Proteomics Approaches to Study Genetic and Metabolic Disorders

ARTICLE *in* JOURNAL OF PROTEOME RESEARCH · MARCH 2007

Impact Factor: 4.25 · DOI: 10.1021/pr060487w · Source: PubMed

Journal of reviews proteome research

Proteomics Approaches to Study Genetic and Metabolic Disorders

Jolein Gloerich,[†] Ron A. Wevers,[*,†] Jan A. M. Smeitink,[‡] Baziel G. van Engelen,[§] and Lambert P. van den Heuvel[†]

Laboratory for Pediatrics and Neurology, Nijmegen Centre for Mitochondrial Disorders, Department of Pediatrics, Neuromuscular Center Nijmegen, Department of Neurology, Radboud University Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

Received September 18, 2006

Several proteomics approaches to study different aspects of genetic and metabolic diseases are presented. The choice of technique is strongly dependent on the biological question to be addressed and the availability and amount of sample. In general, there are three approaches that may be used to study genetic and metabolic diseases: protein profiling of complex biological samples, identification of affected proteins, or a functional proteomics approach to study protein interactions and function.

Keywords: genetic disease • metabolic disorders • mass spectrometry • protein profiling • protein identification • functional proteomics

Introduction

Sequencing of the human genome revealed approximately 23 000 genes that are believed to encode more than a million different proteins with distinct functional properties. Genetic alterations can lead to aberrant expression and/or function of any of these proteins, which can result in disease. In some genetic and metabolic disorders, the symptoms can be correlated with the activity of a single gene product, whereas in other diseases, the phenotype is determined by the interactions of many proteins.[1,2] Although genomic studies can provide information on differential gene expression, proteomics approaches are needed to explore the effects of genetic defects on protein expression. Besides abundance-based proteomics studies, also post-translational modification (PTM) of proteins, such as phosphorylation and glycosylation, can be studied.[3] Furthermore, functional studies can be performed using proteomics approaches.[4]

When studying the differences in the proteome of the diseased state versus the normal state, it is important to consider several points. First of all, one has to be aware that there are major differences in the proteome of different tissues and cells and that the choice of tissue, cell, or cell organelle is very important. In general, the proteome of tissues that are most affected by the disorder will show most differences between patient and control subjects. Furthermore, when differences in protein expression or activity are observed, it is of importance whether this protein is primarily affected by a mutation in the gene, or if the differential expression is secondary, due to, for example, defective import of proteins into an organelle, defective PTM of the protein in question, or a secondary effect on a metabolic pathway. Finally, if one wants to use a proteomics approach for studying the primarily affected protein, the protein has to be relatively abundant, or extensive purification of the protein is needed. Also, the mutation has to cause a difference in either protein level, isoelectric point, or mass that can be distinguished via the chosen technique.

Many different mass spectrometry (MS) based strategies can be used to study the differences between the healthy and diseased state. The choice of technique is strongly dependent on the biological question to be addressed and the availability and amount of tissues, cells, or body fluids. In general, there are three approaches to study the proteome: protein profiling of complex biological samples, identification of affected proteins, or a functional proteomics approach to study protein interactions and function. Proteomics is already widely used in cancer research, but its techniques are also very well suited to study different aspects of inborn errors of metabolism. Some proteomics studies have been performed to investigate genetic and metabolic diseases. In this review, an overview is given of proteomics approaches that can be used to address different biological questions, and examples of their application in studies on genetic and metabolic disorders are described (Table 1).

Protein Profiling. Differences in protein expression of specific proteins between healthy subjects and patients suffering from a genetic and/or metabolic disorder are classically studied using techniques such as immunoblotting and immunoprecipitation. Although these techniques are highly specific, they do not address the total proteome and they are not suited to study proteomic profiles in complex biological samples. To search for differences in protein patterns, rather than the expression of one particular protein, between patient and control samples, several protein profiling techniques have been developed. Comparative two-dimensional gel electrophoresis

* To whom correspondence should be addressed. Prof. Dr. Ron Wevers, Laboratory of Pediatrics and Neurology, Radboud University Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands; Tel: +31-(0)24-3614567; Fax: +31-(0)24-3668754; E-mail: R.Wevers@cukz.umcn.nl.

[†] Laboratory for Pediatrics and Neurology.

[‡] Nijmegen Centre for Mitochondrial Disorders, Department of Pediatrics.

[§] Neuromuscular Center Nijmegen, Department of Neurology.

10.1021/pr060487w CCC: $37.00
Published on Web 12/07/2006

Table 1. Overview of Described Disorders, the Genetic Defects, the Proteomics Approach, and the Techniques Used

disorder	affected protein	approach	technique
Niemann-Pick C2 disease[a]	HE1	protein profiling	2DGE, Edman degradation
Methylmalonic acidemia[b]	L-methylmalonyl-CoA mutase	protein profiling	DIGE, MALDI-TOF MS
Amyotrophic lateral sclerosis[c]	genetically heterogeneous	protein profiling	SELDI-TOF MS, LC-FT−MS/MS
Gaucher disease[d]	acid beta-glucosidase	protein identification	SELDI-TOF MS
Transthyretin-associated hereditary amyloidosis[e]	transthyretin	protein identification	LC-FT−MS, LC-MS
Congenital disorders of glycosylation type I[f]	genetically heterogeneous	protein identification	SELDI-TOF MS, MALDI-TOF MS, LC-FT−MS/MS
Cystic fibrosis[g]	cystic fibrosis trans-membrane conductance regulator	functional proteomics	antibody microarray
Thrombocytopenic purpura[h]	ADAMTS−13	functional proteomics	SELDI-TOF MS
Fanconi anemia[i]	genetically heterogeneous	functional proteomics	immunoprecipitation, LC-MS/MS

[a] Reference 15. [b] Reference 19. [c] References 26 and 28. [d] Reference 34. [e] References 38−41. [f] References 47−52. [g] Reference 69. [h] Reference 71. [i] Reference 73.

