Published in final edited form as:
Proteomics Clin Appl. 2008 July 10; 2(7-8): 964. doi:10.1002/prca.200800024.

CE-MS analysis of the human urinary proteome for biomarker discovery and disease diagnostics

Joshua J. Coon[1,2,#], **Petra Zürbig**[3,*,#], **Mohammed Dakna**[3], **Anna F. Dominiczak**[4], **Stéphane Decramer**[5,6,7], **Danilo Fliser**[8,+], **Moritz Frommberger**[9], **Igor Golovko**[3], **David M. Good**[1], **Stefan Herget-Rosenthal**[10,+], **Joachim Jankowski**[11,+], **Bruce A. Julian**[12], **Markus Kellmann**[13], **Walter Kolch**[14], **Ziad Massy**[15,+], **Jan Novak**[12], **Kasper Rossing**[16], **Joost P. Schanstra**[5,6], **Eric Schiffer**[3], **Dan Theodorescu**[17], **Raymond Vanholder**[18,+], **Eva M. Weissinger**[19,+], **Harald Mischak**[3,+], and **Philippe Schmitt-Kopplin**[9]

[1] Department of Chemistry, University of Wisconsin-Madison, Madison, WI, USA

[2] Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA

[3] Mosaiques diagnostics & therapeutics, Hannover, Germany

[4] BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, UK

[5] Department of Renal and Cardiac Remodelling, Inserm, U858/I2MR, Toulouse, France

[6] Université Toulouse III Paul Sabatier, Institut de Médecine Moléculaire de Rangueil, Toulouse, France

[7] Pediatric Nephrology Unit, Hopital des Enfants, CHU de Toulouse, Toulouse, France

[8] Department of Nephrology, Hannover Medical School, Hannover, Germany

[9] Department of Ecological Chemistry, Helmholtz Center Munich German Research Center for Health and Environment, Neuherberg, Germany

[10] Department of Nephrology, University Hospital Essen, University Duisburg-Essen, Germany

[11] Nephrology Departement, Charité Hospital, Berlin, Germany

[12] University of Alabama at Birmingham, Birmingham, AL, USA

[13] Thermo Fisher Scientific, Bremen, Germany

[14] The Beatson Institute for Cancer Research & Sir Henry Wellcome Functional Genomics Facility, University of Glasgow, Glasgow, UK

[15] Departments of Clinical Pharmacology and Nephrology, INSERM, ERI-12, and Amiens University Hospital, UPJV, Amiens, France

[16] Steno Diabetes Centre, Gentofte, Denmark

[17] Department of Urology, University of Virginia, Charlottesville, VA, USA

[18] Department of Internal Medicine, University Hospital Ghent, Ghent, Belgium

[19] Department of Hematology, Hemostasis and Oncology, Hannover Medical School, Hannover, Germany

*Corresponding author at: Mosaiques diagnostics & therapeutics AG, Mellendorfer Str. 7-9, 30625 Hannover, Germany, Tel: +49 511 554744 15, Fax: +49 511 554744 31, zuerbig@mosaiques-diagnostics.com.
[#]These authors have contributed equally to this work.
[+]On behalf of the EUTox consortium

Abstract

Owing to its availability, ease of collection, and correlation with pathophysiology of diseases, urine is an attractive source for clinical proteomics. However, many proteomic studies have had only limited clinical impact, due to factors such as modest numbers of subjects, absence of disease controls, small numbers of defined biomarkers, and diversity of analytical platforms. Therefore, it is difficult to merge biomarkers from different studies into a broadly applicable human urinary proteome database. Ideally, the methodology for defining the biomarkers should combine a reasonable analysis time with high resolution, thereby enabling the profiling of adequate samples and recognition of sufficient features to yield robust diagnostic panels. Capillary electrophoresis coupled to mass spectrometry (CE-MS), which was used to analyze urine samples from healthy subjects and patients with various diseases, is a suitable approach for this task. The database of these datasets compiled from the urinary peptides enabled the diagnosis, classification, and monitoring of a wide range of diseases. CE-MS exhibits excellent performance for biomarker discovery and allows subsequent biomarker sequencing independent of the separation platform. This approach may elucidate the pathogenesis of many diseases, and better define especially renal and urological disorders at the molecular level.

Keywords

Capillary electrophoresis; database; mass spectrometry; proteomics; urine

1 Introduction

Human urine plays a central role in clinical diagnostics. Throughout the centuries, physicians have examined urinary samples from patients to diagnose various disorders. Hermogenes described the color and other attributes of urine as indicators of certain diseases [1]. The human urinary proteome has been investigated extensively to analyze disease processes affecting the kidney and the urogenital tract [2–4]. Whenever the function of these tissues is compromised, changes in the urinary proteome may reflect the role in the generation of the urine. However, urinary proteins originate not only from glomerular filtration, but also from tubular secretion, shed epithelial cells, secreted exosomes, and semen [5–7]. Thus, in principle, urine is a rich source of biomarkers for a wide range of diseases that alter the urinary proteome (proteinuria) [4,8,9]). To realize this potential, large-scale studies are necessary to analyze the human urine proteome, quantitatively and in sufficient detail. The approaches have included a variety of techniques (for further details, see Fliser et al. [10] and Thongboonkerd [11]), *e.g.*, two-dimensional electrophoresis with mass spectrometric and/or immunochemical identification of proteins (2-DE-MS) [12–16], liquid chromatography coupled to mass spectrometry (LC-MS) [14,15,17,18], and surface-enhanced laser desorption ionization mass spectrometry (SELDI-MS) [19].

To date, urinary proteome analyses have revealed more than 1,500 different proteins/peptides (see *e.g.* Castagna et al. [20], Adachi et al. [21]). Although these studies did not attempt to define urinary biomarkers for a specific disease, they clearly highlighted the plethora of information in the urinary proteome and provided some insight for its potential use as a clinical tool. Moreover, the data have been helpful for planning proteomic analyses, by identifying disease-specific proteins and peptides in the urine. This approach has been applied to patients with prostate cancer [14,22–24], bladder cancer [2,25–27], diabetic nephropathy [28,29], chronic renal disease, transplant-associated complications [30–33], myeloma [34,35], and renal dysfunction due to heavy-metal toxicity [36] (for detailed lists see [11]).

