Naturally Occurring Human Urinary Peptides for Use in Diagnosis of Chronic Kidney Disease*⑤

David M. Good,[a,b] **Petra Zürbig,**[b,c,d,e,ww] **Àngel Argilés,**[f] **Hartwig W. Bauer,**[g]
Georg Behrens,[h] **Joshua J. Coon,**[a,i] **Mohammed Dakna,**[c] **Stéphane Decramer,**[d,j,k]
Christian Delles,[l] **Anna F. Dominiczak,**[d,l,m] **Jochen H. H. Ehrich,**[n] **Frank Eitner,**[o]
Danilo Fliser,[d,p,q] **Moritz Frommberger,**[r] **Arnold Ganser,**[s] **Mark A. Girolami,**[t]
Igor Golovko,[c] **Wilfried Gwinner,**[u] **Marion Haubitz,**[u] **Stefan Herget-Rosenthal,**[q,v]
Joachim Jankowski,[d,q,w,x] **Holger Jahn,**[y] **George Jerums,**[z] **Bruce A. Julian,**[aa]
Markus Kellmann,[bb] **Volker Kliem,**[cc] **Walter Kolch,**[dd,ee] **Andrzej S. Krolewski,**[ff,gg]
Mario Luppi,[hh] **Ziad Massy,**[q,ii] **Michael Melter,**[jj] **Christian Neusüss,**[kk] **Jan Novak,**[aa]
Karlheinz Peter,[ll] **Kasper Rossing,**[mm] **Harald Rupprecht,**[nn] **Joost P. Schanstra,**[d,j,oo]
Eric Schiffer,[c,ww] **Jens-Uwe Stolzenburg,**[pp] **Lise Tarnow,**[mm,qq] **Dan Theodorescu,**[d,rr,ss]
Visith Thongboonkerd,[d,tt] **Raymond Vanholder,**[q,uu] **Eva M. Weissinger,**[q,s]
Harald Mischak,[c,d,l,q,vv,ww] and **Philippe Schmitt-Kopplin**[r]

Because of its availability, ease of collection, and correlation with physiology and pathology, urine is an attractive source for clinical proteomics/peptidomics. However, the lack of comparable data sets from large cohorts has greatly hindered the development of clinical proteomics. Here, we report the establishment of a reproducible, high resolution method for peptidome analysis of naturally occurring human urinary peptides and proteins, ranging from 800 to 17,000 Da, using samples from 3,600 individuals analyzed by capillary electrophoresis coupled to MS. All processed data were deposited in an Structured Query Language (SQL) database. This database currently contains 5,010 relevant unique urinary peptides that serve as a pool of potential classifiers for diagnosis and monitoring of various diseases. As an example, by using this source of information, we were able to define urinary peptide biomarkers for chronic kidney diseases, allowing diagnosis of these diseases with high accuracy. Application of the chronic kidney disease-specific biomarker set to an independent test cohort in the subsequent replication phase resulted in 85.5% sensitivity and 100% specificity. These results indicate the potential usefulness of capillary electrophoresis coupled to MS for clinical applications in the analysis of naturally occurring urinary peptides. *Molecular & Cellular Proteomics 9:2424–2437, 2010.*

From the Departments of [a]Chemistry and [i]Biomolecular Chemistry, University of Wisconsin, Madison, Wisconsin 53706, [c]Mosaiques Diagnostics GmbH, 30625 Hannover, Germany, [f]RD Néphrologie, 34090 Montpellier, France, [g]Ludwig Maximilians University Munich, 80539 Munich, Germany, Departments of [h]Clinical Immunology, [s]Hematology, Hemostasis, Oncology, and Stem Cell Transplantation, and [u]Nephrology and [n]Children's Hospital, Hannover Medical School, 30625 Hannover, Germany, [j]U858/I2MR, Renal and Cardiac Remodeling, INSERM, 31432 Toulouse, France, [k]Pediatric Nephrology Unit, Hopital des Enfants, Centre Hospitalier Universitaire de Toulouse, 31059 Toulouse, France, [l]British Heart Foundation Glasgow Cardiovascular Research Centre, Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8TA, UK, [o]Department of Nephrology and Clinical Immunology, RWTH University Hospital, 52074 Aachen, Germany, [p]Department of Internal Medicine IV, Saarland University Hospital, 66421 Homburg/Saar, Germany, [r]Institute of Ecological Chemistry, Helmholtz Center Munich-German Research Center for Environmental Health, 85764 Neuherberg, Germany, [t]Department of Computing Science, University of Glasgow, Glasgow G12 8QQ, United Kingdom, [v]Department of Medicine, Red Cross Hospital, 28199 Bremen, Germany, [w]Nephrology Department, Charité Hospital, 12200 Berlin, Germany, [y]Department of Psychiatry, University Hospital Hamburg-Eppendorf, 20246 Hamburg, Germany, [z]Department of Medicine, Austin Health and Northern Health, University of Melbourne, Heidelberg, Victoria 3081, Australia, [aa]Departments of Microbiology and Medicine, University of Alabama, Birmingham, Alabama

35294, [bb]Thermo Fisher Scientific, 28199 Bremen, Germany, [cc]Department of Nephrology, Transplantation Centre, Lower Saxony Centre for Nephrology, 34346 Hannoversch Münden, Germany, [dd]Systems Biology Ireland, University College Dublin, Dublin 4, United Kingdom, [ff]Section on Genetics and Epidemiology, Joslin Diabetes Center, Boston, Massachusetts 02215, [gg]Department of Medicine, Harvard Medical School, Boston, Massachusetts 02215, [hh]Department of Oncology Hematology and Respiratory Diseases, University of Modena and Reggio Emilia, 41100 Modena, Italy, [ii]Departments of Clinical Pharmacology and Nephrology, INSERM, ERI-12, and Amiens University Hospital, Université de Picardie Jules Verne, 80054 Amiens, France, [jj]Children's Hospital, University of Regensburg, 93053 Regensburg, Germany, [kk]Department of Analytical Chemistry, University of Aalen, 73430 Aalen, Germany, [ll]Department of Atherothrombosis and Vascular, Baker Heart Research Institute, Melbourne, Victoria 8008, Australia, [mm]Steno Diabetes Centre, 2820 Gentofte,