(2DGE) can be used to address proteomic profiles.[5] Furthermore, several MS based methods have been developed to study protein profiles of complex samples, such as surface-enhanced laser desorption/ionization time-of-flight (SELDI-TOF),[6,7] matrix assisted laser desorption/ionization time-of-flight (MALDI-TOF),[8] and capillary electrophoresis electrospray ionization mass spectrometry (CE-ESI-MS).[9] Also, a combination of liquid chromatography, either (nano)-reversed-phase or ion exchange, with MS techniques such as Fourier transform (FT) ion cyclotron resonance mass spectrometry can be used to screen complex samples.[10]

2-D Gel Electrophoresis. In proteomics, 2DGE is one of the oldest and most used separation techniques for complex protein mixtures.[11,12] The separation of proteins is based on both mass and iso-electric point. An advantage of 2DGE lies in its ability to resolve and investigate the abundance of thousands of proteins in a single sample and the possibility to directly detect PTM changes. In addition, the use of difference gel electrophoresis (DIGE)[13,14] allows a quantitative comparison between related samples, such as those from patients and control subjects, using different fluorescent labels for patient and control samples. After attachment of a specific fluorescent dye to each sample, samples are combined and separated on one 2D-gel (see Figure 1, control subject is labeled with Cy3 and patient with Cy5). Protein expression of each sample is visualized by fluorescence imaging using the specific excitation wavelength of each dye. Superimposing the images allows the detection of changes in protein expression between the samples.

A disadvantage of 2DGE is the fact that relatively low abundance proteins are easily missed in complex mixtures such as plasma, where 90% of the total amount of protein consists of a small set of very abundant proteins. Furthermore, 2DGE is a laborious technique, not suited for high throughput proteomic analysis. To identify the differentially expressed proteins on 2DGE, different techniques can be used such as immunoblotting, Edman degradation or MALDI-TOF MS.

An example of the use of 2DGE to identify the primarily affected protein in an inborn error of metabolism is the identification of HE1 as the protein that is deficient in Niemann-Pick C2 disease (NPC2).[15] NPC2 is an autosomal recessive lipid storage disorder, characterized by a defect in intracellular trafficking of exogenous cholesterol that leads to the lysosomal accumulation of unesterified cholesterol.[16,17] To exclusively study soluble lysosomal proteins, these proteins were purified from cell lysates based on the fact they acquire a PTM that distinguishes them from most other types of proteins, the mannose-6-phosphate marker. This PTM is recognized by mannose-6-phosphate receptors, which target the protein to the endolysosomal system. Isolation of soluble lysosomal proteins from cell lysates was performed using purified mannose-6-phosphate receptors to specifically bind these proteins. HE1 was identified as a lysosomal protein in a 2DGE proteomics study directed at characterizing the lysosomal proteome.[15,18] In combination with a study in which the porcine homolog of HE1 was shown to specifically bind cholesterol, this led to the hypothesis that HE1 might be involved in NPC2. Indeed, HE1 was undetectable in NPC2 fibroblasts after immunoblotting and mutations in the gene encoding HE1 were found in several NPC2 patients.

DIGE was used to investigate changes in mitochondrial protein expression of human fibroblasts of control subjects and patients suffering from methylmalonic acidemia (MMA).[19] MMA is an inborn error of metabolism, caused by impaired

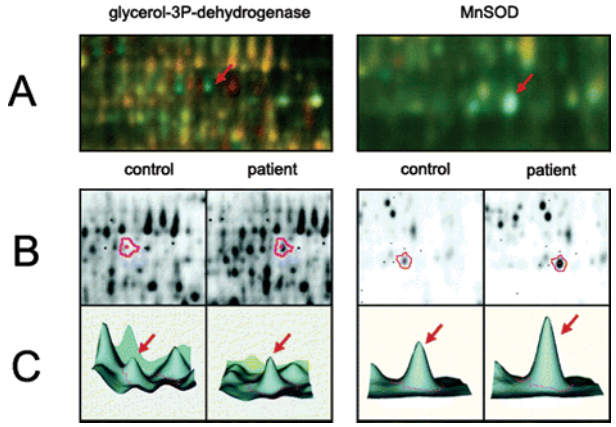

Figure 1. 2-D DIGE image analysis of two differentially expressed mitochondrial proteins, the mitochondrial precursor of glycerol-3P-dehydrogenase and mitochondrial manganese superoxide dismutase (MnSOD). Proteins of mitochondrial fractions of fibroblasts from a control subject and an MMA patient were labeled with Cy3 and Cy5, respectively. These samples were mixed and subjected to 2DGE. (A) Fluorescent image of the gel; Cy3-labeled proteins fluoresce green, Cy5-labeled proteins fluoresce red, and yellow spots indicate the presence of equal amounts of Cy3- and Cy5-labeled protein. In (B) and (C), the left panel corresponds to the image of the Cy3-labeled proteins (control sample) and the right panel corresponds to the Cy5-labeled proteins (MMA patient sample). (B) Separated images of control and patient samples, (C) 3-D view of DeCyder (GE Healthcare) software analysis used to quantify the amount of protein of the spot encircled in (B). Figure adapted from ref 19.

isomerization of L-methylmalonyl-CoA to succinyl-CoA.[20,21] In this study, a set of differentially expressed mitochondrial proteins was identified using MALDI-TOF MS. In contrast to the above-mentioned study on NPC2, where the protein directly related to the disease was identified, the differentially expressed proteins were part of cellular pathways, such as apoptosis and oxidative stress, which are indirectly affected by the metabolic defect and may play a role in the pathogenesis of the disease. In Figure 1, a 2-D DIGE image analysis is shown for two of the proteins that were differentially expressed.