Unfortunately, most of these studies have been compromised due to the small numbers of samples (100 samples, at best), and restriction generally to only two diagnostic groups, patients with a single disease and healthy individuals. Therefore, only a few novel potential biomarkers have been validated in these studies. However, proteomes are highly dynamic and directly react to actual (patho-)physiological situations and environmental influences. This feature is an invaluable advantage, as it reflects the current health state of the organism, but it also poses enormous challenges. The associated high degree of heterogeneity suggests that it is crucial to identify panels of markers rather than individual markers [37–39]. A useful method to define such diagnostic panels, therefore, must combine a reasonable analysis time with high analytical resolution for testing many samples and recognition of sufficient features to yield robust diagnostic panels.

Capillary electrophoresis coupled to mass spectrometry (CE-MS) comprises a fast analysis method capable of resolving 1,000 to 4,000 different peptides per sample within approximately 45 minutes [40]. As outlined in more detail in several recent reviews [10,40,41], some of the advantages of CE-MS in comparison to LC-MS are the robustness of CE (towards interfering compounds, precipitates, etc.) and the high comparability of the datasets obtained. This approach was employed to analyze urine samples from healthy volunteers and patients with a variety of different diseases (reviewed in [10,42]). Suitable software solutions are necessary to facilitate processing of standardized raw data, including peak detection, charge assignment, calibration, and database deposition [40,43]. The resulting database consisted of more than 5,000 different polypeptides, characterized by their mass, CE migration time, and MS signal amplitude. These data represent a comprehensive description of the urinary proteome in patients with various (patho)-physiological conditions.

2 CE-MS Methodology

The clinical application of CE-MS demands high reproducibility and comparability of acquired data [11,44]. Previous studies demonstrated that, in contrast to blood, urine is stable for several hours at room temperature [38,45]. This finding is due, in part, to the fact that proteolytic degradation by endogenous proteases is essentially complete once urine is voided. Although CE allows separation of even crude urine samples, salt and higher-molecular-weight proteins interfere with this separation; hence, it is advantageous to remove these compounds in preparation of the sample for analysis. To serve this purpose, the sample is subjected to ultrafiltration in the presence of urea and sodium dodecylsulfate to eliminate protein-protein interactions, and then desalted by size exclusion chromatography. This protocol reliably removes polypeptides larger that 20 kDa and salts [38,46].

The reproducibility of the CE-MS approach was achieved in part by preparing urine samples under standardized preparation/analysis conditions, and stringent quality control. For the detection of narrow CE-separated analyte zones, a fast and sensitive mass spectrometer is necessary. Modern electrospray time-of-flight mass spectrometers (ESI-TOF-MS) provide resolution >10,000 and mass accuracy <10 ppm, suggesting CE-ESI-TOF-MS is a well-suited setup. Each CE-MS analysis consists of about 1,500 single mass spectra. The essential information that must be extracted is identity and quantity of detected polypeptides. The data were evaluated using MosaiquesVisu software (www.proteomiques.com) [40], resulting in a list (raw data list) of peptides/proteins defined by mass, migration time, and ion-counts, serving as a measure of relative abundance.

Different charge states of identical peptides were combined as a single entity, resulting in a list of 1,200–2,000 peptides/proteins per sample. Key to the comparative examination of samples is the ability to reliably retrieve identical polypeptides in consecutive samples. To this end, CE-migration time and mass are used to assign tentative identity to a peptide, enhancing the

resolution of the analysis by utilizing two independent and reproducible parameters. CE-migration time and molecular mass are normalized using ‘internal standards’, peptides found with high frequency in urine [38,47]. Finally, a list of unambiguously identified and standardized peptides of a given sample enables digital compilation of individual data sets to specific polypeptide panels that are used for biomarker definition.

To improve mass accuracy, TOF-MS-derived masses were calibrated using 80 precisely FT-ICR-characterized reference masses (mass deviations <0.5 ppm) [48]. High FT-ICR MS resolution enabled an accurate analysis of the first isotope signal (z>6), which is crucial for determination of the exact mass of high molecular weight peptides. Therefore, the mean mass deviation of the ‘FT-ICR-calibrated’ TOF masses improved from 19±30 ppm to 3±9 ppm, referring to theoretical masses.

3 Human Urinary Proteome Database

The calibrated data sets of currently 3,687 human urine samples (with an average, 1,724 peptides/proteins were detected in each individual urine sample, ranging from 983 to 4,094) were deposited in a Microsoft SQL database, enabling digital data compilation [49]. Subsequently, data clustering defined 116,869 different peptides/proteins. Each peptide was assigned a unique identification number (Protein ID). To eliminate peptides of apparently low significance that appeared sporadically, only those peptides present in more than 20% of the urine samples in at least one group (samples from patients with same disease) were further investigated. This noise-filtering process significantly reduced the number of peptides for analysis to 5,010 “relevant” different peptides, characterized by molecular mass and normalized CE-migration time. The filtered data of all individual samples are available on the mosaiques diagnostics webpage[a]. Currently (see version 2.0 at the mosaiques webpage), the database contains datasets from patients of 28 different pre-selected pathophysiological conditions (see figure 1). In contrast to the SELDI technology and as outlined in more detail elsewhere [10,40,42],, the high resolution obtained by the use of two identifying parameters in CE-MS allow the reproducible definition of a potential biomarker based on accurate mass and migration time. These parameters enable not only robust definition of potential biomarkers, but in general also targeted sequencing, which appears impossible in SELDI-based approaches. The avoidance of any pre-analytical manipulation that will result in high variability of datasets (e.g. affinity matrices, ion-exchange or reversed phase material) also clearly distinguishes this approach from the SELDI-based studies, where the different conditions and chip surfaces preclude comparison of the data from different experiments. Further, the inacceptably low resolution of the mass spectrometer used in combination with a missing second identifying parameter (as the migration time here)

3.1 Identification of Naturally Occurring Peptides

Recently, several groups reported the sequencing of an array of urinary proteins [20,21]. While these data impressively demonstrated a vast number of urinary proteins, the potential information critical, or even mandatory, for their application in the definition of biomarkers was unfortunately absent.

