclear pore complex, which lead to two back-to-back publications in Nature. Timmers and Heck introduced a novel on SILAC based method to study specifically the dynamics of protein complexes. The group of Poolman reported on the optimization and quality control of membrane protein expression. Moreover, the NPC was involved in the initiation and public announcement of the HUPO Stem Cell Initiative and in defining standards for reporting proteomics data in scientific journals.

**Quality control of overexpressed membrane proteins**

Geertsma ER, Groeneveld M, Slotboom DJ, Poolman B.


Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, The Netherlands.

Overexpression of membrane proteins in Escherichia coli frequently leads to the formation of aggregates or inclusion bodies, which is undesirable for most studies. Ideally, one would like to optimize the expression conditions by monitoring simultaneously and rapidly both the amounts of properly folded and aggregated membrane protein, a requirement not met by any of the currently available methods. Here, we describe a simple gel-based approach with green fluorescent protein as folding indicator to detect well folded and aggregated proteins simultaneously. The method allows for rapid screening and, importantly, pinpointing the most likely bottlenecks in protein production.

**High-throughput cloning and expression in recalcitrant bacteria**

Geertsma ER, Poolman B.


Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, the Netherlands.

We developed a generic method for high-throughput cloning in bacteria that are less amenable to conventional DNA manipulations. The method involves ligation-independent cloning in an intermediary Escherichia coli vector, which is rapidly converted via vector-backbone exchange (VBEx) into an organism-specific plasmid ready for high-efficiency transformation. We demonstrated VBEx proof of principle for Lactococcus lactis, but the method can be adapted to all organisms for which plasmids are available.

**The molecular architecture of the nuclear pore complex**


Department of Bioengineering and Therapeutic Sciences, and California Institute for Quantitative Biosciences, Mission Bay QB3, 1700 4th Street, Suite 503B, University of California at San Francisco, San Francisco, California 94158-2330, USA.

Nuclear pore complexes (NPCs) are proteinaceous assemblies of approximately 50 MDa that selectively transport cargoes across the nuclear envelope. To determine the molecular architecture of the yeast NPC, we collected a diverse set of biophysical and proteomic data, and developed a method for using these data to localize the NPC’s 456 constituent proteins (see the accompanying paper). Our structure reveals that half of the NPC is made up of a core scaffold, which is structurally analogous to vesicle-coating complexes. This scaffold forms an interlaced network that coats the entire curved surface of the nuclear envelope membrane within which the NPC is embedded. The selective barrier for transport is formed by large numbers of proteins with disordered regions that line the inner face of the scaffold. The NPC consists of only a few structural modules that resemble each other in terms of the configuration of their homologous constituents, the most striking of these being a 16-fold repetition of “columns”. These findings provide clues to the evolutionary origins of the NPC.
Affinity purification in combination with isotope labeling of proteins has proven to be a powerful method to discriminate specific from non-specific interactors. However, in the standard SILAC (Stable isotope labeling by amino acids in cell culture) approach dynamic components may easily be assigned as non-specific. We compared two affinity purification protocols, which in combination revealed information on the dynamics of protein complexes. We focused on the central component in eukaryotic transcription, the human TATA binding protein (TBP), which is involved in different complexes. All known TBP associated factors (TAFs) were detected as specific interactors. Interestingly, one of them, BTAF1, exchanged significantly in cell extracts during the affinity purification. The other TAFs did not display this behavior. Cell cycle synchronization showed that BTAF1 exchange was regulated during mitosis. The combination of the two affinity purification protocols allows a quantitative approach to identify transient components in any protein complex.

The minimum information about a proteomics experiment (MIAPE)


Both the generation and the analysis of proteomics data are now widespread, and high-throughput approaches are commonplace. Protocols continue to increase in complexity as methods and technologies evolve and diversify. To encourage the standardized collection, integration, storage and dissemination of proteomics data, the Human Proteome Organization’s Proteomics Standards Initiative develops guidance modules for reporting the use of techniques such as gel electrophoresis and mass spectrometry. This paper describes the processes and principles underpinning the development of these modules; discusses the ramifications for various interest groups such as experimentalists, funders, publishers and the private sector; addresses the issue of overlap with other reporting guidelines; and highlights the criticality of appropriate tools and resources in enabling ‘MIAPE-compliant’ reporting.