SELDI-TOF MS. A technique that is very suitable for high throughput proteomic analysis of complex mixtures of proteins is SELDI-TOF MS.[6,7] Proteins are retained on solid-phase chromatographic surfaces with specific properties and are subsequently ionized and detected by TOF MS. Separation is based on specific physical properties of the proteins, such as hydrophobicity, charge state, PTMs, etc. With SELDI-TOF MS, only limited sample preparation is needed, and the system is ideally suited for profiling low molecular weight proteins (<20 kDa) in a large variety of complex biological materials, such as serum, blood, plasma, intestinal fluid, urine, and cell lysates. This technique is often used to screen for biomarkers for several diseases and has been applied to the diagnosis of cancer[22−24] and neurological disorders.[25,26] For example, cerebrospinal fluid (CSF) of patients having amyotrophic lateral sclerosis (ALS), a motor neuron disorder that exists in both sporadic and familial forms, was screened to identify biomarkers for this disease.[26] This resulted in identification of 30 mass ion peaks with statistically significant differences between control and ALS subjects. Using an algorithm, a biomarker panel was created with an 89% positive predictive value for ALS. Furthermore, from this panel three candidate biomarkers for ALS were identified as either decreased (transthyretin, cytostatin C) or increased (carboxy-terminal fragment of neuroendocrine protein 7B2). SELDI-TOF results were validated using immunoblotting and immunohistochemistry.

Liquid Chromatography Mass Spectrometry (LC-MS). The combination of (nano)-liquid chromatography (LC) and electrospray ionization (ESI) tandem-MS provides a powerful approach for proteome analysis. Without the need of prepurification, complex protein mixtures can be digested, and the resulting peptides can be directly analyzed by nanoLC-MS/MS. Especially when FT−MS is used as mass spectrometric technique, the combination with nanoLC is very suitable for the identification of low-abundant proteins in complex mixtures, because the resolving power and mass accuracy of FT−MS is extremely high.[27] Furthermore, only very small amounts of sample are needed, especially when compared to 2DGE. Despite these advantages, LC-MS/MS is not ideal for high throughput protein profiling of large numbers of samples, because LC-MS/MS data processing is a very laborious and time-consuming task. However, LC-MS/MS is a very promising technique with respect to the discovery of (low abundance) biomarkers for many diseases.[10]

In another study on CSF of ALS patients, LC-FT−MS/MS was used in combination with a pattern recognition program to reveal peak patterns specific for ALS.[28] First, the authors evaluated this research strategy by comparing the peptide patterns of CSF spiked *in vitro* with myoglobin from horse heart, with control CSF. The patterns of these samples were clearly separated, and the peptides of myoglobin were identified as characteristic peaks. Application of this strategy to CSF from ALS and control subjects did not lead to the identification of a single protein with altered expression, but 80% of the unknown samples could be assigned correctly, based on the specific peak pattern of CSF from ALS patients. Although this approach has to be tested on larger numbers of samples, it indicates that the combination of LC-FT−MS/MS together with pattern recognition software to perform multivariate statistical analysis is a powerful tool in development of diagnostic assays. However, special care should be taken when small sample numbers are used for multivariate analysis to avoid over-fitting of the data.

Another technique that utilizes LC together with MS is differential peptide display (DPD; BioVisioN). This technique generates a comprehensive peptide map of ~3000 peptides covering a mass range of 950−15 000 Da from a biological sample.[29,30] First, the biological sample is digested and peptides are separated by reversed-phase HPLC. Eluting peptides are collected in 96 fractions, which are each subjected to MALDI-TOF MS. The mass spectra of all fractions are combined in a two-dimensional map in which the first dimension is determined by the retention time on the HPLC, and the second dimension is determined by the peak intensity of the MALDI-TOF MS spectrum at a specific m/z value. Tammen et al.[31] has shown that this technique can be successfully applied to screen blood samples for a specific polymorphism of blood coagulation factor XIII, Val34Leu, which correlates with an increased risk for hemorrhagic stroke and lower incidence of myocardial infarction and ischemic stroke.[32,33] This indicates that this technique might also be applicable to reveal new mutations in proteins affected in inborn errors of metabolism, as long as the mutation causes a change in m/z-value and/or retention time on the LC column.

Protein Identification. Besides protein profiling of biological samples to screen for differences between patients and control subjects, identification of proteins that show a disease-specific expression pattern is one of the main goals in clinical proteomics. Often a combination of profiling and protein identification is used either in the search for biomarkers for specific diseases, to identify the protein that is deficient in a disorder, or to assess changes in PTMs. First, a profiling approach is used to quickly screen for differences between patients and control subjects. Subsequently, proteins or peptides that cause these differences in the profile are identified and characterized. In most cases in which a protein or PTM is identified, other methods such as immunoblotting, activity measurements or sequence analysis are used to validate the results of the proteomics approach.