All of these studies used tryptic digests of urinary proteins, and the sequences of the peptides allow, with variable degree of confidence, the preliminary assignment of a protein to this sequence. However, due to the tryptic digest (or similar manipulation), it is not possible to define the exact nature of the proteins actually present in the urine at the time of sampling. The naturally occurring protein(s) will generally not be identical to the theoretical protein in the

[a]http://mosaiques-diagnostics.de/diapatpcms/mosaiquescms/front_content.php?idcat=257/

database (*e.g.*, albumin precursor), but one or several variably post-translationally modified proteins. This also implies that several different proteins (originating from the same pre-protein and a result of different post-translational processing) will in fact give rise to identical tryptic peptides. Consequently, differentially modified proteins cannot be distinguished. However, such modifications are often the hallmark of the potential biomarker, *e.g.*, advanced glycation end-products are markers for uremia [50], repetitive urinary albumin and alpha-1-antitrypsin fragments have been described as potential biomarkers for distinct nephrotic syndrdoms [51], distinct urinary collagen fragments appear to be markers for diabetes and diabetic nephropathy [37,52]. As a consequence, the information on the exact protein actually present is required. In fact, the definition of a potential biomarker by several physical parameters (e.g. mass, retention time, isoelectric point, etc.) appears more advantageous than the mere definition of the biomarker by the sequence of a (theoretical) precursor [53].

While the exact sequence of a biomarker is not an absolute requirement for its clinical/diagnostic use, it may offer further insight into the pathogenesis of the disease, (patho-) physiological mechanisms, and aid in the design of relevant therapy. Further, in the absence of sequence the assessment of the potential biomarker is restricted solely to the technology employed in the tentative identification, CE-MS. Hence, sequence analysis of naturally occurring urinary peptides completes the content of the human urinary proteome database.

For further validation of polypeptides listed within the urine proteome database (all ongoing identified naturally occurring urinary peptides can be accessed on the mosaiques webpage[a], version 2.0), different MS/MS technologies have been applied for sequencing [54,55]. In this context, the direct and strict dependance of CE migration time on the charge density of the analyte represents a valuable key feature of the technology for validation of sequences obtained via MS/MS analysis. At the assay pH of 2.2, the effective charge of the analyzed polypeptides depends strictly on the number of basic amino acid residues, including the free N-terminus [55]. Therefore, it is not a prerequisite to use CE-separation for MS/MS sequencing, as the number of basic amino acids in combination with accurate mass permits the correlation of the sequence (including number of positively charged residues) to a peptide in the database. Most frequently, different fragments of collagens, common blood proteins (*e.g.* alpha-1-antitrypsin, hemoglobin, serum albumin, and fibrinogen), and uromodulin were identified (table 1). Many precursor proteins were also found by other research groups [20,21]. However, in general the proteolytically processed native peptides in urine detected by the ‘top-down’ approach [56] would remain undetected by the ‘bottom-up’ techniques.

Most of these naturally occurring urinary peptides are the result of proteolytic activity. Extracellular proteases may reflect the activity of a specific disease or its progression [57]. Complex changes in protease activities may be more readily recognized by the pattern of proteolytic fragments generated, rather than by direct assessment of the activity of a specific protease [58]. CE-MS analysis may be suitable to indicate the regulated activity of proteases and protease inhibitors by displaying potential products and monitoring their concentrations.

Therefore, top-down approaches appear to be suitable to show the regulated activity of proteases and protease inhibitors by displaying potential products and enabling monitoring of their concentrations. This feature outweighs several obstacles encountered in the sequencing of native peptides: Major obstacles are the frequently occurring post-translational modifications (PTM) that change the mass, that then differs from the theoretical mass in the database, and the higher degree of freedom, as pre-set terminal arginine or lysine cleavage sites used by trypsin cannot be employed in the database search. Further, search algorithms are generally adapted to the needs of tryptic digests, which differ greatly from the requirements for *de novo* sequencing of naturally occurring peptides or proteins (see also *e.g.*, [10,42]). Identification of the full sequence may be challenging due to elimination of water from same

amino acid residues (asparagine, aspartic acid, glutamine, glutamic acid) [59,60] or loss of proline residues caused by partial fragmentation. In this case, typically only MS^n methods using an ion trap device provide satisfactory results [55].

3.2 Biomarkers for Disease

Although CE-MS analysis of urine samples can identify biomarkers for a variety of diseases of the kidney and urogenital tract [47,61–64], the variability of polypeptides presents a serious handicap. The excretion of some polypeptides varies significantly during the day, most likely as a consequence of physical activity, diet, or medication effect [65,66]. As a result, reproducibility at the level of single polypeptides is limited. Hence, the clinical usefulness of a single biomarker may be of only modest value, even if the accuracy and reproducibility of the test are optimal. In contrast, a polypeptide panel consisting of an array of well defined biomarkers is more robust, as changes in individual analytes will not lead to marked changes of the panel, consequently the classification result. This becomes evident when comparing single biomarkers with the scoring of a biomarker panel in blinded assessment (e.g. Theodorescu et al., [38], Zimmerli et al., [54] Rossing et al. [52]).

The comparability of the datasets from the CE-MS analysis enables the tentative definition of biomarkers that show statistically significant (even when adjustment for multiple testing are made) changes in a certain disease. As shown exemplarily in figure 2, potential biomarkers can be defined by comparing data obtained from controls and patients with different, distinct diseases. These biomarkers that are validated by appropriate statistics can be combined to a panel of biomarkers from human urine that enable distinction of patients with a certain disease from healthy subjects and of patients with various other disorders. Diseases affecting urine composition may be directly related to the urogenital tract, but this association may at first glance be obscure. Acute graft-versus-host disease and cardiovascular disease are two examples of processes that fit this category. Subsequently these panels can be used for diagnostic purposes. The validation of CE-MS-identified urinary peptide panels for diagnosis and prognosis in blinded studies has already been published in several recent reports [31,38, 39,46,52,54,67]. Theodorescu *et al.* [38] recently used CE-MS to assay more than 600 samples, including 180 samples as a validation set that were examined in a blinded manner. The discovered biomarkers correctly classified all blinded urothelial cancer samples and normal controls; however, nine of 138 patients with various chronic kidney diseases or nephrolithiasis were incorrectly classified as having urothelial cancer.