Chronic kidney disease (CKD)[1] is often characterized by a slow, progressive loss of renal function with a loss of glomerular filtration over a period of months or years that may eventually lead to end stage renal disease (ESRD). Patients with ESRD require renal replacement therapy (dialysis or kidney transplantation). The most common causes of CKD in North America, Europe, and Japan are diabetic nephropathy, hypertension, and glomerulonephritis (1). Together, these diseases account for ~75% of all adult cases of ESRD. Historically, kidney diseases were classified according to the anatomical compartment of the kidney that is involved. On this basis, vascular diseases include large and small vessel diseases, such as hypertensive nephropathy and vasculitis. Glomerular diseases comprise a diverse group of histologically defined primary glomerulopathies (*e.g.* focal segmental glomerulosclerosis (FSGS), membranous glomerulonephritis (MGN), minimal change disease (MCD), and IgA nephropathy (IgAN)) and secondary glomerulopathies due to diabetes mellitus (diabetic nephropathy), systemic autoimmune disorders (*e.g.* lupus erythematosus and hemolytic-uremic syndrome), or chronic viral infection (*e.g.* hepatitis and HIV). Tubular diseases are characterized by low molecular weight proteinuria and multiple transport defects (*e.g.* DeToni-Debré-Fanconi syndrome).

In clinical practice, renal damage is generally detected by proteinuria/albuminuria on urinalysis or quantitative measurement, changes in serum creatinine concentration for estimation of glomerular filtration rate, or both. However, these methods have major limitations as they are nonspecific and frequently are also late manifestations of renal damage. Therefore, we have sought to define alternative biomarkers for renal damage that may enable earlier and more accurate disease assessment.

Analysis of urine plays a central role in clinical diagnostics as it can be collected non-invasively. Urine as a body fluid for clinical analysis is relatively stable, probably due to the fact that it is "stored" for hours in the bladder; hence, proteolytic degradation by endogenous proteases may be essentially complete by the time of voiding (2). This is in sharp contrast to blood for which the activation of proteases and, consequently, generation of an array of proteolytic breakdown products are inevitably associated with its collection (3). The human urinary peptidome has been extensively investigated to gain insight about disease processes affecting the kidney and the urogenital tract (4–6). Urinary proteins and peptides originate not only from glomerular filtration but also from tubular secretion, epithelial cells shed from the kidney and urinary tract, secreted exosomes, and seminal secretions (7–9). Urine is a rich source of biomarkers for a wide range of diseases due to specific changes in its proteome (10–13). To test the feasibility of urinary proteomics as a non-invasive diagnostic tool, large scale studies are needed to analyze urine samples with reliable and quantitative experimental procedures. Various techniques have been applied to this effort, including two-dimensional electrophoresis combined with mass spectrometry (MS) and/or immunochemical identification of proteins (14–16), liquid chromatography coupled to mass spectrometry (LC-MS) (17, 18), and surface-enhanced laser desorption/ionization mass spectrometry (SELDI-MS) (19).

Because of mostly technical challenges, studies relying on proteomics experimental procedures are often restricted to the comparison of two groups of subjects (*i.e.* healthy controls *versus* patients with a well defined disease entity) with only a few individuals in each group. The lack of comparability severely limits the suitability of such data for a meta-analysis approach to define broadly applicable biomarkers. Consequently, findings from several studies cannot be used to explore the human urinary proteome/peptidome in its entirety. In addition, the health state of patients with kidney disease is often too heterogeneous to be reliably classified by biomarkers identified by such a strictly single disease-oriented approach. Diagnosis of individuals with different stages or types of kidney disease (disease controls) is conceivable by multiplex screening of proteomics data. Realization of such an approach critically depends on the use of a measurement platform allowing analysis of proteomic profiles within a reasonably short time and with high resolution and on the generation of a reference database for the human urinary proteome/peptidome.

Capillary electrophoresis coupled to mass spectrometry (CE-MS) enables reproducible and robust high resolution analysis of several thousand low molecular weight urinary proteins/peptides in less than an hour (5). In comparison with other proteomics methods, CE-MS offers several advantages. (i) It provides fast separation with high resolution. (ii) It is robust: capillaries are inexpensive and can be reconditioned efficiently using NaOH. (iii) It is compatible with most volatile buffers and analytes generally required for ESI. (iv) It provides a stable constant flow, avoiding the necessity of buffer gradients (for more details, see recent reviews (3, 5, 20)).

Denmark, [nn]Department of Nephrology, Clinical Centre Bayreuth, 95444 Bayreuth, Germany, [pp]Department of Urology, University of Leipzig, 04103 Leipzig, Germany, [rr]Department of Urology, University of Virginia, Charlottesville, Virginia 22908, [tt]Faculty of Medicine at Siriraj Hospital and Center for Research in Complex Systems Sciences, Mahidol University, Bangkok 10700, Thailand, and [uu]Department of Internal Medicine, University Hospital Ghent, 9000 Ghent, Belgium

[1] The abbreviations used are: CKD, chronic kidney disease; ANCA, anti-neutrophil cytoplasmic antibody; CAD, coronary artery disease; CE, capillary electrophoresis; ESRD, end stage renal disease; FSGS, focal segmental glomerulosclerosis; HC, healthy controls; IgAN, IgA nephropathy; MCD, minimal change disease; MGN, membranous glomerulonephritis; ROC, receiver operating characteristic; SVM, support vector machine; ID, identification; CI, confidence interval; FDR, false discovery rate; BPH, benign prostatic hyperplasia.

