

Thongboonkerd V, Klein JB (eds): Proteomics in Nephrology.
Contrib Nephrol. Basel, Karger, 2004, vol 141, pp 1–10

.....

Overview of Proteomics

Jon B. Klein, Visith Thongboonkerd

Core Proteomics Laboratory, Kidney Disease Program, Department of Medicine,
University of Louisville, Louisville, Ky., USA

What Is Proteomics?

In 1975, two-dimensional polyacrylamide gel electrophoresis (2-DE) was simultaneously described by O'Farrell [1] and by Klose [2]. Their techniques allowed, for the first time, the separation of complex mixtures of proteins into individual components. Using this technique, initially 1,100 protein components were resolved from *Escherichia coli* on a 2-D gel [1]. Because techniques for protein identification were limited at that time and due to a lack of standardized reagents, 2-DE was not extensively used. However, recent advances in mass spectrometry (MS), in parallel with successes in genomic analysis, have made it possible to identify proteins separated by 2-DE. The high-throughput capability of MS, coupled with 2-DE separations of proteins, has made proteomic analysis practical. This field of research has acquired the name '*Proteomics*'. 'PROTEins expressed by the genOME' have been termed '*Proteome*', a term firstly used in the public by Marc Wilkins at the first Siena proteomic conference in 1994 [3]. Proteomics is a field that Anderson and Anderson [4] have defined as: 'the use of quantitative protein-level measurement of gene expression to characterize biological processes (e.g., disease processes and drug effects) and decipher the mechanisms of gene expression control'. It is a field of study that is separated from, but complementary to genomics. The field has expanded rapidly as the number of published proteomic articles has been increasing rapidly. The first publication related to proteomics appeared in 1995 and there are approximately 3,400 publications related to proteomics through the first quarter of 2003 (fig. 1).

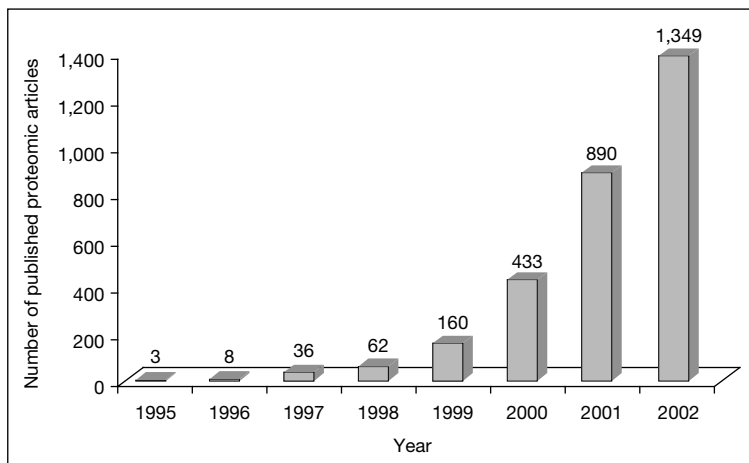


Fig. 1. Number of published proteomic articles. The PubMed search was performed using the keyword ‘proteomics’, ‘proteome’, or ‘proteomic’. The article with more than one of these keywords found was counted as only one. Rapid growing of the publications related to proteomics is observed by upward parabolic curve of the increasing in the number.

Why Proteomics?

Advances in molecular biology techniques during the past two decades have led to better understanding the genetic and molecular bases of a wide variety of human diseases. In nephrology, genetic analysis has explained inherited renal disorders with primary renal involvement, including primary glomerular diseases, cystic renal diseases, renal tubular diseases, and inherited renal neoplasms [5]. Recently, the term *functional genomics* has been used to describe analysis of changes in gene expression in response to various experimental conditions. These types of analyses have played an important role in studying renal physiology and pathophysiology [6, 7]. However, there are some limitations of genomic analysis that it cannot provide complete information of cellular, subcellular, and supracellular functions, in which *proteins*, not *genes*, govern the functions. Simply put, differential developmental stages of the animals are resulted from changes in global protein expression or the proteome, while gene expression remains static (fig. 2). This can be explained by dynamic processes, especially posttranslational modifications (PTMs) that modify gene products and cause functional changes in the proteins. Indeed, studies of protein expression, rather than gene expression, have generally been used for several decades for biological study. One of the most successful techniques in protein