Urinary proteome analysis may also be an excellent tool for fast, noninvasive, and unbiased monitoring of disease progression or response to therapy. In a randomized, double-blinded study, Rossing *et al.* [64] evaluated the treatment of macroalbuminuric patients with daily doses of 8, 16, and 32 mg of Candesartan or placebo for 2 month. Candesartan treatment resulted in a significant change in 15 of 113 proteins that are characteristic for diabetic renal damage.

Kaiser *et al.* [30] defined biomarkers for graft-*versus*-host disease after bone marrow transplantation using CE-MS based urine proteomics. This preliminary observation was validated in a recent prospective multicenter study with more than 600 urine samples from more than 100 patients [68].

In addition, the work of Decramer *et al.* [67] can be interpreted as first proof of the capability of CE-MS based proteomics for early diagnosis. The authors analyzed urinary polypeptides from infants with ureteropelvic junction (UPJ) obstruction to predict a need for surgical correction. As evident from the results, the authors identified and, in a prospective blinded study, validated polypeptides markers that enable diagnosis of the severity of the obstruction; this resulted in the correct prediction of clinical evolution of 34/36 neonates (resulting in a correct prediction in 94% of the cases) with UPJ obstruction several months in advance.

4 Conclusion and Outlook

Proteome analysis of urine for biomarker discovery mandates analytical methods with high reproducibility and comparability. Besides biomarker defining experimental parameters (*e.g.* migration time, molecular weight, and amplitude), biomarker sequence is an indispensable cornerstone for deeper insights into the pathological pathways or for made-to-measure therapeutic drug design.

CE-MS enables reproducible and robust high-resolution analysis of several thousand low molecular weight urinary proteins/peptides within a reasonable time frame. The analysis of more than 3,500 urine samples from diseased and healthy individuals enabled the establishment of a database of naturally occurring urinary peptides. This unique database serves as a broad basis for the definition and validation of biomarkers for diagnosis/prognosis/monitoring of a wide range of diseases using biomarker patterns.

These signature patterns seem to reflect primary pathogenetic changes as well as the reaction of the organism to diseases. Hence, their usefulness extends far beyond the applicability to diseases of the urogenital tract, and may be universally applicable to any disease that produces systemic changes.

While genetic analysis may predict the risk of a disease, proteomics with its potential for dynamic monitoring may define at which point the risk manifests as disease, and also allow assessment of the response to therapy. Thus, these two methods are complementary, but we anticipate that proteomics may have a greater role in individualized medicine. As we begin to understand the unique differences between patients in their response to therapy, methods to objectively measure these responses will become of prime importance to tailor the therapy to the individual patient. In this effort, proteomics has the advantage that monitoring of therapy in real time, and adjustments can be made accordingly. This vision is within reach, but its realization depends entirely on establishment of databases that allow investigators to quickly compare patients' profiles against those of other patients or healthy controls in a robust manner. Thus, we contend that the human urinary proteome database derived from CE-MS analysis is a seminal step in this direction; we anticipate that the availability of such databases will significantly improve the diagnostic and therapeutic options available to many patients.

Acknowledgments

This work was supported in part by grant #0312939 from BioProfil 'Funktionelle Genomanalyse' and by grant #203.19-32329-5-461 from the Lower Saxony Ministry of Economy. Eurotransbio partially funded by grant #ETB-2006-016 to the Urosysteomics consortium (www.urosysteomics.com) to PZ, MD, ES, and HM. From the European Union HM and KR were supported by grant #LSHM-CT-2005-018733 from PREDICTIONS (PREvention of DIabetic ComplicaTIONS), and HM and AFD were supported by grant #LSHM-CT-2006-037093 from InGenious HyperCare. Furthermore, AFD was supported by grant # RG/02/012 from the British Heart Foundation (BHF). WK was supported by grant # 505520 from the FP6 EU project 'Interaction Proteome' and the Wellcome Trust JIF proteomics grant #29240. BAJ and JN were supported in part by grants #DK61525, #DK78244, and #DK71802 from the NIH and by General Clinical Research Centre of the University of Alabama at Birmingham (#M01 RR00032).

Abbreviations

MIAPE	Minimum Information About a Proteomics Experiment
ppm	parts per million
SQL	Structured Query Language
UPJ	ureteropelvic junction