This approach has recently been used to analyze urine samples from healthy individuals and patients with various chronic kidney diseases in several independent masked studies (21–27), including IgAN (28), diabetic nephropathy (29), and ANCA-associated vasculitis (30). The high number of data sets analyzed under identical conditions using the same technological platform allows comprehensive characterization of the low molecular weight proteome (peptidome) that can then become a primary source of information for the diagnosis, classification, and monitoring of a wide range of diseases. Here, we report the analysis of the human urinary peptidome by CE-MS and the identification of peptide urinary biomarkers for the detection of pathological changes in the kidney during the development of many forms of CKD. Furthermore, we have replicated these findings in an independent cohort.

EXPERIMENTAL PROCEDURES

Sample Collection—Since 2004, urine samples were collected at >20 clinical centers (Europe, America, and Australia) using consistent standard operating procedures for sample collection, storage, and transport.[2] All samples were collected as a midstream portion of a spontaneously voided, second morning urine (with the exception of urine samples for the analysis of prostate cancer where an initial portion of a second morning urine was collected) and were stored immediately at −20 °C until analysis. The accepted diagnostic reference standard of the respective disease was used for the characterization of disease status (*e.g.* cystoscopy for the analysis of bladder cancer and renal biopsy for diagnosis of the type of CKD). All healthy controls (HC) were checked with a dipstick urinalysis for proteinuria. Clinical data for this study are shown in Table I. Furthermore, clinical parameters of the majority of individual samples have been published previously (25, 28, 30) and are available upon request. The study complied with the guidelines of the Declaration of Helsinki (www.wma.net/en/30publications/10policies/b3/index.html). Informed consent was obtained from all subjects, and ethical approval was obtained from the appropriate Institutional Review Boards.

Sample Preparation—A 0.7-ml aliquot of urine was thawed immediately prior to use and diluted with 0.7 ml of an aqueous solution supplemented with 2 M urea, 10 mM NH_4OH, and 0.02% SDS. To remove proteins of higher molecular mass, the sample was filtered using Centrisart ultrafiltration devices (20-kDa molecular mass cutoff; Sartorius, Goettingen, Germany) at $3,000 \times g$ for 45 min at 4 °C. Subsequently, 1.1 ml of filtrate was applied onto a PD-10 desalting column (GE Healthcare) equilibrated in 0.01% NH_4OH in HPLC grade water to remove urea, electrolytes, and salts. Finally, all samples were lyophilized, stored at 4 °C, and resuspended in HPLC grade water shortly before CE-MS analysis as described previously (27). The resuspension volume was adjusted to yield 0.8 $\mu g/\mu l$ protein as measured by BCA assay (Interchim, Montlucon, France).

CE-MS Analysis—CE-MS analysis was performed with a P/ACE™ MDQ capillary electrophoresis system (Beckman Coulter, Brea, CA) coupled on line to micro-TOF-MS instrument (Bruker Daltonics, Bremen, Germany) (27). The electroionization sprayer (Agilent Technologies, Santa Clara, CA) was grounded, and the ion spray interface potential was set between −4.0 and −4.5 kV. Data acquisition and MS acquisition experimental procedures were automatically controlled by the CE via contact closure. Spectra were accumulated every 3 s over a range of m/z 350–3,000.

The analytical characteristics of the CE-MS system have been extensively investigated by Kolch *et al.* (3) and Theodorescu *et al.* (27). Briefly, the average recovery of the sample preparation was ~85% with a detection limit of ~1 fmol. Monoisotopic mass signals were resolved for $z \leq 6$. Mass accuracy of the CE-TOF-MS method was determined to be <25 ppm for monoisotopic resolution and <100 ppm for unresolved peaks ($z > 6$). Mass accuracy of the CE-TOF-MS method was determined to be <25 ppm for monoisotopic resolution and <100 ppm for unresolved peaks ($z > 6$).

CE-FT-ICR-MS Analysis—For CE-FT-ICR-MS, a Bruker Daltonics Apex Qe instrument equipped with a 12-tesla magnet, and an Apollo II ion source was used. Coupling of the P/ACE™ 5510 capillary electrophoresis system (Beckman Coulter) via the Agilent ESI sprayer was performed as above. The instrument was tuned with a peptide standard mixture (27) and externally mass-calibrated on arginine clusters (<0.1-ppm calibration errors). Mass spectra were acquired over an m/z range of 300–2,000. Ions were stored in the collision cell for 500 ms, and five spectra were accumulated for each scan, resulting in a scan interval of 5 s.

A set of 20 samples was reanalyzed. The number of FT-ICR-detectable features was significantly lowered (by a factor of ~10). After normalization of the CE migration time, the analysis using CE-FT-ICR resulted in ~500 features, which can be matched to the CE-TOF-MS entities, which are in the urinary database. Eighty of these features were selected, enabling alignment with high confidence (because their sequence could be identified or they were found to be present at relatively high abundance) or providing a good coverage of features over the whole CE time and mass area. For mass calibration of the CE-TOF-MS data, the precise masses of these features detected by CE-FT-ICR-MS (see supplemental Table III) were used as references in a linear regression (see also Ref. 31).