Biomarker Discovery. In recent years, one of the hot topics of proteomics research has been the identification of biomarkers for various diseases. Especially in the field of cancer research, many studies have been performed that use proteomics approaches to search for specific markers for various types of cancers in biological samples to aid in the early detection of these cancers.[22−24] The reason for the interest in proteomics within the oncology field is most obvious: in tumors, mutated genes lead to aberrant protein expression. Because inborn errors of metabolism also originate from mutations in specific genes, the application of proteomics to search for biomarkers for these disorders is a logical step.

Identification of a biomarker for Gaucher disease has been described by Boot et al.[34] Gaucher disease is a lysosomal storage disorder, which is caused by a deficiency in lysosomal glucocerebrosidase activity, leading to an accumulation of glucosylceramide in the lysosomes of macrophages and other cell

types.[35] Accumulation of these glycerolipids causes the clinical symptoms of the disease. Treatment of Gaucher disease consists of enzyme replacement therapy. The availability of sensitive plasma biomarkers for Gaucher disease is of great value for monitoring the efficacy of enzyme replacement therapy.[36] The disease state is monitored by measurement of plasma activity of chitotriosidase, which is exclusively excreted by activated macrophages. However, a frequently occurring recessive inherited chitotriosidase deficiency, which does not cause any clinical symptoms, hampers the use of chitotriosidase activity as a marker for Gaucher disease in some patients.[37] Therefore, SELDI-TOF MS was used to screen for other specific plasma biomarkers to monitor the disease. The chemokine CCL18/PARC was found to be a sensitive and specific marker for Gaucher disease.[34] These results were confirmed by immunohistochemistry and ELISA assays. Thus, CCL18/PARC levels in plasma can serve as an alternative biomarker to monitor Gaucher disease.

Identification of Mutations. Besides the identification of disease-specific biomarkers for genetic disorders, a proteomics approach can also be used to directly study the protein that is mutated. In this way, not only expression of the mutated protein can be investigated, but also the specific mutations in the protein can be recognized, which is shown in several studies on transthyretin-associated hereditary amyloidosis (ATTR).[38−41] Amyloidosis is characterized by deposition of insoluble fibril deposits of normally soluble proteins, which eventually leads to organ dysfunction and ultimately death. In ATTR, the disease is caused by deposits of variant transthyretin (TTR, prealbumin), which is the protein that is affected through mutations in the according gene.[42] To diagnose ATTR, either DNA sequence analysis is done to show mutations in the gene or changes in the protein can be investigated by iso-electric focusing, 2DGE, or MS techniques.[38−41,43,44] When using MS to identify TTR variants, TTR is first isolated from plasma or serum by immunoaffinity purification and reduced by tris(2-carboxymethyl)phosphine (TCEP) to reduce TTR sulfite adducts. The intact protein is then subjected to MS analysis (either LC-FT−MS,[41] or single quadrupole MS[40]) and the mass difference between the wild-type and mutated protein is used to determine the variant TTR as is shown in Figure 2.[40] However, one has to keep in mind that MS analysis is not able to identify amino acid substitutions that do not lead to a mass difference (e.g., Leu to Ile). Furthermore, when analyzing intact proteins, a mass shift of ≤10 Da can only be detected using MS techniques with a high resolving power. If all potential single amino acid substitutions are considered, 12.6%, have a mass shift ≤ 10 Da, which gives a theoretical sensitivity of ∼87% for screening of intact proteins. However, these mutations can easily be detected, when peptide analysis is performed. Fortunately, for the investigation of ATTR by MS analysis of intact TTR, the known TTR variants with a mass shifts of ≤10 Da are not common, and they constitute only 10% of the 100 known variants. Using this setup allows for quick screening for ATTR in a broad patient population and eliminates the necessity of DNA sequencing.[40,41]

Identification of Changes in PTMs. Besides identification of amino acid substitutions in the protein of interest, changes in PTMs of proteins can also be investigated using proteomics. Protein modifications such as methylation, phosphorylation, and glycosylation (both *N*- and *O*-glycosylation) can be detected using MS techniques.[3] Several MS techniques have been employed in diagnostic testing for congenital disorders of

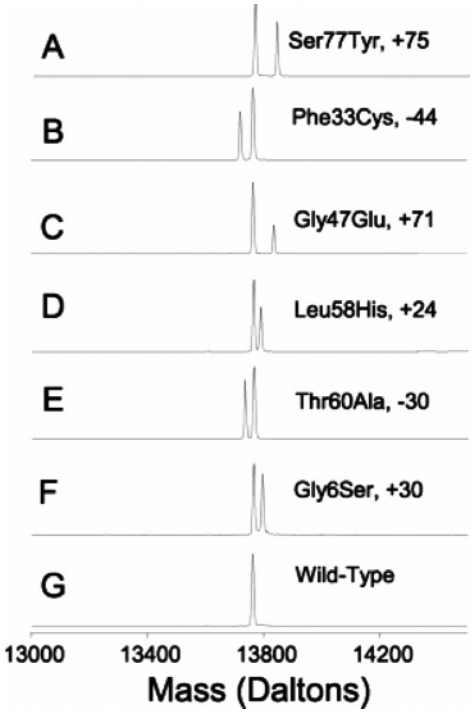

Figure 2. Transformed spectra of TCEP-reduced TTR. The measured mass differences and the actual variants are shown on the right. Reprinted with permission from ref 40.