References

1. Iorio L, Avagliano F. Observations on the Liber medicine orinalibus by Hermogenes. Am J Nephrol 1999;19:185–188. [PubMed: 10213817]
2. Celis JE, Rasmussen HH, Vorum H, Madsen P, et al. Bladder squamous cell carcinomas express psoriasin and externalize it to the urine. J Urol 1996;155:2105–2112. [PubMed: 8618345]
3. Delanghe J. Use of specific urinary proteins as diagnostic markers for renal disease. Acta Clin Belg 1997;52:148–153. [PubMed: 9350295]
4. Marshall T, Williams K. Two-dimensional electrophoresis of human urinary proteins following concentration by dye precipitation. Electrophoresis 1996;17:1265–1272. [PubMed: 8855415]
5. Oh J, Pyo JH, Jo EH, Hwang SI, et al. Establishment of a near-standard two-dimensional human urine proteomic map. Proteomics 2004;4:3485–3497. [PubMed: 15529407]
6. Pieper R, Gatlin CL, McGrath AM, Makusky AJ, et al. Characterization of the human urinary proteome: a method for high-resolution display of urinary proteins on two-dimensional electrophoresis gels with a yield of nearly 1400 distinct protein spots. Proteomics 2004;4:1159–1174. [PubMed: 15048996]
7. 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. [PubMed: 12234320]
8. Shihabi ZK, Konen JC, O'Connor ML. Albuminuria vs urinary total protein for detecting chronic renal disorders. Clin Sci 1991;37:621–624.
9. Yudkin JS, Forrest RD, Jackson CA. Microalbuminuria as predictor of vascular disease in non-diabetic subjects. Islington Diabetes Survey. Lancet 1988;2:530–533. [PubMed: 2900920]
10. Fliser D, Novak J, Thongboonkerd V, Argiles A, et al. Advances in urinary proteome analysis and biomarker discovery. J Am Soc Nephrol 2007;18:1057–1071. [PubMed: 17329573]
11. Thongboonkerd V. Recent progress in urinary proteomics. Proteomics Clin Appl 2007;1:780–791.
12. Anderson NG, Anderson NL, Tollaksen SL. Proteins of human urine. I. Concentration and analysis by two-dimensional electrophoresis. Clin Chem 1979;25:1199–1210. [PubMed: 455639]
13. Bueler MR, Wiederkehr F, Vonderschmitt DJ. Electrophoretic, chromatographic and immunological studies of human urinary proteins. Electrophoresis 1995;16:124–134. [PubMed: 7537657]
14. M'koma AE, Blum DL, Norris JL, Koyama T, et al. Detection of pre-neoplastic and neoplastic prostate disease by MALDI profiling of urine. Biochem Biophys Res Commun 2007;353:829–834. [PubMed: 17194448]
15. Spahr CS, Davis MT, McGinley MD, Robinson JH, et al. Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. Proteomics 2001;1:93–107. [PubMed: 11680902]
16. Zerefos PG, Vougas K, Dimitraki P, Kossida S, et al. Characterization of the human urine proteome by preparative electrophoresis in combination with 2-DE. Proteomics 2006;6:4346–4355. [PubMed: 16807940]
17. Baer JC, Hjelm M. Analysis of urinary proteins by liquid chromatography. Ann Clin Biochem 1994;31 (Pt 4):315–326. [PubMed: 7979095]
18. Jurgens M, Appel A, Heine G, Neitz S, et al. Towards characterization of the human urinary peptidome. Comb Chem High Throughput Screen 2005;8:757–765. [PubMed: 16464162]
19. Cadieux PA, Beiko DT, Watterson JD, Burton JP, et al. Surface-enhanced laser desorption/ionization-time of flight-mass spectrometry (SELDI-TOF-MS): a new proteomic urinary test for patients with urolithiasis. J Clin Lab Anal 2004;18:170–175. [PubMed: 15103681]
20. Castagna A, Cecconi D, Sennels L, Rappsilber J, et al. Exploring the hidden human urinary proteome via ligand library beads. J Proteome Res 2005;4:1917–1930. [PubMed: 16335936]
21. Adachi J, Kumar C, Zhang Y, Olsen JV, Mann M. The human urinary proteome contains more than 1500 proteins including a large proportion of membranes proteins. Genome Biol 2006;7:R80.1–R80.16. [PubMed: 16948836]
22. Edwards JJ, Anderson NG, Tollaksen SL, von Eschenbach AC, Guevara J Jr. Proteins of human urine. II. Identification by two-dimensional electrophoresis of a new candidate marker for prostatic cancer. Clin Chem 1982;28:160–163. [PubMed: 7055903]

23. Grover PK, Resnick MI. High resolution two-dimensional electrophoretic analysis of urinary proteins of patients with prostatic cancer. Electrophoresis 1997;18:814–818. [PubMed: 9194613]
24. Rehman I, Azzouzi AR, Catto JW, Allen S, et al. Proteomic analysis of voided urine after prostatic massage from patients with prostate cancer: a pilot study. Urology 2004;64:1238–1243. [PubMed: 15596215]
25. Irmak S, Tilki D, Heukeshoven J, Oliveira-Ferrer L, et al. Stage-dependent increase of orosomucoid and zinc-alpha2-glycoprotein in urinary bladder cancer. Proteomics 2005;5:4296–4304. [PubMed: 16196100]
26. Rasmussen HH, Orntoft TF, Wolf H, Celis JE. Towards a comprehensive database of proteins from the urine of patients with bladder cancer. J Urol 1996;155:2113–2119. [PubMed: 8618346]
27. Saito M, Kimoto M, Araki T, Shimada Y, et al. Proteome analysis of gelatin-bound urinary proteins from patients with bladder cancers. Eur Urol 2005;48:865–871. [PubMed: 15964125]
28. Sharma K, Lee S, Han S, Lee S, et al. Two-dimensional fluorescence difference gel electrophoresis analysis of the urine proteome in human diabetic nephropathy. Proteomics 2005;5:2648–2655. [PubMed: 15984045]
29. Thongboonkerd V, Barati MT, McLeish KR, Benarafa C, et al. Alterations in the renal elastin-elastase system in type 1 diabetic nephropathy identified by proteomic analysis. J Am Soc Nephrol 2004;15:650–662. [PubMed: 14978167]
30. Kaiser T, Kamal H, Rank A, Kolb HJ, et al. Proteomics applied to the clinical follow-up of patients after allogeneic hematopoietic stem cell transplantation. Blood 2004;104:340–349. [PubMed: 15054046]
31. Wittke S, Haubitz M, Walden M, Rohde F, et al. Detection of acute tubulointerstitial rejection by proteomic analysis of urinary samples in renal transplant recipients. Am J Transplant 2005;5:2479–2488. [PubMed: 16162198]
32. Schaub S, Rush D, Wilkins J, Gibson IW, et al. Proteomic-based detection of urine proteins associated with acute renal allograft rejection. J Am Soc Nephrol 2004;15:219–227. [PubMed: 14694176]
33. Schaub S, Wilkins JA, Antonovici M, Krokhin O, et al. Proteomic-based identification of cleaved urinary beta2-microglobulin as a potential marker for acute tubular injury in renal allografts. Am J Transplant 2005;5:729–738. [PubMed: 15760396]
34. Harrison HH. The "ladder light chain" or "pseudo-oligoclonal" pattern in urinary immunofixation electrophoresis (IFE) studies: a distinctive IFE pattern and an explanatory hypothesis relating it to free polyclonal light chains. Clin Chem 1991;37:1559–1564. [PubMed: 1909941]
35. Tichy M, Stulik J, Kovarova H, Mateja F, Urban P. Analysis of monoclonal immunoglobulin light chains in urine using two-dimensional electrophoresis. Neoplasma 1995;42:31–34. [PubMed: 7617072]
36. Myrick JE, Caudill SP, Robinson MK, Hubert IL. Quantitative two-dimensional electrophoretic detection of possible urinary protein biomarkers of occupational exposure to cadmium. Appl Theor Electrophor 1993;3:137–146. [PubMed: 8512944]
37. Rossing K, Mischak H, Rossing P, Schanstra JP, et al. The urinary proteome in diabetes and diabetes-associated complications: new ways to assess disease progression and evaluate therapy. Proteomics Clin Appl. 2008 in press.
38. Theodorescu D, Wittke S, Ross MM, Walden M, et al. Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. Lancet Oncol 2006;7:230–240. [PubMed: 16510332]
39. Theodorescu D, Schiffer E, Bauer HW, Douwes F, Eichhorn F, Polley R, Schmidt T, Schöfer W, Zurbig P, Good DM, Coon JJ, Mischak H. Discovery and validation of urinary biomarkers for prostate cancer. Proteomics Cli ppl. 2008
40. Kolch W, Neususs C, Pelzing M, Mischak H. Capillary electrophoresis-mass spectrometry as a powerful tool in clinical diagnosis and biomarker discovery. Mass Spectrom Rev 2005;24:959–977. [PubMed: 15747373]
41. Mischak H, Julian BA, Novak J. High-resolution proteome/peptidome analysis of peptides and low-molecular-weight proteins in urine. Proteomics Clin Appl 2007;1:792–804. [PubMed: 20107618]
42. Mischak H, Julian BA, Novak J. High-resolution proteome/peptidome analysis of peptides and low-molecular-weight proteins in urine. Proteomics ClinAppl 2007;1:792–804.