Data Processing—Mass spectral ion peaks representing identical molecules at different charge states (m/z with z = 1, 2, 3,…) were deconvoluted into single masses using MosaiquesVisu software (32) (www.proteomiques.com). Only signals observed in a minimum of three consecutive spectra with a signal-to-noise ratio of at least 4 were considered. Signals with a calculated charge of 1+ were automatically excluded to minimize interference with matrix compounds or drugs. MosaiquesVisu used a probabilistic clustering algorithm and used both isotopic distributions and conjugated masses for charge state determination of the entities. TOF-MS data were calibrated utilizing Fourier transform-ion cyclotron resonance mass spectrometry (FT-ICR-MS) data as reference masses and applying linear regression. Both CE migration time and ion signal intensity (amplitude) showed high variability, mostly due to different amounts of salt and peptides in the sample, and were consequently normalized. Reference signals of over 1,700 urinary entities were used for CE time calibration by local regression. For normalization of analytical and urine dilution variances, MS signal intensities were normalized relative to 29 "housekeeping" peptides with small relative standard. For calibration, linear regression was performed (27, 33). The resulting peak list characterized each feature by its molecular mass (Da) and normalized CE migration time (min). Normalized signal intensity was used as a measure for relative abundance.

Data sets were accepted only if the following quality control criteria were met. A minimum of 950 chromatographic features (mean number of features minus one standard deviation) must be detected with a minimal MS resolution of 8,000 (required resolution to resolve ion signals with z = 6) in a minimal migration time interval (the time window in which separated signals can be detected) of 10 min. After calibration, the mean deviation of migration time (compared with reference standards) must be below 0.35 min.

Urinary Peptidome Database—All detected features in the urine samples that passed the quality control criteria (on average, 1,724

features were detected in each urine sample, ranging from 983 to 4,094) were deposited in a Microsoft Structured Query Language (SQL) database and subsequently matched for further analysis and comparison of individual samples. For clustering, features in different samples were considered identical if the mass deviation was lower than ± 50 ppm for 800-Da entities, gradually increasing to ± 75 ppm for 15-kDa features. Due to analyte diffusion effects, CE peak widths increase with CE migration time. In the data clustering process, this effect was considered by linearly increasing cluster widths over the entire electropherogram (19–45 min) from 2 to 5%. A feature that can be detected in a specific cluster will be assigned to a respective protein ID with its amplitude value. Feature signals that could not be related to a specific cluster possess a value of 0. After calibration, deviation of migration time was controlled to be below 0.35 min. This process resulted in the tentative definition of 116,869 different features. Each feature (presumably peptide) was assigned a unique identification number (protein ID). As described previously (25, 34), several of these features appeared sporadically, being observed in only one or a few samples. To eliminate such entities of apparently low significance from the analysis, only those features detected in more than 20% of the urine samples in at least one group (samples from patients with the same disease) were further investigated. This noise-filtering process reduced the number of features available for analysis significantly. Applying these limits, 5,010 "relevant" different entities characterized by molecular mass (Da) and normalized CE migration time (min) were detected. The filtered data of all individual samples are available.[3]

Statistical Analysis—Estimates of sensitivity and specificity were calculated based on tabulating the number of correctly classified samples. Confidence intervals (95% CI) were based on binomial calculations performed with MedCalc version 8.1.1.0 (MedCalc Software, Mariakerke, Belgium; www.medcalc.be). The receiver operating characteristic (ROC) plot was evaluated as it provides a single measure of overall accuracy that is not dependent upon a particular threshold (35).

For statistical differential analysis, we set a frequency threshold of 30% for a feature to be deemed valid in one of the considered groups and to be included in downstream analysis. For each of the entities, q values were estimated from the permutation method implemented in the Cran package called "samr" (36) (available at: http://cran.at.r-project.org/web/packages/index.html). We retained all peptides with a q value <0.05 from the samr output for downstream analysis. The q values were adjusted p values found using an optimized FDR approach. The FDR approach is optimized by using characteristics of the p value distribution to produce a list of q values. A p value of 0.05 implies that 5% of all tests will result in false positives. An FDR-adjusted p value (a q value) of 0.05 implies that 5% of significant tests will result in false positives. The latter is clearly a far smaller quantity. To test the results of the samr analysis, we calculated a p value for each feature using the standard Wilcoxon rank sum test. From those p values, the q values were obtained using the Cran package "fdrtool" (available at http://cran.at.r-project.org/web/packages/index.html) and the Bioconductor package "fdrame" (37, 38) (available at http://www.bioconductor.org/packages/release/Software.html). The agreement between the lists of differentially expressed markers (*i.e.* q value <0.05) using the different packages was higher than 94%. Supervised classification was performed on the training data using Mosa-Cluster software (Biomosaiques Software GmbH) for support vector machines (SVM) classification (23).

Classification—MosaCluster (version 1.6.5) was developed for the discrimination between different patient groups. This software tool allows the classification of samples in the high dimensional parameter space by using SVM. For this purpose, MosaCluster generates polypeptide models, which rely on peptides displaying statistically significant differences when comparing data from patients with a specific disease with those from healthy controls or patients with other diseases. Each of these peptides allegorizes one dimension in the *n*-dimensional parameter space (25, 39–41).

CE-MS Platform Validation—For statistical evaluation, F-test for comparison of standard deviations was used. For statistical testing, the classification score of the CKD classification model was used, reflecting the overall variability of the system (detection, deconvolution, clustering, and classification). Relative standard deviations were calculated with the entire observed classification range of the CKD model (lowest observed values >-1.4 and highest observed values $<+1.6$). Hence, a standard deviation of S.D. = 3 corresponds to 100% (statistical spread).

To assess intra- and intersample variations, reproducibility was based on the analysis of three individual urine samples (healthy controls) that were processed in 15 replicates. Therefore, each urine sample was prepared by one operator 15 times, and the 15 replicates were measured with the same device on 1 day. Over the next days, the second and the third batches of 15 replicates were processed as independent sets of experiments.