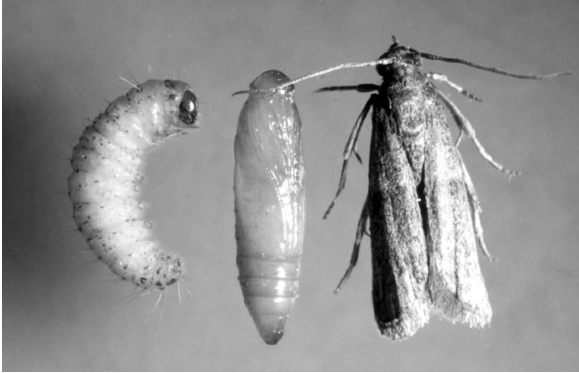


Fig. 2. Why proteomics? The figure shows different stages of life cycle in an insect. With the same genes, they have different morphologies in various stages of life cycle. The morphological differences accompany differential cellular and organelle functions are directly determined by the variety of their proteome, not by the genome that can be modified by various PTMs.

studies is Western blotting, which determines protein expression by using a specific antibody to bind to the protein of interest. This technique, however, is limited because specific antibodies must be available and only a few proteins can be tested in a single experiment. Moreover, Western blotting techniques require a prior assumption from previous studies about the protein to be identified.

How to Perform Proteomic Analysis?

Proteomic analysis can be simply classified as three main categories: expression proteomics, bioinformatic analysis, and functional proteomics. While the most important part of proteomic analysis is the functional study, expression proteomics may be a necessary initial step. Unless a hypothesis has been generated by prior study, expression proteomics can be used to screen biological samples to identify candidate proteins for further functional studies. To date, most proteomic studies have compared protein expression in normal and disease states in cells and tissues. Expression proteomics alone, however, does not provide any functional or physiological significance. Functional proteomics and other functional studies will take role after candidates or hypotheses have been generated from initial expression studies. A summary of this typical proteomic approach is shown in figure 3.

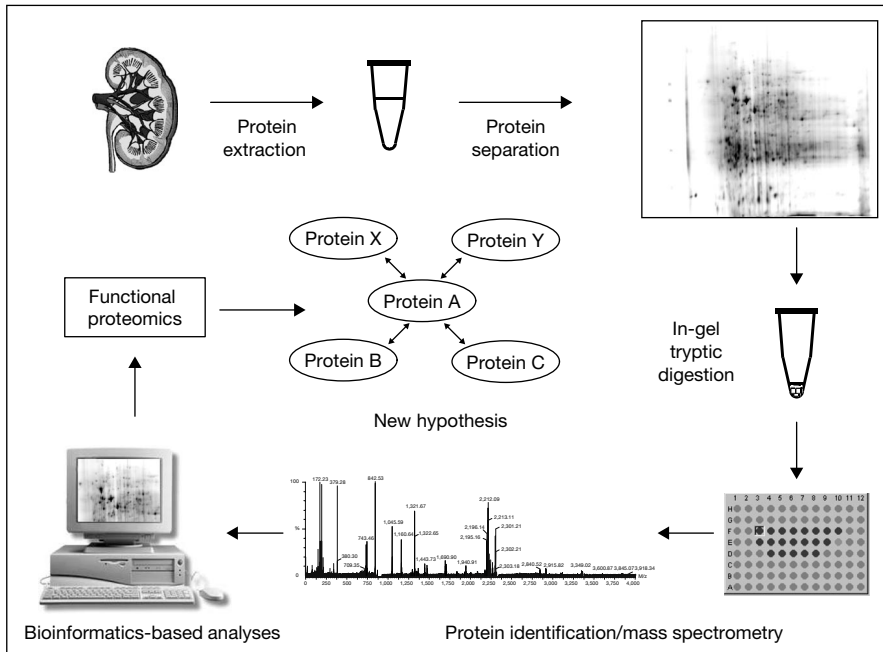


Fig. 3. The typical methodology in proteomic analysis. Typical processes in expression proteomics are: protein extraction and separation, protein identification, and bioinformatics-based analyses. Protein extraction is to isolate the proteins from tissues. Protein separation can be performed either by 2-DE or by LC techniques. The proteins are then identified by various MS methods. Bioinformatics-based analyses are used to obtain protein information to guide further functional proteomic studies. The final step in proteomic study is to generate a new hypothesis from both expression and functional proteomic data. The novel hypothesis is then examined by various methods.