43. Weissinger EM, Hertenstein B, Mischak H, Ganser A. Online coupling of capillary electrophoresis with mass spectrometry for the identification of biomarkers for clinical diagnosis. Expert Rev Proteomics 2005;2:639–647. [PubMed: 16209644]
44. Mischak H, Apweiler R, Banks RE, Conaway M, et al. Clinical Proteomics: a need to define the field and to begin to set adequate standards. Proteomics Clin Appl 2007;1:148–156.
45. Schaub S, Wilkins J, Weiler T, Sangster K, et al. Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. Kidney Int 2004;65:323–332. [PubMed: 14675066]
46. Julian BA, Wittke S, Novak J, Good DM, et al. Electrophoretic methods for analysis of urinary polypeptides in IgA-associated renal diseases. Electrophoresis 2007;28:4469–4483. [PubMed: 18004714]
47. Weissinger EM, Wittke S, Kaiser T, Haller H, et al. Proteomic patterns established with capillary electrophoresis and mass spectrometry for diagnostic purposes. Kidney Int 2004;65:2426–2434. [PubMed: 15149356]
48. Frommberger M, Zürbig P, Jantos J, Krahn T, et al. Peptidomic analysis of rat urine using capillary electrophoresis coupled to mass spectrometry. Proteomics Clin Appl 2007;1:650–660.
49. Schiffer E, Mischak H, Novak J. High resolution proteome/peptidome analysis of body fluids by capillary electrophoresis coupled with MS. Proteomics 2006;6:5615–5627. [PubMed: 16991199]
50. Thornalley PJ, Argirova M, Ahmed N, Mann VM, et al. Mass spectrometric monitoring of albumin in uremia. Kidney Int 2000;58:2228–2234. [PubMed: 11044246]
51. Candiano G, Musante L, Bruschi M, Petretto A, et al. Repetitive fragmentation products of albumin and alpha1-antitrypsin in glomerular diseases associated with nephrotic syndrome. J Am Soc Nephrol 2006;17:3139–3148. [PubMed: 17005933]
52. Rossing K, Mischak H, Dakna M, Zurbig P, et al. Proteomic discovery and validation of urinary biomarkers for diabetes and chronis renal disease. J Am Soc Nephrol. 2008 in press.
53. Good DM, Thongboonkerd V, Novak J, Bascands JL, et al. Body Fluid Proteomics for Biomarker Discovery: Lessons from the Past Hold the Key to Success in the Future. J Proteome Res 2007;6:4549–4555. [PubMed: 17970587]
54. Zimmerli LU, Schiffer E, Zurbig P, Kellmann M, et al. Urinary proteomics biomarkers in coronary artery disease. Mol Cell Proteomics 2008;7:290–298. [PubMed: 17951555]
55. Zurbig P, Renfrow MB, Schiffer E, Novak J, et al. Biomarker discovery by CE-MS enables sequence analysis via MS/MS with platform-independent separation. Electrophoresis 2006;27:2111–2125. [PubMed: 16645980]
56. Pisitkun T, Johnstone R, Knepper MA. Discovery of Urinary Biomarkers. Mol Cell Proteomics 2006;5:1760–1771. [PubMed: 16837576]
57. Candiano G, Musante L, Bruschi M, Petretto A, et al. Repetitive fragmentation products of albumin and alpha1-antitrypsin in glomerular diseases associated with nephrotic syndrome. J Am Soc Nephrol 2006;17:3139–3148. [PubMed: 17005933]
58. Villanueva J, Shaffer DR, Philip J, Chaparro CA, et al. Differential exoprotease activities confer tumor-specific serum peptidome patterns. J Clin Invest 2006;116:271–284. [PubMed: 16395409]
59. Geiger T, Clarke S. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. J Biol Chem 1987;262:785–794. [PubMed: 3805008]
60. Stephenson RC, Clarke S. Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. J Biol Chem 1989;264:6164–6170. [PubMed: 2703484]
61. Haubitz M, Wittke S, Weissinger EM, Walden M, et al. Urine protein patterns can serve as diagnostic tools in patients with IgA nephropathy. Kidney Int 2005;67:2313–2320. [PubMed: 15882273]
62. Mischak H, Kaiser T, Walden M, Hillmann M, et al. Proteomic analysis for the assessment of diabetic renal damage in humans. Clin Sci (Lond) 2004;107:485–495. [PubMed: 15281910]
63. Neuhoff N, Kaiser T, Wittke S, Krebs R, et al. Mass spectrometry for the detection of differentially expressed proteins: a comparison of surface-enhanced laser desorption/ionization and capillary electrophoresis/mass spectrometry. Rapid Commun Mass Spectrom 2004;18:149–156. [PubMed: 14745763]