After assessment of variances over a short period of time, for one of the three urine samples (see reproducibility analysis below), the time point of analysis and operator/device dependences were assessed (intermediate precision). Therefore, at 15 different days the urine sample was prepared twice by two different operators; one aliquot was measured at device 1, and the second aliquot was measured at device 2.

To determine the statistical variances observed for CKD diagnosis of individuals over the time course of a month, classification results were assessed for four healthy volunteers (two males and two females). For each individual, second morning urine samples were taken on Wednesday (reflecting working activities) and on Sunday (reflecting weekend activities) for 4 weeks. The duration of sampling reflects variances according to work-life balance and according to menstrual cycle for females. All urine samples were immediately stored at -20 °C after sampling for at least 24 h until time of preparation. All samples were prepared together as a single batch by one operator and were measured using the same device over a short period of time.

Sequencing of Peptides—Candidate biomarkers and other native peptides from urine were sequenced using CE-MS/MS or LC-MS/MS analysis from different sequencing platforms (42). Sequencing using MALDI-TOF-TOF was performed on an Ultraflex instrument (Bruker Daltonics). To retain the information on migration time, the entire CE run (see "CE-MS Analysis") was spotted onto a MALDI target plate using a Probot microfraction collector (LC Packings, San Francisco, CA), depositing one spot every 15 s. Matrix solution (2 mg/ml α-cyano-4-hydroxycinnamic acid in 50% acetonitrile and 0.1% TFA) was added to the CE eluate during application onto the plate with a flow rate of 4 μl/min. The plate was initially examined in MS mode. Subsequently, peptides were fragmented using the tandem mass spectrometry (MS/MS) mode. MALDI-TOF-TOF-MS/MS was carried out using MALDI postsource decay (PSD) in combination with LIFT-cell TOF/TOF setup of the Ultraflex mass spectrometer.

For sequencing using ESI-Q-TOF, a Qq-orthogonal time-of-flight mass spectrometer (micrOTOF-Q, Bruker Daltonics) in MS and MS/MS modes directly coupled to a CE system via a coaxial sheath liquid sprayer (Agilent Technologies) was used for the experiments. The CE conditions were adapted to the conditions described under "CE-MS Analysis" to have comparable migration time of the peptides. The MS analysis was performed in positive electrospray mode with an ESI-TOF sprayer kit from Agilent Technologies (Palo Alto, CA). The ESI sprayer was grounded, and the sheath liquid, consisting of 30% (v/v) isopropanol (Sigma-Aldrich Chemie GmbH) and 0.5% (v/v) for-

---

[3] http://mosaiques-diagnostics.de/diapatpcms/mosaiquescms/front_content.php?idcat=257/; version 2.0; password: Mosa573X.

mic acid in HPLC grade water, was applied coaxially with a rate of 2 μl/min. The MS system operates in data-dependent MS/MS mode with one full-scan MS spectrum (50–3,000 m/z) followed by three MS/MS spectra. The repetition rate of the TOF was set to 1.5 Hz in the full-scan mode and 1 Hz in the MS/MS mode. The instrument was operated at 15,000 resolution. Argon was used as collision gas, and the collision energies were set automatically, depending on the m/z value and the charge state of the peptide. All data were externally calibrated using a sodium formate cluster. The calibration was used for both MS modes.

For sequencing using ESI ion trap, an aliquot corresponding to ~25–50 μg (2.5–5 μl) of a sample was loaded onto a preanalytical column that consisted of a 10-cm-long piece of 360-μm-outer diameter × 100-μm-inner diameter fused silica (Polymicro Technologies, Phoenix, AZ) packed with C_{18} particles (5–15-μm YMC ODS-A, Waters, Milford, MA). After sample loading, the column was washed extensively with 0.1% acetic acid and then butt-connected using a Teflon sleeve to an analytical column, which consisted of a 10-cm-long piece of 360-μm-outer diameter × 50-μm-inner diameter fused silica with an integrated nanospray tip (1–2-μm orifice) and packed with C_{18} resin (5-μm YMC ODS-AQ). Peptides were eluted using high performance liquid chromatography (Agilent 1100, Agilent Technologies, Palo Alto, CA) with a flow rate of ~100 nl/min using a splitter. Peptides were eluted directly into an ESI interface and analyzed with a Finnigan LTQ linear quadrupole ion trap mass spectrometer (Thermo Electron Corp., San Jose, CA). The LTQ was operated with a "top 10" data-dependent analysis method consisting of a repeated data acquisition cycle of one full mass spectrum (m/z 400–2000) followed by 10 MS/MS spectra, which corresponded to fragmentation mass spectra of the top 10 most abundant ions from the MS.

MS/MS experiments were also performed on an Ultimate 3000 nanoflow system (Dionex/LC Packings, Sunnyvale, CA) connected to an LTQ Orbitrap hybrid mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nanoelectrospray ion source. The mass spectrometer was operated in data-dependent mode to automatically switch between MS and MS/MS acquisition. Survey full-scan MS spectra (from m/z 300 to 2,000) were acquired in the Orbitrap. Ions were sequentially isolated for fragmentation in the linear ion trap using collision-induced dissociation. General mass spectrometric conditions were as follows: electrospray voltage, 1.6 kV; no sheath and auxiliary gas flow; ion transfer tube temperature, 225 °C; collision gas pressure, 1.3 millitorrs; normalized collision energy, 32% for MS/MS. Ion selection threshold was 500 counts for MS/MS. In addition, samples were analyzed using electron transfer dissociation (43–45). Peptides were separated by nano-reversed phase HPLC (Agilent 1100; flow split by a tee to ~60 nl/min) and introduced into an electron transfer dissociation-capable LTQ quadrupole linear ion trap (Thermo Fisher Scientific, San Jose, CA) via nano-ESI using instrumental parameters described previously (46).