glycosylation (CDG).[45] CDG constitutes a group of diseases affecting glycosylation pathways. CDG-I results from deficiencies in the early *N*-glycosylation pathway in the endoplasmatic reticulum (ER). This concerns the biosynthesis of a lipid-linked core oligosaccharide and its transfer to proteins in the ER. Defects in this pathway lead to underglycosylation of serum and other glycoproteins. On the other hand, CDG-II defects affect subsequent processing steps of the protein-bound *N*-glycan structure in the Golgi.[46]

MS analysis of the glycosylation state of transferrin, a plasma protein with two biantennary *N*-glycans, can already be used routinely to screen for CDG-I.[47−50] The intact transferrin is analyzed using SELDI-TOF with anti-transferrin coated protein chips, or after partial purification of transferrin from plasma or serum,[47] LC-MS[48,51,52] or MALDI-TOF MS[49,50] is used to assess the glycosylation state. Mass differences account for the loss of glycan-moieties, but structural analysis of the glycans is not possible in this way. Besides analysis of intact glycoproteins, MS analysis of glycopeptides can be performed. Especially the use of a combination of electron capture dissociation (ECD), collision induced dissociation (CID) and infrared multiphoton dissociation (IRMPD) as fragmentation techniques after LC-FT−MS, provides information on both the site of glycosylation and the structure of the glycan-moiety.[3,53−56] The glycan structure can also be analyzed separately after enzymatic cleavage of the glycan.[57,58] These relatively new techniques offer many opportunities to study CDGs. Not only the aberrant glycosylation in CDG-I, but also the abnormal glycan structure in CDG-II and in *O*-glycosylation defects can be investigated.

Quantitative Proteomics Approaches. More recently, quantitative proteomics approaches have been developed, in which samples are differentially labeled and protein expression levels can be compared quantitatively. Proteins can be labeled at different stages during sample preparation. Labeling can be performed *in vivo* by stable isotope labeling by amino acids in

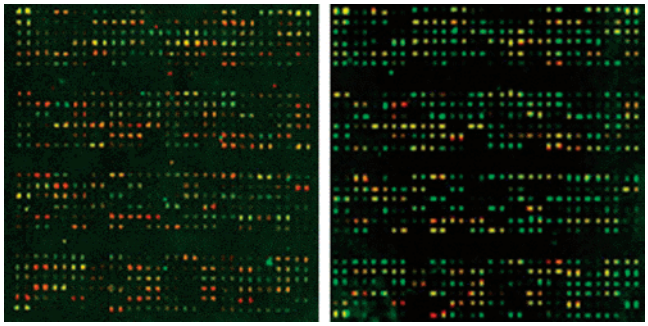

Figure 3. Antibody microarray analysis comparing pooled sera samples of control subjects and CF patients, reacted with Cy3 and Cy5. The inversely labeled samples were applied to the other array. Proteins that are elevated in one source compared to the other either fluoresce green or red, whereas proteins with similar levels in both samples appear in yellow. Reprinted from ref 69.

cell culture (SILAC)[59] or *in vitro* by several different methods, such as DIGE, isotope-coded affinity tag (ICAT),[60] isobaric tag for relative and absolute quantification (ITRAQ),[61] or N-terminal labeling of peptides.[62] These methods are all very well suited to study differences between health and disease. Therefore, they are very well applicable to study inborn errors of metabolism and probably they will become more prominent in future proteomics research on genetic and metabolic disorders.

Functional Proteomics. Besides studies on protein expression, proteomics approaches can also be used to study the functionality of proteins. Functional proteomics approaches are based on interactions of proteins or specific activities of proteins. Protein microarrays have been developed in analogy to DNA microarrays to compare expression of large sets of proteins between diseased state and control state, but also to study the function of proteins. Furthermore, when using specifically designed surfaces for SELDI-TOF, protein activities can be measured. Interactions between a specific bait protein and its unknown partners can be investigated using MS techniques in combination with immunoaffinity methods[4,63] or surface-plasmon resonance.[64−66]

Protein Microarrays. Antibody-based microarrays are a novel technology in the proteomics field that holds great promise for use in clinical applications. Microarrays can be printed with thousands of recombinant antibodies carrying the desired specificities, which makes it very well suited for high throughput screening of biological samples for specific disease markers.[67,68] Srivastava et al. used this approach to screen for specific expression patterns in serum of cystic fibrosis (CF) patients compared to control subjects.[69] Serum proteins of patients and controls were labeled with Cy3 or Cy5 respectively, and equal volumes of labeled sera were applied to the antibody microarray, which is shown in Figure 3. This resulted in a list of differentially expressed proteins in serum of CF patients compared to control subjects. Although larger numbers of samples have to be measured, their results show that one can discriminate between serum of CF patients and control subjects using antibody microarrays.

In contrast to using chips with immobilized antibodies to detect specific proteins, protein chips carrying the proteome of a specific organism or cell type can be made by cloning and purification of these proteins.[70] This protein microarray can then be screened on the ability to bind specific ligands or interact with specific proteins. Although this approach has not been used with respect to inborn errors of metabolism, it could be useful in identifying affected proteins, protein complexes, or pathways in disorders of unknown etiology.

Protein Activity Measurements using SELDI-TOF. SELDI-TOF MS can also be used to design specific tests for the diagnosis of genetic diseases, such as thrombotic thrombocytopenic purpura (TTP).[71] TTP is a thrombotic microangiopathy caused impairment of ADAMTS-13 by either a hereditary or an acquired deficiency. An acquired deficiency of ADAMTS-13 is caused by autoantibodies that specifically inhibit ADAMTS-13 function. ADAMTS-13 is a metalloprotease enzyme that cleaves von Willebrand Factor (VWF) multimers into smaller protein units.[72] A diagnostic assay for ADAMTS-13 activity in plasma was developed. In this assay, a recombinant VWF substrate containing a His-tag was cleaved by ADAMTS-13. After incubation, the cleaved peptide was bound to an IMAC/Nickel protein chip and quantified using SELDI-TOF MS. Thus, a rapid assay to diagnose TTP has been developed using SELDI-TOF MS.