64. Rossing K, Mischak H, Parving HH, Christensen PK, et al. Impact of diabetic nephropathy and angiotensin II receptor blockade on urinary polypeptide patterns. Kidney Int 2005;68:193–205. [PubMed: 15954909]

65. Fliser D, Wittke S, Mischak H. Capillary electrophoresis coupled to mass spectrometry for clinical diagnostic purposes. Electrophoresis 2005;26:2708–2716. [PubMed: 15966014]

66. Sniehotta M, Schiffer E, Zurbig P, Novak J, Mischak H. CE - a multifunctional application for clinical diagnosis. Electrophoresis 2007;28:1407–1417. [PubMed: 17427258]

67. Decramer S, Wittke S, Mischak H, Zurbig P, et al. Predicting the clinical outcome of congenital unilateral ureteropelvic junction obstruction in newborn by urinary proteome analysis. Nat Med 2006;12:398–400. [PubMed: 16550189]

68. Weissinger EM, Schiffer E, Hertenstein B, Ferrara JL, et al. Proteomic patterns predict acute graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. Blood 2007;109:5511–5519. [PubMed: 17339419]

69. Westfall, PH.; Young, SS. Resampling-based Multiple Testing: Examples and Methods for P-Value Adjustment. Wiley; New York: 1993.

70. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc B (Methodological) 1995:125–133.

71. Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics 2003;19:368–375. [PubMed: 12584122]

72. Abdi, H. Bonferroni and Sidak corrections for multiple comparisons. Sage; Thousand Oaks (CA): 2007.

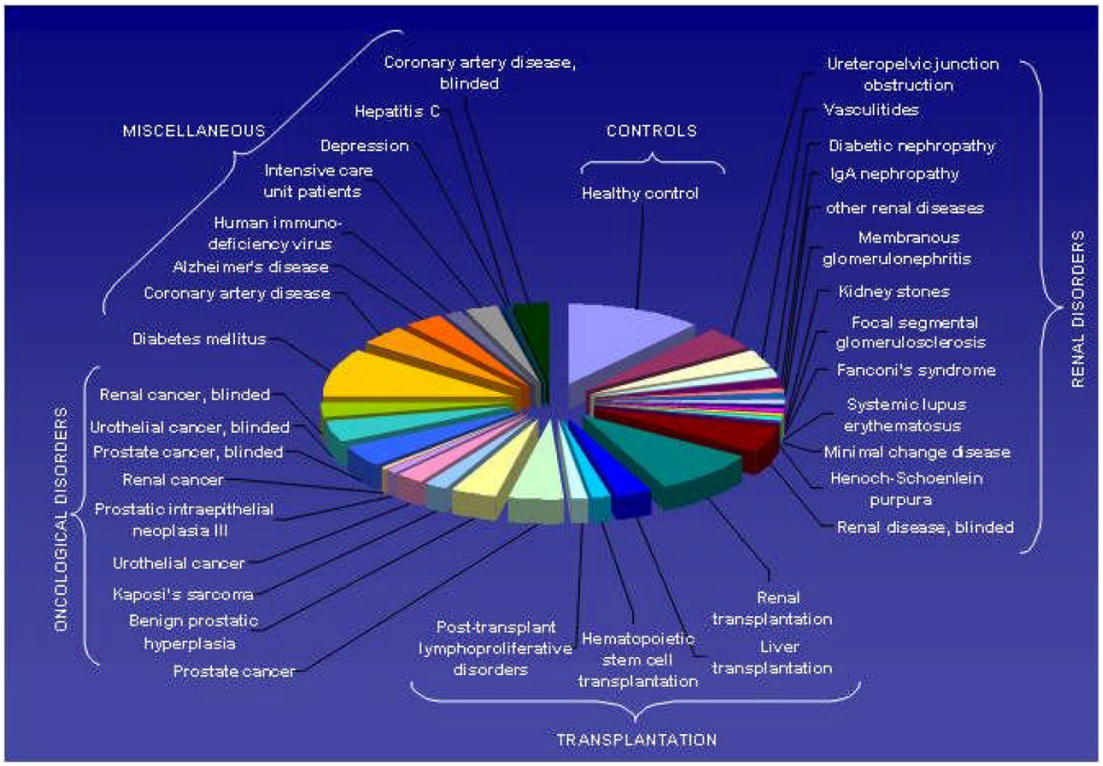

Figure 1.
Disease conditions currently represented in the human urinary proteome database.

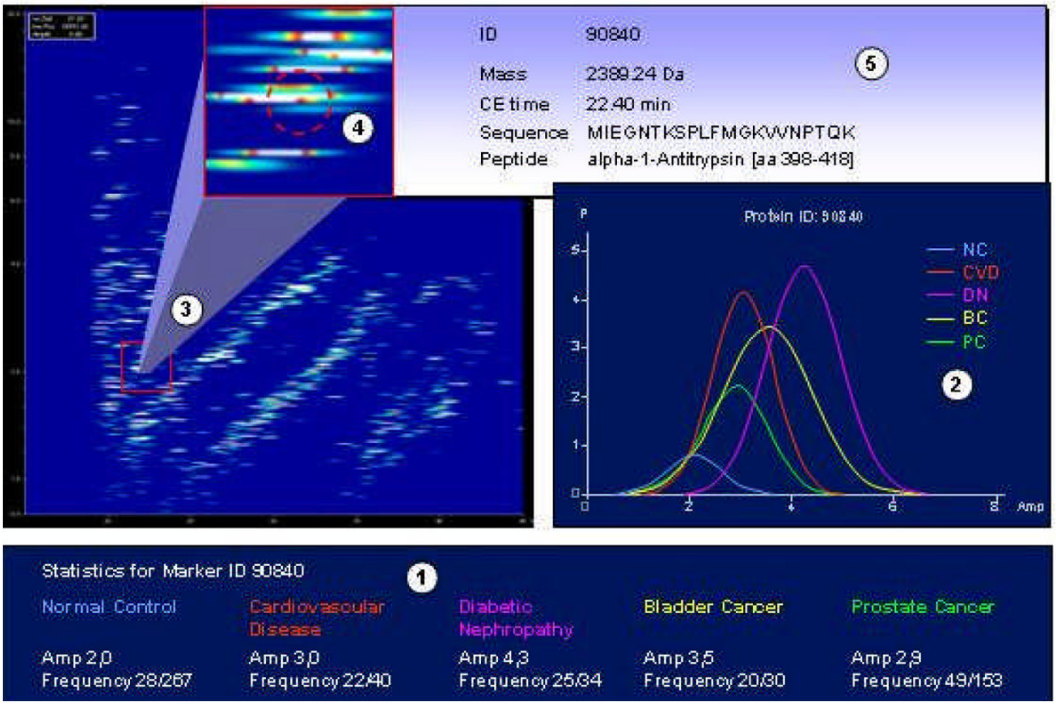

Figure 2.