Raw data files were either converted into dta files (RAW files generated by ion traps from Thermo Fisher Scientific) with the use of DTA Generator (47, 48) or into mgf files (data derived from MALDI-TOF and Q-TOF analyses) with the use of DataAnalysis (version 4.0; Bruker Daltonics). All resultant MS/MS data were submitted to MASCOT (www.matrixscience.com; release number, 2.3.01) for a search against human entries (20,295 sequences) in the Swiss-Prot database (Swiss-Prot Number 2010.06) without any enzyme specificity and with up to one missed cleavage. No fixed modification was selected, and oxidation products of methionine, proline, and lysine residues were set as variable modifications. The accepted parent ion mass deviation was 0.5 Da (20 ppm for all Orbitrap spectra); the accepted fragment ion mass deviation was 0.7 Da. Only search results with a MASCOT peptide score equal to or higher than the MASCOT score threshold were included (see supplemental Table I).

Additionally, ion coverage was controlled to be related to main spectral fragment features (b/y or c/z ion series) (see also the supplemental figures). For further validation of obtained peptide identifications, the strict correlation between peptide charge at pH 2 and CE migration time was utilized to minimize false positive identification rates (42, 49). As depicted in Fig. 1, the polypeptides are arranged in four to five lines. The members of each line are characterized by the numbers of basic amino acids (arginine, histidine, and lysine) included in the peptide sequence. Specifically, the peptides in the *right* line contain no basic amino acids; only the N terminus of the peptide is positively charged at pH 2. In contrast, peptides in the other lines (from *right* to *left*) show increasing numbers of basic amino acids in addition to their N-terminal ammonium group (49). The calculated CE migration time of the sequence candidate based on its peptide sequence (number of basic amino acids) was compared with the experimental migration time. A peptide was accepted only if it had a mass deviation below ±25 ppm and a CE migration time deviation less than ±2 min. The full list of identified peptides and MS/MS spectra is available at Human ProteinPedia (http://www.humanproteinpedia.org; accession number HuPA_00670).

RESULTS

All urine samples were prepared using a standard protocol and analyzed by CE-MS, resulting in individual data sets containing information on generally 1,200–2,000 features (presumably peptides) per sample. All information recommended by the "minimum information about proteomics experiments" guidelines (50) for proteome analysis was recorded and is available upon request. The resulting list of peptides was defined by mass, migration time, and ion counts, the last of which served as a measure of the relative abundance of a peptide.[3]

To improve mass accuracy, 20 samples were reanalyzed using CE on line-coupled to FT-ICR-MS. Because of the high costs and the lower sensitivity (higher detection limit) of FT-ICR-MS instruments in comparison with TOF instruments, it was not practical to analyze all samples by CE-FT-ICR-MS. The high FT-ICR-MS resolution also enabled an accurate analysis of the first isotope signal of ions with $z > 6$, which is crucial for determination of the exact mass of proteins and high molecular weight peptides. These data were used to refine the TOF-MS masses of the urinary peptides.

During additional calibration steps using "internal standard" peptides (27, 33) for reference, CE migration time and signal intensity were corrected for analytical variances and urine dilution effects. All calibrated data are publicly accessible on the Internet.[3] A subsequent, thorough data calibration (see "Data Processing") allowed digital compilation (matching of data from respective samples) of individual data sets to an averaged "case-specific" data set. This compiled data set can be compared with any compiled "control group" data set, enabling the identification of statistically significant changes for biomarker definition.

The main purpose of this "human urinary peptidome database" is to serve as a universal platform for definition and validation of biomarkers for a variety of diseases and (patho)-physiological changes. For the selection of CKD-specific bi-

TABLE I
Demographic data of all patients used in this study

The first part shows the clinical data of patients with different diseases (see column "Group"), which represents the training set. The total number (n) of patients, gender ratio (male/female), age, serum (S) creatinine, and proteinuria status are shown. All patients included in the test set are listed in the second part. Patients representing further controls are listed in the last part of the table where blood pressure values for the hypertensive patients are given in the last column. ND, not done; DN, diabetic nephropathy; SLE, systemic lupus erythematosus; SBP, systolic blood pressure; DBP, diastolic blood pressure.

Group	n	Gender (male/female)	Age (mean ± S.D.)	S creatinine (mean ± S.D.)	Proteinuria (mean ± S.D.)	Blood pressure (SBP/DBP)
			years	μmol/liter	g/d	mm Hg
Training set						
Healthy controls	379	196/183	33 ± 16	ND	<0.1[a]	
ANCA	30	15/15	61 ± 7	179 ± 85	0.89 ± 0.53	
SLE	21	9/12	39 ± 12	109 ± 56	1.10 ± 2.24	
MGN	30	24/6	40 ± 23	97 ± 23	1.27 ± 1.00	
MCD	22	10/12	51 ± 18	113 ± 73	0.32 ± 0.44	
IgAN	44	34/10	43 ± 14	163 ± 96	1.88 ± 2.42	
FSGS	25	17/8	46 ± 26	126 ± 49	1.64 ± 2.53	
DN	58	46/12	53 ± 9	133 ± 82	1.71 ± 2.25	
Test set						
Healthy controls	34	21/13	54 ± 13	ND	<0.1[a]	
FSGS	12	9/3	59 ± 14	130 ± 22	0.82 ± 1.40	
MGN	15	9/5	57 ± 13	100 ± 14	5.75 ± 2.58	
SLE	9	4/5	42 ± 17	170 ± 104	1.40 ± 1.74	
ANCA	23	11/12	69 ± 4	230 ± 152	0.80 ± 0.83	
IgAN	18	12/6	46 ± 18	156 ± 112	1.81 ± 2.12	
MCD	4	2/2	51 ± 11	182 ± 84	1.06 ± 2.34	
DN	10	7/3	65 ± 13	86 ± 24	4.22 ± 2.79	
Other CKD	18	10/8	56 ± 16	132 ± 81	1.73 ± 2.10	
Additional controls						
HIV	29	22/7	44 ± 3	65 ± 3	ND	ND
Diabetes mellitus II	22	9/13	66 ± 7	82 ± 50	16.3 ± 20.0[b]	ND
Diabetes mellitus I	36	16/20	53 ± 7	ND	ND	ND
BPH	34	34/0	64 ± 8	ND	ND	ND
Hypertension	13	10/3	61 ± 9	ND	0.25 ± 0.19	147 ± 13/80 ± 10
MCD, remission	3	2/1	20 ± 2	ND	0.09 ± 0.05	ND