An experimental correction for arginine-to-proline conversion artifacts in SILAC-based quantitative proteomics

Van Hoof D, Pinkse MW, Oostwaard DW, Mummery CL, Heck AJ, Krijgsve J.

Netherlands Institute of Developmental Biology, Hubrecht Laboratory, Uppsalaalan 8, 3584 CT Utrecht, The Netherlands.


In the application of SILAC labeling to embryonic stem cells, we noted substantial conversion of labeled arginine to proline, thereby preventing proper protein quantitation in these cells. We have therefore designed a novel method that circumvents the problem of Arginine-to-Proline conversion, while maintaining Arginine levels and thus preventing any undesired side effects related to Arginine deprivation. The principle is based on the use of two different forms of isotope-coded Arginine ([15N3]-Arg and [13C6,15N4]-Arg) for the ‘light’ and ‘heavy’ conditions, respectively. Although both will be converted to Proline, this will occur at the same rate, thus providing an internal correction for Arginine conversion. This now offers the possibility for accurate protein quantification in embryonic stem cells.

Structural biology of RNA polymerase III: mass spectrometry elucidates subcomplex architecture

Lorenzen K, Vannini A, Cramer P, Heck AJ.


RNA polymerases (Pol) II and III synthesize eukaryotic mRNAs and tRNAs, respectively. The crystal structure of the 12 subunit Pol II is known, but only limited structural information is available for the 17 subunit Pol III. Using mass spectrometry (MS), we correlated masses of Pol II complexes with the Pol II subunit Pol II is known, but only limited structural information is available for the 17 subunit Pol III. Using mass spectrometry (MS), we correlated masses of Pol II complexes with the Pol II subunit form lacking the Pol III-specific subcomplex C53/37. DMSO treatment dissociated the C17/25 heterodimer of Pol III, confirming a peripheral location as its counterpart in Pol II. Tandem MS revealed the Pol III-specific subunits C82 and C34 dissociating as a heterodimer. C11 was retained, arguing against a stable trimeric subcomplex, C53/37/11. These data suggest that Pol III consists of a 10 subunit Pol II-like core; the peripheral heterodimers C17/25, C53/37, and C82/34; and subunit C31, which bridges between C82/34, C17/25, and the core.
Identification of glioma neovascularization-related proteins by using MALDI-FTMS and nano-LC fractionation to microdissected tumor vessels


Department of Pathology, Erasmus Medical Center, Dr. Molewaterplein 50, 3015 GD Rotterdam, The Netherlands.

The identification of angiogenesis-related proteins is important for the development of new antiangiogenic therapies, and such proteins are potential new biomarkers for gliomas. The aim of this study was to identify proteins that are exclusively present in glioma neovasculature and not in the vasculature of normal brain. We combined advanced proteomics techniques to compare the expression profiles of microdissected blood vessels from glioma with blood vessels of normal control brain samples. We measured the enzymatic generated peptide profiles from these microdissected samples by MALDI-FTMS. Subsequently, the samples were fractionated by nano-LC prior to MALDI-TOF/TOF. This combined approach enabled us to identify four proteins that appeared to be exclusively expressed in the glioma blood vessels. Two of these proteins, fibronectin and colligin 2, were validated on tissue sections using specific antibodies. We found that both proteins are present in active angiogenesis in glioma, other neoplasms, and reactive conditions in which neoangiogenesis takes place. This work proves that gel-free mass spectrometric techniques can be used on relatively small numbers of cells generated by microdissection procedures to successfully identify differentially expressed proteins.

Other highlighted publications

Determining the architectures of macromolecular assemblies

Endogenous phosphotyrosine signaling in zebrafish embryos

A high-quality catalog of the Drosophila melanogaster proteome

Direct observation of chaperone-induced changes in a protein folding pathway

Arabidopsis JACKDAW and MAGPIE zinc finger proteins delimit asymmetric cell division and stabilize tissue boundaries by restricting SHORT-ROOT action

The BR1 Associated receptor Kinase 1, BAK1, has a brassinolide-independent role in plant cell death control

The fluxes through glycolytic enzymes in Saccharomyces cerevisiae are predominantly regulated at posttranscriptional levels

Proteome biology of stem cells: a new joint HUPO and ISSCR initiative

The time is right: proteome biology of stem cells
Proteomics and Biomarkers at HUPO 6th Annual World Congress 2007

Han Roelofsen (UMC Groningen)

The 6th Human Proteome Organization conference was held from 6 to 10 October 2007 at the COEX Centre in the heart of the business district of Seoul, Korea. It was organized by Young-Ki Paik, Richard Simpson and Young Mok Park. The theme this year was how to go from technology development to biomarker application.