Affinity Pull Down Strategies to Study Protein Interactions. Besides the high throughput protein microarray technology, more conventional methods to study interactions of a protein can be used in concert with MS based techniques.[4,63] In general, a bait protein or ligand is immobilized on agarose beads. After incubation of these beads with the biological sample of interest (plasma or cell lysate) and extensive washing to remove nonspecifically bound proteins, the complexes are eluted from the beads and analyzed using MS methods. A disadvantage of this technique is that the protein complex is formed *in vitro*. To circumvent this problem, the tagged bait protein can be expressed in an appropriate cell line and protein complexes are allowed to form *in vivo*. Cell extracts are then immunoprecipitated using antibodies directed against the tag of the protein of interest. The latter approach has been used to study the components of the Fanconi Anemia (FA) core complex.[73] FA is an autosomal recessive disorder affecting all bone marrow elements and is associated with cardiac, renal, and limb malformations as well as dermal pigmentary changes. The molecular mechanism of the disease remains elusive, but at least 11 complementation groups have been defined and 8 genes have been cloned that are involved in FA. At least 6 FA proteins are known to be part of the FA core complex. The complex has not yet been fully characterized and its physiological function remains elusive. To study the complex and possibly identify novel complex components, one of the FA core complex proteins, FANCA, was expressed with a Flag-tag in HeLa cells and FANCA deficient cells. Immunoprecipitation was performed with anti-Flag-affinity resin. MS analysis of the eluted complex identified all known binding partners of the FA core complex. Using this approach in combination with cell fractionation, it was shown that the FA core complex forms four complexes of different sizes in different subcellular compartments. By using this approach, further research can be done to characterize the FA core complex in more detail.

Concluding remarks

With the increase in different proteomics techniques over the past few years, a large variety of biological questions can be answered using these approaches. Especially in cancer research, proteomics methods are widely used to study protein expression and function, but these techniques are also very well suited to investigate genetic and metabolic disorders, as can be seen from the examples above. In the near future, the development of reliable and fast quantification methods for

MS based proteomics will aid in the growing applications of proteomics techniques in research and diagnosis of genetic and metabolic disorders.

Because proteomics approaches specifically aim at changes in proteins and protein levels, the techniques described are complementary to metabolomics approaches, as NMR spectroscopy or LC-MS methods, which are targeted on changes in metabolite profiles or levels. Both proteomics and metabolomics strategies may be relevant for the diagnosis and the unraveling of the pathophysiology of specific genetic and metabolic diseases.