Expression Proteomics

The initial process of expression proteomics is to prepare the samples to be ready for proteomic analysis, either by gel-based or gel-free methods. Because of differences in protein physical and chemical properties, different extraction protocols may favor proteins from differential compartments of the entire proteome. For example, most of the extraction protocols for 2-DE separation have a limitation in solubilizing membrane proteins. After extraction or isolation of proteins from cells and tissues, the proteins can be separated either by 2-DE or by liquid chromatography (LC) that requires specific conditionings for each procedure. Thereafter, those proteins are identified, mostly by MS-based methods.

Expression proteomics can be performed as to identify proteins expressed in a specific cell or tissue. However, comparing protein expression between two

(or more) differential experimental conditions or disease states will provide more information to facilitate understanding physiology and pathogenic mechanisms. *Quantitative proteomics* takes role to determine differential protein abundance in various conditions. Quantitation can be performed by indirect method (intensity analysis) or by direct measurement using isotope labeling [details in the chapter ‘Mass spectrometric approaches to quantitative proteomics’ by Sechi et al.]. While the intensity analysis is simple and widely used, it is limited by sensitivity and accuracy, which allow new technologies of MS-based quantitative techniques to be developed.

Bioinformatic Analysis

After the proteins are identified, bioinformatic analysis will take role in expanding the initial protein information. Additionally, bioinformatic analysis can be used to guide the future direction of functional proteomic studies. The following are additional information that can be obtained by bioinformatic analysis: (1) primary, secondary, and tertiary structures; (2) sequence alignment and homology; (3) motifs and domains; (4) protein interactions and networks; (4) potential PTMs; (5) potential transmembrane regions; (6) subcellular locations, and (7) miscellaneous. Several bioinformatic analytical tools are freely accessible at <http://us.expasy.org/tools/>. Additional information obtained from bioinformatic analysis may facilitate and make functional analysis be more focused.

Functional Proteomics

To understand the role of targeted proteins in cellular functions, high-throughput comprehensive analyses of protein-protein interactions, protein complexes, and PTMs are necessary. Protein interactions and complexes can be analyzed by co-immunoprecipitation techniques followed by 2-DE or by affinity chromatography, protein chip [8, 9], tandem (sequential) affinity tag purification (TAP) [10], biomolecular interaction analysis-mass spectrometry (BIA-MS) [11], etc. All of the applications are based on ‘ligand fishing’ by using the specific antibody or tag-specific receptor as a ‘bait’ and the ‘prey’ or interacting proteins from ligand (analyte)-containing elution can be further analyzed by MS techniques.

In addition to the aforementioned screening steps for determination of PTMs, a confirmation method may be needed in some circumstances. The methods employed depend on the chemistry involved in the PTMs. For example, phosphoproteins can be detected by MS-based methods or by Western blot analysis using a phosphoprotein-specific antibody that can be performed in parallel and compared with stained 2-D gel [12]. An alternative approach to phosphoproteomics is using enzymatic radiolabeling, using the protein of interest as

a substrate, and using ^{32}P phosphate (^{32}ATP) as a reagent. The protein sample is treated with the appropriate substrate containing ^{32}P , which is incorporated into the protein. Subsequent separation with 2-DE allows identification and quantitation by autoradiography [13]. Fluorescence is another means to detect protein-bound phosphate group by staining with phosphate-specific fluorescence reagents. BO-INI, a fluorescent label, provides a rapid screening method for phosphoproteins without the use of more dangerous radioactive materials [14]. Protein kinases and phosphatases can be compared as the ‘yin and yang’ of protein phosphorylation and cell signaling [15]. Significant functional roles of phosphorylation can be determined by exposing the cells to various conditionings of protein kinases, phosphatases, and phosphatase inhibitors.

Applications of Proteomics to Nephrology

The earliest reference to the application of *renal proteomics* may be in 1995 when Witzmann et al. [16] used 2-DE to identify renal stress proteins as the biomarkers for chemical toxicity. Proteomic analysis of the urine, *urinary proteomics*, was begun in 1996 when Marshall and Williams [17] identified urinary proteins on 2-D gel using a dye precipitation method. There were only a few proteomic studies in nephrology during the late 1990s by a variety of reasons including the capital and labor-intensive nature of proteomics. More rapid progress began with the new era in MS analyses, when Patterson et al. [18] identified components of the urinary proteome using LC coupled with tandem MS (LC-MS/MS). More recently, we successfully isolated and identified a number of normal human urinary proteins using two different isolation methods [19]. Proteomics has now become a very fruitful field in nephrological research and several renal centers have initiated proteomic analysis applied to their research areas. Applications of proteomics in nephrology can be divided into three categories: normal physiology, pathophysiology of diseases, and biomarker discovery, which are detailed in individual chapters in this book.