The buzz words of the conference were targeted proteomics and validation. There is more and more interest in targeted proteomics approaches to find markers. It is now realized that, without a special toolbox, the dynamic range of plasma is far too large to get down to the interesting proteins that have biomarker potential. This is illustrated by the appeal by Gil Omenn, the head of the HUPO Plasma Proteome Project, to groups that have data on tissue/organ secretomes to share them with the HUPO-PPP to extend their plasma data and, additionally for validation.

Glyco-proteomics is another important topic in the biomarker field. It was covered as part of the conference in a session by the HUPO Glyco-Proteomics Initiative and by a symposium. Glycosylation of proteins is altered in cancer cells. Therefore, glyco-proteins may be important cancer biomarkers.

Our group represented the NPC research theme 6 on biomarker discovery by presenting two posters and one oral presentation of our work on adipose tissue. Adipose tissue is not only a fat storage depot but is also an active endocrine organ involved in regulation of whole-body energy metabolism via secretion of various peptide hormones (adipokines). As such, it is an important source of biomarkers for insulin resistance, metabolic syndrome and type-2 diabetes. One of our posters on a new method for quantifying insulin-related changes in the adipose tissue secretome won the poster prize.

The meeting was closed by a speech by the President of the HUPO, Rolf Apweiler, followed by a presentation by our NPC president Albert Heck who introduced the 7th HUPO congress to be held from 16 to 20 August 2008 in Amsterdam.

The meeting was rounded off by a dinner during which the congress attendants were treated on a special show consisting of old Korean dances followed by a compilation of parts of Broadway musicals to emphasize the old and the modern Korea.

We cordially invite you to attend and contribute to the major proteomics event of 2008, the HUPO 7th Annual World Congress.

The theme of the meeting is “Proteome Biology” and reflects the need for proteomics to progress beyond identifying the building blocks of the proteome towards understanding of how the complex biological functions of a cell are orchestrated.

We look forward to seeing you in Amsterdam!

On behalf of the Congress Chairs,
Albert Heck
At the interface of proteomics and bioinformatics

The Protein and Nucleotide Database (PANDA) group is a large group providing all the sequence resources at the EBI, from DNA through to Protein and encompassing the associated databases, such as genome databases (eg, Ensembl), protein families (eg, InterPro), protein function (eg, UniProtKB), proteomics (eg, IntAct and PRIDE), ontologies (eg, GOA), and pathways (eg, Reactome).

Of interest to all proteomics researchers, and thus the NPC, is the Proteomics Services Team inside PANDA (headed by Henning Hermjakob). This team provides databases and tools for the deposition, distribution and analysis of proteomics and proteomics-related data. We contribute to the development of community standards for proteomics data in the context of the HUPO Proteomics Standards Initiative (PSI), and develop reference implementations for these standards. The PRIDE Proteomics IDENTifications database is a centralized, standards compliant, public data repository for proteomics data. It has been developed to provide the proteomics community with a public repository for protein and peptide identifications together with the evidence supporting these identifications.

In 2006, EBI and NPC started a collaboration on the PRIDE repository, after which a satellite version of this database was setup at the NPC in Utrecht. The main idea here is to further extend PRIDE to become a central part of the proteomics pipeline as developed by the NPC. Additionally effort is put into turning PRIDE-NPC into a full satellite of PRIDE at the EBI. Recently a more formal collaboration with the PRIDE team was initiated to further extend PRIDE such that it can capture quantitative information about peptides and proteins. Here the NPC will both deliver data as well as manpower for development and testing.

At EBI we are very pleased to collaborate with proteomics researchers worldwide. The intense collaboration with the bioinformatics team at the NPC is just starting, but will certainly lead to a fruitful cross-fertilization, and can be seen as a model for future international partnerships at the interface of proteomics and bioinformatics. I wish you success with the second phase of NPC II and look forward to further collaborations.

Dr. Rolf Apweiler
President HUPO
Joint Head of Protein and Nucleotide Data (PANDA) Group
European Bioinformatics Institute (EBI)
Wellcome Trust Genome Campus
Hinxton, Cambridge
United Kingdom