[a] Tested with dipstick method.
[b] Urinary albumin.

omarkers, data from individual samples were compiled (21, 27) and grouped according to patients' clinical profiles (diagnostic group). These groups included healthy subjects (n = 379) and patients with various biopsy-proven kidney diseases (n = 230) (for details, see Tables I and II). All peptides were statistically analyzed using multiple testing correction (51). This approach resulted in 634 peptides of statistical significance. To reduce this large number of variables, we chose only a subset of these peptides with known sequence information. This strategy resulted in the identification of a panel of 273 peptides listed in supplemental Table I. Subsequently, an SVM classification model was generated based on these peptides to distinguish between healthy subjects and individuals with biopsy-proven kidney diseases (Fig. 1). In the training set, a sensitivity of 98.7% and a specificity of 100% were observed (for ROC curve analysis, see also Fig. 2, "Training Set"). To avoid any bias introduced by the use of samples from only one clinical center, samples from patients with different diseases or (patho)physiological conditions collected at different centers were included in this training set.

In agreement with the recently published guidelines for clinical proteomics (52), we reproduced this biomarker pattern in a multicenter prospective study using an independent blinded cohort (test set) of 144 individuals, including patients with different kidney diseases (n = 110) and controls (n = 34). Upon unblinding, all controls and 94 patients with CKD were correctly classified, resulting in a sensitivity of 85.5% (95% CI, 77.5–91.4) and a specificity of 100% (95% CI, 89.6–100.0) (see Fig. 2, "Test Set," and supplemental Table IIa, "prospective study").

CKD is frequently associated with underlying pathologies due to complications of diabetes, hypertension, or chronic viral infection. Therefore, we examined the specificity of the identified markers using further "disease controls" present in the human urinary peptidome database. This test set consisted of patients with diabetes mellitus types I and II, untreated HIV patients, patients with benign prostatic hyperplasia (BPH), and patients with arterial hypertension. Furthermore, we included in this test set additional patients with prior MCD who now have complete remission induced by treatment

TABLE II

Compendium of all patients and healthy controls who were used for establishment of training and test sets

Additionally, all sensitivities and specificities are shown. Clinical details on the patients and controls utilized in training and test sets are shown in Table I. DN, diabetic nephropathy; SLE, systemic lupus erythematosus.

Disease condition: CKD (cutoff, 0.343)							
Training set				Test set			
Case	Control	Sensitivity	Specificity	Case	Control	Sensitivity	Specificity
		%	%			%	%
n = 230; 30 ANCA, 21 SLE, 30 MGN, 22 MCD, 44 IgAN, 25 FSGS, 58 DN	n = 379; 379 HC	98.7	100.0	n = 110; 12 FSGS, 15 MGN, 9 SLE, 23 ANCA, 18 IgAN, 5 MCD, 10 DN, 18 other CKD	n = 34; 34 HC	85.5	100.0

with glucocorticoids. All these patients showed no evidence of renal damage based on clinical history, serum creatinine, or urinary protein levels (see Table I). The results of the classification with the CKD model are shown in supplemental Table IIb, "additional controls." None of the 29 patients with HIV, one of the 22 patients with diabetes mellitus type II, one of the 36 patients with diabetes mellitus type I, none of the 34 patients with BPH, one of the 13 patients with arterial hypertension, and none of the five patients with MCD in remission were classified as having CKD, resulting in an overall specificity of 97.8%.

To precisely gauge the analytical variability of the CE-MS-based CKD platform, a set of experiments for platform validation was performed: temperature stability, postpreparation stability, frost/defrost stability, reproducibility, intermediate precision, and time course were evaluated. In all tests, urine samples were processed as described under "Experimental Procedures."

The aim of the stability analyses was the assessment of sample stability at room temperature, at 4 °C, and after repeated freeze/thaw cycles. Additionally, postpreparation stability of the prepared sample for 24 h at 4 °C in an autosampler was investigated. Both prolonged storage (at room temperature for 6 h and at 4 °C for 24 h) and repeated freeze/thaw cycles did not yield significant differences in standard deviations (see Fig. 3A). Furthermore, all analyses resulted in SVM scores. This finding implied a consistent classification result for these patients in all experiments. All observed statistical spreads were below 11%. Variances observed for postpreparation stability were in the range of 6%.

Reproducibility and intermediate precision analyses were used to determine the statistical variances of the SVM classification result 1) under the same operating conditions over a short period of time (reproducibility or intra-assay precision) and 2) at different days by different operators with different devices in a long period of time (intermediate precision or within-laboratory precision). As depicted in Fig. 3B, the relative intra-assay standard deviation was below 7% for all tested samples. As expected, relative interassay standard deviation increased ($p = 0.005$); however, it was still below 10%. Intermediate precision included different lot numbers of buffers, solvents, and chemicals. Both devices underwent biannual maintenance service.

Statistical variances were determined for time course experiments (see Fig. 3C) observed for CKD diagnosis of individuals over the course of a month. The variances were in the range of the intermediate precision, suggesting stable expression of the CKD pattern and, thus, a small biological variance over time.