Limitations of Proteomics and Potential Solutions

Although proteomic analysis is a powerful investigative tool, it still suffers from significant technical limitations. Until now, methods available for proteomics analyses can be counted as only the first step. Various advances in proteomic technologies are now being developed to solve these problems. Below, we have summarized some of the major limitations, as applied to renal and urinary proteomics, and also the possible solutions in each problem.

Extraction of Integral Membrane Proteins. Because receptors and channel proteins are the typical membrane proteins, extraction of integral membrane proteins from the kidney is very important to study the renal physiology. The difficulty in extraction of these proteins is a consequence of their low solubility and the need to produce dispersion of hydrophobic proteins for 2-D analysis. Sequential extraction methods described by Molloy et al. [20] are a possible solution. For the first step, Tris-base is used to solubilize most of the cytosolic proteins. The conventional buffer containing urea, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate, DTT, Tris, and carrier ampholytes are used for solubilizing the remaining pellet. At this point, 85% of initial protein masses can be solubilized. Finally, the remaining pellet is solubilized with a combination of urea, thiourea, tributyl phosphine and multiple zwitterion surfactants. Although this technique is able to extract almost all of protein components from the cell, it is laborious and takes a long time to complete all extraction steps, with the consequent loss of materials. We have demonstrated a simpler way to extract the integral membrane proteins, using sodium dodecyl sulfate (SDS) as a primary reagent as described in SDS/Cymal method [details in the chapter ‘Sample preparation for 2-D proteomic analysis’ by Thongboonkerd et al.]. Although this technique is satisfactory for extracting membrane proteins from the kidney, the hydrophilic or cytosolic proteins are partly obscured. To achieve the broadest representation of all the proteins in the tissues, an applied protocol should be a compromise between complete extraction of both membrane and cytosolic proteins and the ease for application. Alternative approach is using LC-based techniques, which enhance the identification of membrane and low-abundance proteins.

Separation and Identification of Low-Abundance Proteins. Expression of low-abundance proteins separated by 2-DE is normally obscured by higher abundance proteins. A large number of obscured protein spots can be visualized by using overlapped narrow-range (zoom-in) immobilized pH gradient (IPG) strips in the first dimensional separation. Using zoom-in gel, the overall number of protein spots, excluding overlapped spots between each zoom-in gel, is 2.3- to 3.5-fold greater than using the ordinary wide-range (pH 3–10) IPG strip [21, 22]. Longer length IPG strips (using 24 cm instead of ordinary 18 cm long) can be applied for better separation, especially those that are obscured by the higher abundance proteins. This approach also facilitates the separation of multiform proteins that are present with very little changes in their isoelectric points (pI) or molecular weights [22]. Alternatively, protein samples can be prefractionated before isoelectric focusing (IEF) to overcome the diversity and dynamic range of the expressed proteins [23–26]. All these techniques may be suitable for separation of low-abundance proteins. To identify the low-abundance protein, LC-MS/MS, electrospray ionization-tandem MS (ESI-MS/MS), and

multidimensional protein identification technology (MudPIT) methods can be used to identify proteins as low as femtomole levels (down to 10 fmol) [27–29]. Additionally, immunological methods can be combined to magnify the results.

Highly Basic Proteins. A broad-range (pH 3–10) IPG strip is generally used for IEF separation. However, our experience is that, mostly, only proteins in the pI ranges from 3.5 to 8.5 are present in the 2-D gel. In this case, expression of highly basic proteins (pI > 8.5) cannot be visualized by using conventional broad range IPG strips. Recently, basic range IPG strips (pH 3–12, 6–12, 8–12, 9–12, and 10–12) have been developed and are commercially available [30, 31]. Additionally, nonequilibrium pH gradient electrophoresis (NEPHGE) is another successful method to visualize the highly basic proteins [32]. During the IEF separation, while the pH gradient is forming, positive charges of basic proteins move toward the negative end of the gel. Since IEF running to the equilibrium point would result in many highly basic proteins exiting from the basic end of the gel (the cathode end), the electrophoresis can be stopped at a critical point before the equilibrium occurs (thus, ‘nonequilibrium’) [33].