References

(1) Lanpher, B.; Brunetti-Pierri, N.; Lee, B. *Nat. Rev. Genet.* **2006**, *7*, 449−459.
(2) Dipple, K. M.; McCabe, E. R. *Am. J. Hum. Genet.* **2000**, *66*, 1729−1735.
(3) Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom. Rev.* **2005**, *24*, 201−222.
(4) Monti, M.; Orru, S.; Pagnozzi, D.; Pucci, P. *Clin. Chim. Acta* **2005**, *357*, 140−150.
(5) Lilley, K. S.; Razzaq, A.; Dupree, P. *Curr. Opin. Chem. Biol.* **2002**, *6*, 46−50.
(6) Hutchens, T. W.; Yip, T. T. *Rapid Commun. Mass Spectrom.* **1993**, 576−580.
(7) Issaq, H. J.; Veenstra, T. D.; Conrads, T. P.; Felschow, D. *Biochem. Biophys. Res. Commun.* **2002**, *292*, 587−592.
(8) Caldwell, R. L.; Caprioli, R. M. *Mol. Cell. Proteomics* **2005**, *4*, 394−401.
(9) Kolch, W.; Neususs, C.; Pelzing, M.; Mischak, H. *Mass Spectrom. Rev.* **2005**, *24*, 959−977.
(10) Bischoff, R.; Luider, T. M. *J. Chromatogr., B: Biomed. Sci. Appl.* **2004**, *803*, 27−40.
(11) O'Farrell, P. H. *J. Biol. Chem.* **1975**, *250*, 4007−4021.
(12) Klose, J. *Humangenetik* **1975**, *26*, 231−243.
(13) Unlu, M.; Morgan, M. E.; Minden, J. S. *Electrophoresis* **1997**, *18*, 2071−2077.
(14) Tonge, R.; Shaw, J.; Middleton, B.; Rowlinson, R.; Rayner, S.; Young, J.; Pognan, F.; Hawkins, E.; Currie, I.; Davison, M. *Proteomics* **2001**, *1*, 377−396.
(15) Naureckiene, S.; Sleat, D. E.; Lackland, H.; Fensom, A.; Vanier, M. T.; Wattiaux, R.; Jadot, M.; Lobel, P. *Science* **2000**, *290*, 2298−2301.
(16) Vanier, M. T.; Millat, G. *Clin. Genet.* **2003**, *64*, 269−281.
(17) Patterson, M. C.; Vanier, M. T.; Suzuki, K.; Morris, J. A.; Carstea, E.; Neufeld, E. B.; Blanchette-Mackie, J. E.; Pentchev, P. G. In *The Metabolic & Molecular Bases of Inherited Disease*, 8th ed.; Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., Eds.; McGraw-Hill: New York, 2001; pp 3611−3634.
(18) Valenzano, K. J.; Kallay, L. M.; Lobel, P. *Anal. Biochem.* **1993**, *209*, 156−162.
(19) Richard, E.; Monteoliva, L.; Juarez, S.; Perez, B.; Desviat, L. R.; Ugarte, M.; Albar, J. P. *J. Proteome. Res.* **2006**, *5*, 1602−1610.
(20) Deodato, F.; Boenzi, S.; Santorelli, F. M.; onisi-Vici, C. *Am. J. Med. Genet C. Semin. Med. Genet.* **2006**, *142*, 104−112.
(21) Fenton, W. A.; Gravel, R. A.; Rosenblatt, D. S. In *The Metabolic & Molecular Bases of Inherited Disease*, 8th ed.; Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., Eds.; McGraw-Hill: New York, 2001; pp 2165−2194.
(22) Paweletz, C. P.; Liotta, L. A.; Petricoin, E. F., III *Urology* **2001**, *57*, 160−163.
(23) Srinivas, P. R.; Srivastava, S.; Hanash, S.; Wright, G. L., Jr. *Clin. Chem.* **2001**, *47*, 1901−1911.
(24) von Eggeling, F.; Junker, K.; Fiedle, W.; Wollscheid, V.; Durst, M.; Claussen, U.; Ernst, G. *Electrophoresis* **2001**, *22*, 2898−2902.
(25) Austen, B. M.; Frears, E. R.; Davies, H. *J. Pept. Sci.* **2000**, *6*, 459−469.
(26) Ranganathan, S.; Williams, E.; Ganchev, P.; Gopalakrishnan, V.; Lacomis, D.; Urbinelli, L.; Newhall, K.; Cudkowicz, M. E.; Brown, R. H., Jr.; Bowser, R. *J. Neurochem.* **2005**, *95*, 1461−1471.
(27) Ishihama, Y. *J. Chromatogr. A* **2005**, *1067*, 73−83.
(28) Ramstrom, M.; Ivonin, I.; Johansson, A.; Askmark, H.; Markides, K. E.; Zubarev, R.; Hakansson, P.; Aquilonius, S. M.; Bergquist, J. *Proteomics* **2004**, *4*, 4010−4018.
(29) Schulz-Knappe, P.; Zucht, H. D.; Heine, G.; Jurgens, M.; Hess, R.; Schrader, M. *Comb. Chem. High Throughput. Screen.* **2001**, *4*, 207−217.
(30) Schulte, I.; Tammen, H.; Selle, H.; Schulz-Knappe, P. *Expert. Rev. Mol. Diagn.* **2005**, *5*, 145−157.
(31) Tammen, H.; Mohring, T.; Kellmann, M.; Pich, A.; Kreipe, H. H.; Hess, R. *Clin. Chem.* **2004**, *50*, 545−551.
(32) Kohler, H. P.; Schroder, V. *Hamostaseologie* **2002**, *22*, 53−58.
(33) Balogh, I.; Szoke, G.; Karpati, L.; Wartiovaara, U.; Katona, E.; Komaromi, I.; Haramura, G.; Pfliegler, G.; Mikkola, H.; Muszbek, L. *Blood* **2000**, *96*, 2479−2486.
(34) Boot, R. G.; Verhoek, M.; de Fost, M.; Hollak, C. E. M.; Maas, M.; Bleijlevens, B.; van Breemen, M. J.; van Meurs, M.; Boven, L. A.; Laman, J. D.; Moran, M. T.; Cox, T. M.; Aerts, J. M. F. G. *Blood* **2004**, *103*, 33−39.
(35) Beutler, E.; Grabowski, G. A. In *The Metabolic & Molecular Bases of Inherited Disease*, 8th ed.; Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., Eds.; McGraw-Hill: New York, 2001; pp 3635−3668.
(36) Beutler, E.; Kay, A.; Saven, A.; Garver, P.; Thurston, D.; Dawson, A.; Rosenbloom, B. *Blood* **1991**, *78*, 1183−1189.
(37) Hollak, C. E.; van Weely, S.; van Oers, M. H.; Aerts, J. M. *J. Clin. Invest.* **1994**, *93*, 1288−1292.
(38) Wada, Y.; Matsuo, T.; Katakuse, I.; Suzuki, T.; Azuma, T.; Tsujino, S. i.; Kishimoto, S.; Matsuda, H.; Hayashi, A. *Biochim. Biophys. Acta* **1986**, *873*, 316−319.
(39) Ranlov, I.; Ando, Y.; Ohlsson, P. I.; Holmgren, G.; Ranlov, P. J.; Suhr, O. B. *Eur. J. Clin. Invest.* **1997**, *27*, 956−959.
(40) Bergen, H. R., III; Zeldenrust, S. R.; Butz, M. L.; Snow, D. S.; Dyck, P. J.; Dyck, P. J.; Klein, C. J.; O'Brien, J. F.; Thibodeau, S. N.; Muddiman, D. C. *Clin. Chem.* **2004**, *50*, 1544−1552.
(41) Nepomuceno, A. I.; Mason, C. J.; Muddiman, D. C.; Bergen, H. R., III; Zeldenrust, S. R. *Clin. Chem.* **2004**, *50*, 1535−1543.
(42) Benson, M. In *The Metabolic & Molecular Bases of Inherited Disease*, 8th ed.; Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., Eds.; McGraw-Hill: New York, 2001; pp 5345−5378.
(43) Altland, K.; Winter, P.; Sauerborn, M. K. *Electrophoresis* **1999**, *20*, 1349−1364.
(44) Connors, L. H.; Ericsson, T.; Skare, J.; Jones, L. A.; Lewis, W. D.; Skinner, M. *Biochim. Biophys. Acta.* **1998**, *1407*, 185−192.
(45) Wada, Y. *J. Chromatogr., B. Analyt. Technol. Biomed. Life Sci.* **2006**, *838*, 3−8.
(46) Jaeken, J.; Matthijs, G.; Carchon, H.; van Schaftingen, E. In *The Metabolic & Molecular Bases of Inherited Disease*, 8th ed.; Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., Eds.; McGraw-Hill: New York, 2001; pp 1601−1622.
(47) Mills, K.; Mills, P.; Jackson, M.; Worthington, V.; Beesley, C.; Mann, A.; Clayton, P.; Grunewald, S.; Keir, G.; Young, L.; Langridge, J.; Mian, N.; Winchester, B. *Proteomics* **2006**, *6*, 2295−2304.
(48) Bergen, H. R.; Lacey, J. M.; O'Brien, J. F.; Naylor, S. *Anal. Biochem.* **2001**, *296*, 122−129.
(49) Nakanishi, T.; Okamoto, N.; Tanaka, K.; Shimizu, A. *Biol. Mass Spectrom.* **1994**, *23*, 230−233.
(50) Wada, Y.; Gu, J.; Okamoto, N.; Inui, K. *Biol. Mass Spectrom.* **1994**, *23*, 108−109.
(51) Kleinert, P.; Kuster, T.; Durka, S.; Ballhausen, D.; Bosshard, N. U.; Steinmann, B.; Hanseler, E.; Jaeken, J.; Heizmann, C. W.; Troxler, H. *Clin. Chem. Lab Med.* **2003**, *41*, 1580−1588.
(52) Lacey, J. M.; Bergen, H. R.; Magera, M. J.; Naylor, S.; O'Brien, J. F. *Clin. Chem.* **2001**, *47*, 513−518.
(53) Bakhtiar, R.; Guan, Z. *Biochem. Biophys. Res. Commun.* **2005**, *334*, 1−8.
(54) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530−4536.
(55) Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256−3262.
(56) Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431−4436.
(57) Mills, P. B.; Mills, K.; Mian, N.; Winchester, B. G.; Clayton, P. T. *J. Inherited Metab. Dis.* **2003**, *26*, 119−134.
(58) Sagi, D.; Kienz, P.; Denecke, J.; Marquardt, T.; Peter-Katalinic, J. *Proteomics* **2005**, *5*, 2689−2701.
(59) Ong, S. E.; Mann, M. *Nat. Chem. Biol.* **2005**, *1*, 252−262.
(60) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994−999.
(61) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. *Mol. Cell. Proteomics* **2004**, *3*, 1154−1169.