Graphic depiction of the discovery of potential biomarkers for diabetic nephropathy. CE-MS datasets from control and patients with prostate cancer, bladder cancer, and cardiovascular disease are compared to data obtained from patients with diabetic nephropathy (1) using appropriate statistics (adjustments for multiple testing as described in e.g [69–72]). Potential biomarkers that show significant differences in amplitude and/or distribution (2) are located in the database (3). The clustering of the biomarker (with respect to deviation) is examined in comparison to neighboring peptides (4). If found appropriate, ID, mass, normalized migration time, and, if known, sequence can be retrieved from the database (5).

Table 1

Distribution of 443 native human urinary peptides identified with respect to their protein precursor (described by SwissProt protein name for *Homo sapiens* and gene symbol) derived from the currently available sequence list (version 2.0) of mosaiques diagnostics homepage[a]. Comparison of the located peptides to other references [20,21].

			Proteins detected by:	
Number of peptides	**Protein name**	**Gene symbol**	**Adachi et al.**	**Castagna et al.**
157	Collagen alpha-1 (I) chain	*COL1A1*	yes	no
69	Collagen alpha-1 (III) chain	*COL3A1*	yes	no
24	Alpha-1-antitrypsin	*SERPINA1*	yes	yes
24	Collagen alpha-2 (I) chain	*COL1A2*	yes	no
19	Hemoglobin subunit beta	*HBB*	yes	yes
18	Uromodulin	*UMOD*	yes	yes
17	Hemoglobin subunit alpha	*HBA1, HBA2*	yes	no
16	Serum albumin	*ALB*	yes	yes
14	Fibrinogen alpha chain	*FGA*	yes	no
12	Beta-2-microglobulin	*B2M*	yes	yes
6	Polymeric-immunoglobulin receptor	*PIGR*	yes	yes
3	Alpha-2-HS-glycoprotein	*AHSG*	yes	yes
3	Collagen alpha-1 (II) chain	*COL2A1*	no	no
3	Membrane associated progesterone receptor component 1	*PGRMC1*	yes	no
3	Osteopontin	*SPP1*	yes	no
3	Transthyretin precursor (Prealbumin)	*TTR*	yes	yes
2	Alpha-1-microglobulin	*AMBP*	yes	yes
2	Apolipoprotein A-I	*APOA1*	no	yes
2	CD99 antigen	*CD99*	no	no
2	Clusterin	*CLU*	yes	yes
2	Collagen alpha-1 (XVIII) chain	*COL18A1*	yes	no
2	Epithelial-cadherin	*CDH1*	yes	yes
2	Insulin; includes C peptide	*INS*	no	no
2	Insulin-like growth factor II	*IGF2*	yes	no
2	ProSAAS	*PCSK1N*	yes	no
2	Prostaglandin-H2 D-isomerase	*PTGDS*	yes	yes
1	Alpha-1-acid glycoprotein 1	*ORM1*	yes	yes
1	Alpha-1B-glycoprotein	*A1BG*	yes	yes
1	Antithrombin-III	*SERPINC1*	yes	no
1	Basement membrane-specific heparan sulfate proteoglycan core protein	*HSPG2*	yes	yes
1	Collagen alpha-1 (XIX) chain	*COL19A1*	no	no
1	Collagen alpha-1 (XV) chain	*COL15A1*	yes	no
1	Collagen alpha-1 (XVII) chain	*COL17A1*	no	no
1	Collagen alpha-1 (XXII) chain	*COL22A1*	no	no

			Proteins detected by:	
Number of peptides	**Protein name**	**Gene symbol**	**Adachi et al.**	**Castagna et al.**
1	Collagen alpha-2 (VIII) chain	*COL8A2*	no	no
1	Collagen alpha-3 (IX) chain	*COL9A3*	no	no
1	Complement factor B	*CFB*	yes	no
1	Cystatin-B	*CSTB*	yes	no
1	Fibrinogen beta chain	*FGB*	no	no
1	Fillagrin	*FLG*	yes	no
1	Gelsolin	*GSN*	yes	yes
1	Hemoglobin subunit delta	*HBD*	yes	no
1	Histone H2B type 1	*HIST1H2B*	no	no
1	Ig kappa chain C region	*IGKC*	yes	yes
1	Ig kappa chain V-III region	none	no	yes
1	Ig lambda chain C regions	*IGLC1*	yes	yes
1	Josephin-1	*JOSD1*	no	no
1	Liprin-beta-2	*PPFIBP2*	no	no
1	Microfibrillar-associated protein 5	*MFAP5*	no	no
1	Neurosecretory protein VGF	*VGF*	yes	no
1	Peptidoglycan recognition protein	*PGLYRP1*	yes	yes
1	PREDICTED: similar to Cyclin G-associated kinase	*GAK*	no	no
1	Psoriasis susceptibility 1 candidate gene 2 protein	*PSORS1C2*	yes	no
1	PX domain-containing protein kinase-like protein	*PXK*	no	no
1	Secreted and transmembrane protein 1	*SECTM1*	yes	yes
1	Sodium/potassium-transporting ATPase gamma chain	*FXYD2*	yes	no
1	Zinc finger CCHC domain-containing protein 3	*ZCCHC3*	no	no
1	Zinc finger protein 653	*ZNF653*	no	no