Proteomic Analysis of Renal Biopsy Tissue. The availability of renal tissue for proteomic analyses can become a significant obstacle. For samples obtained from cadaveric tissue, this is not a problem. However, for samples from biopsies, the available tissue is severely limited. Especially in the case of rare diseases, it may not be possible to obtain adequate replicated samples for proteomic study. A temporary solution may be the storage of rare renal tissues at ultra-low temperatures (less than -20°C) until such time as very sensitive identification techniques become available. A recent possible approach is to use high-sensitive identification techniques such as surface-enhanced laser desorption/ionization time-of-flight MS (SELDI-TOF MS), MudPIT, and other ‘shotgun’ techniques [27, 34–36]. However, shotgun technologies may provide only limited information of protein characteristics. Additional tissue is still needed for complete analyses.

Future Directions

To date, there are fewer than 50 publications using proteomic approaches in renal medicine. In the coming years, the authors expect that the number of publications in renal and urinary proteomics will increase rapidly, as has been the case in the study of other organ systems. Future directions are focused not only on renal physiology, pathophysiology, and biomarker discovery, but also on development of new target therapies. Urinary proteomics will be applied more widely because of its non-invasiveness to study the human subjects.

This has been facilitated by recent protein isolation protocols that allow urinary protein samples to be transferred to centralized core proteomics facilities [19]. Renal and urinary proteomic analysis has been applied to the study of normal physiologic processes. The development of normal proteomic profiles of renal tissue and urine can be applied to a wide spectrum of renal diseases. The clinical application of proteomic analysis to primary glomerulopathies, diabetic renal diseases, and tubular disorders holds a great promise.

References

- 1 O'Farrell PH: High-resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 1975; 250:4007–4021.
- 2 Klose J: Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 1975;26: 231–243.
- 3 Wilkins MR, Sanchez JC, Gooley AA, Appel RD, Humphery-Smith I, Hochstrasser DF, Williams KL: Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* 1996;13:19–50.
- 4 Anderson NL, Anderson NG: Proteome and proteomics: New technologies, new concepts, and new words. *Electrophoresis* 1998;19:1853–1861.
- 5 George AL Jr, Neilson EG: Genetics of kidney disease. *Am J Kidney Dis* 2000;35:S160–S169.
- 6 Hieter P, Boguski M: Functional genomics: It's all how you read it. *Science* 1997;278:601–602.
- 7 Takenaka M, Imai E: Functional genomics in nephrology. *Nephrol Dial Transplant* 2000;15:139–141.
- 8 Figeys D, Pinto D: Proteomics on a chip: Promising developments. *Electrophoresis* 2001;22:208–216.
- 9 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 RA, Gerstein M, Snyder M: Global analysis of protein activities using proteome chips. *Science* 2001;293:2101–2105.
- 10 Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B: A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* 1999;17: 1030–1032.
- 11 Nelson RW, Nedelkov D, Tubbs KA: Biosensor chip mass spectrometry: A chip-based proteomics approach. *Electrophoresis* 2000;21:1155–1163.
- 12 Yanagida M, Miura Y, Yagasaki K, Taoka M, Isobe T, Takahashi N: Matrix assisted laser desorption/ionization-time of flight-mass spectrometry analysis of proteins detected by anti-phosphotyrosine antibody on two-dimensional-gels of fibroblast cell lysates after tumor necrosis factor- α stimulation. *Electrophoresis* 2000;21:1890–1898.
- 13 Yan JX, Packer NH, Gooley AA, Williams KL: Protein phosphorylation: Technologies for the identification of phosphoamino acids. *J Chromatogr A* 1998;808:23–41.
- 14 Wang P, Giese RW: Phosphate-specific fluorescence labeling with BO-IMI: Reaction details. *J Chromatogr A* 1998;809:211–218.
- 15 Hunter T: Protein kinases and phosphatases: The yin and yang of protein phosphorylation and signaling. *Cell* 1995;80:225–236.
- 16 Witzmann F, Clack J, Fultz C, Jarnot B: Two-dimensional electrophoretic mapping of hepatic and renal stress proteins. *Electrophoresis* 1995;16:451–459.
- 17 Marshall T, Williams K: Two-dimensional electrophoresis of human urinary proteins following concentration by dye precipitation. *Electrophoresis* 1996;17:1265–1272.
- 18 Spahr CS, Davis MT, McGinley MD, Robinson JH, Bures EJ, Beierle J, Mort J, Courchesne PL, Chen K, Wahl RC, Yu W, Luethy R, Patterson SD: Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. *Proteomics* 2001;1:93–107.