(62) Hoving, S.; Munchbach, M.; Schmid, H.; Signor, L.; Lehmann, A.; Staudenmann, W.; Quadroni, M.; James, P. *Anal. Chem.* **2000**, *72*, 1006−1014.
(63) Monti, M.; Orru, S.; Pagnozzi, D.; Pucci, P. *Biosci. Rep.* **2005**, *25*, 45−56.
(64) Buijs, J.; Franklin, G. C. *Brief Funct. Genomic Proteomic* **2005**, *4*, 39−47.
(65) Nedelkov, D.; Nelson, R. W. *Trends Biotechnol.* **2003**, *21*, 301−305.
(66) Larsericsdotter, H.; Jansson, O.; Zhukov, A.; Areskoug, D.; Oscarsson, S.; Buijs, J. *Proteomics* **2006**, *6*, 2355−2364.
(67) Wingren, C.; Borrebaeck, C. A. *Expert. Rev. Proteomics* **2004**, *1*, 355−364.
(68) LaBaer, J.; Ramachandran, N. *Curr. Opin. Chem. Biol.* **2005**, *9*, 14−19.
(69) Srivastava, M.; Eidelman, O.; Jozwik, C.; Paweletz, C.; Huang, W.; Zeitlin, P. L.; Pollard, H. B. *Mol. Genet. Metab.* **2006**, *87*, 303−310.
(70) Zhu, H.; Bilgin, M.; Bangham, R.; Hall, D.; Casamayor, A.; Bertone, P.; Lan, N.; Jansen, R.; Bidlingmaier, S.; Houfek, T.; Mitchell, T.; Miller, P.; Dean, R. A.; Gerstein, M.; Snyder, M. *Science* **2001**, *293*, 2101−2105.
(71) Jin, M.; Cataland, S.; Bissell, M.; Wu, H. M. *J. Thromb. Haemost.* **2006**, *4*, 333−338.
(72) Levy, G. G.; Nichols, W. C.; Lian, E. C.; Foroud, T.; McClintick, J. N.; McGee, B. M.; Yang, A. Y.; Siemieniak, D. R.; Stark, K. R.; Gruppo, R.; Sarode, R.; Shurin, S. B.; Chandrasekaran, V.; Stabler, S. P.; Sabio, H.; Bouhassira, E. E.; Upshaw, J. D., Jr.; Ginsburg, D.; Tsai, H. M. *Nature* **2001**, *413*, 488−494.
(73) Thomashevski, A.; High, A. A.; Drozd, M.; Shabanowitz, J.; Hunt, D. F.; Grant, P. A.; Kupfer, G. M. *J. Biol. Chem.* **2004**, *279*, 26201−26209.

PR060487W