- 19 Thongboonkerd V, McLeish KR, Arthur JM, Klein JB: Proteomic analysis of normal human urinary proteins isolated by acetone precipitation or ultracentrifugation. *Kidney Int* 2002;62:1461–1469.
- 20 Molloy MP, Herbert BR, Walsh BJ, Tyler MI, Traini M, Sanchez JC, Hochstrasser DF, Williams KL, Gooley AA: Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis. *Electrophoresis* 1998;19:837–844.
- 21 Wildgruber R, Harder A, Obermaier C, Boguth G, Weiss W, Fey SJ, Larsen PM, Gorg A: Towards higher resolution: two-dimensional electrophoresis of *Saccharomyces cerevisiae* proteins using overlapping narrow immobilized pH gradients. *Electrophoresis* 2000;21:2610–2616.
- 22 Gorg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W: The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 2000;21:1037–1053.
- 23 Zuo X, Speicher DW: A method for global analysis of complex proteomes using sample prefractionation by solution isoelectrofocusing prior to two-dimensional electrophoresis. *Anal Biochem* 2000;284:266–278.
- 24 Badock V, Steinhilber U, Bommert K, Otto A: Prefractionation of protein samples for proteome analysis using reversed-phase high-performance liquid chromatography. *Electrophoresis* 2001;22:2856–2864.
- 25 Righetti PG, Castagna A, Herbert B: Prefractionation techniques in proteome analysis. *Anal Chem* 2001;73:320A–326A.
- 26 Corthals GL, Molloy MP, Herbert BR, Williams KL, Gooley AA: Prefractionation of protein samples prior to two-dimensional electrophoresis. *Electrophoresis* 1997;18:317–323.
- 27 Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR III: Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* 1999;17:676–682.
- 28 McCormack AL, Schieltz DM, Goode B, Yang S, Barnes G, Drubin D, Yates JR III: Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. *Anal Chem* 1997;69:767–776.
- 29 Washburn MP, Ulaszek R, Deciu C, Schieltz DM, Yates JR III: Analysis of quantitative proteomic data generated via multidimensional protein identification technology. *Anal Chem* 2002;74:1650–1657.
- 30 Gorg A, Obermaier C, Boguth G, Csordas A, Diaz JJ, Madjar JJ: Very alkaline immobilized pH gradients for two-dimensional electrophoresis of ribosomal and nuclear proteins. *Electrophoresis* 1997;18:328–337.
- 31 Gorg A, Obermaier C, Boguth G, Weiss W: Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: Wide pH gradients up to pH 12, longer separation distances and simplified procedures. *Electrophoresis* 1999;20:712–717.
- 32 Yamaguchi Y, Pfeiffer SE: Highly basic myelin and oligodendrocyte proteins analyzed by NEPHGE-two-dimensional gel electrophoresis: Recognition of novel developmentally regulated proteins. *J Neurosci Res* 1999;56:199–205.
- 33 O'Farrell PZ, Goodman HM, O'Farrell PH: High-resolution two-dimensional electrophoresis of basic as well as acidic proteins. *Cell* 1977;12:1133–1141.
- 34 Hampel DJ, Sansome C, Sha M, Brodsky S, Lawson WE, Goligorsky MS: Toward proteomics in uroscopy: Urinary protein profiles after radiocontrast medium administration. *J Am Soc Nephrol* 2001;12:1026–1035.
- 35 MacCoss MJ, McDonald WH, Saraf A, Sadygov R, Clark JM, Tasto JJ, Gould KL, Wolters D, Washburn M, Weiss A, Clark JI, Yates JR III: Shotgun identification of protein modifications from protein complexes and lens tissue. *Proc Natl Acad Sci USA* 2002;99:7900–7905.
- 36 Wolters DA, Washburn MP, Yates JR III: An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* 2001;73:5683–5690.

Jon B. Klein, MD, PhD
Core Proteomics Laboratory
Kidney Disease Program, University of Louisville
570 S. Preston Street, Suite 102, Louisville, KY 40202 (USA)
Tel. +1 502 8520014, Fax +1 502 8524384, E-Mail jon.klein@louisville.edu