We aimed to identify as many peptides of the human urinary database as possible by applying state-of-the-art top-down MS/MS. Sequence information was obtained for 444 of the 5,010 peptides constituting the human urinary peptidome database.[3] As described previously (42, 49), the migration

FIG. 1. **Peptide patterns distinguishing patients with CKD from HC.** The compiled data sets of urine samples from 230 patients with CKD (*left panel*) and 379 healthy control subjects (*right panel*) included in the training set are shown. Normalized molecular mass (*y axis*) is plotted against normalized CE migration time (*x axis*). The mean signal intensity is represented in three-dimensional depiction.
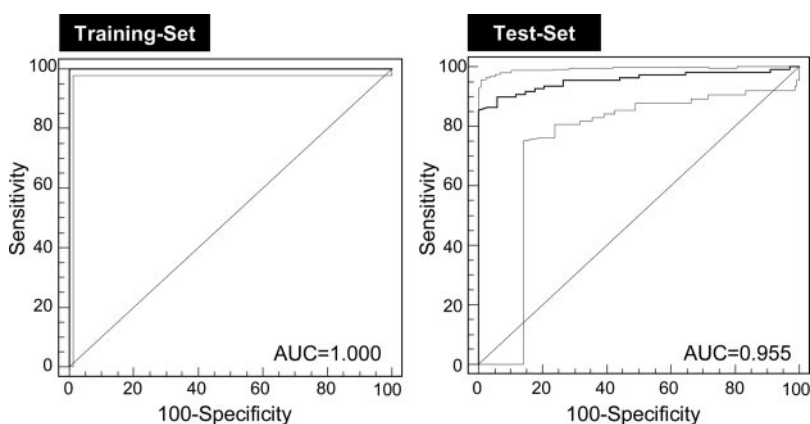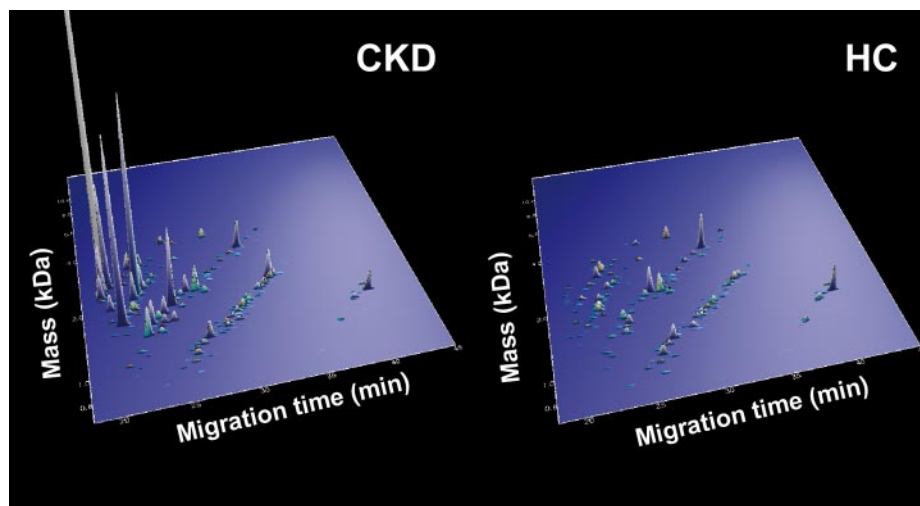

FIG. 2. **ROC curves for classification of patient cohorts with "CKD pattern."** ROC analysis for CKD diagnosis of the training set and the test set after unblinding is shown. *AUC*, area under the curve.

time in CE depends on peptide mass and charge state at pH 2, equaling the number of free amino groups (N-terminal and basic amino acids). Therefore, it is not a prerequisite to use CE separation for MS/MS sequencing as the number of basic amino acids as well as the calculated mass serve to assign a putative sequence to a signal in the CE-MS spectrum. The identification and annotation are associated with a level of uncertainty, but this level appeared to be very low as all data generated by CE-MS/MS were confirmed by LC-MS/MS identification and, thus, confirmed the subsequent annotation of CE-MS data. Furthermore, of the 273 potential CKD biomarkers, 107 were also among the features detected in the FT-ICR data (see supplemental Table I, last two columns), which show mass differences compared with the calculated masses of the sequences of 0.5 ppm ($\pm$0.9 ppm).

The sequenced markers for the diagnosis of CKD were fragments of different collagens, blood proteins (*e.g.* α_1-antitrypsin, serum albumin, hemoglobin α chain, and fibrinogen α chain), and kidney-specific proteins (*e.g.* uromodulin, sodium/potassium-transporting ATPase γ chain, and membrane-associated progesterone receptor component 1) as well as fragments of various secreted proteins (see Table III and supplemental Table I). As depicted in Fig. 1 and supplemental Table I, fragments from serum proteins increased, whereas those of most of the collagens decreased in CKD.

DISCUSSION

Here, we report the application of CE-MS for the analysis of the human urinary peptidome and demonstrate its utility for the definition of biomarkers of CKD in general. These biomarkers apparently reflect primary pathogenetic changes as well as the response to certain disease processes. Hence, their usefulness extends beyond the applicability to diseases of the urogenital tract, and the approach may be applicable to diseases that result in systemic manifestations. Although genetic analysis can predict the risk of a disease, proteomics, with its potential to monitor dynamic processes, may more clearly show at which point in time the risk manifests itself as disease and also facilitates monitoring of the response to therapy. Thus, these methods are complementary in the therapeutic approach to an individual patient (29, 53).

Recently, several groups have reported sequencing of an array of urinary proteins (15, 54). Although these data sets impressively encompass many proteins (and the potential for insightful information), clinically applicable information, *i.e.* for
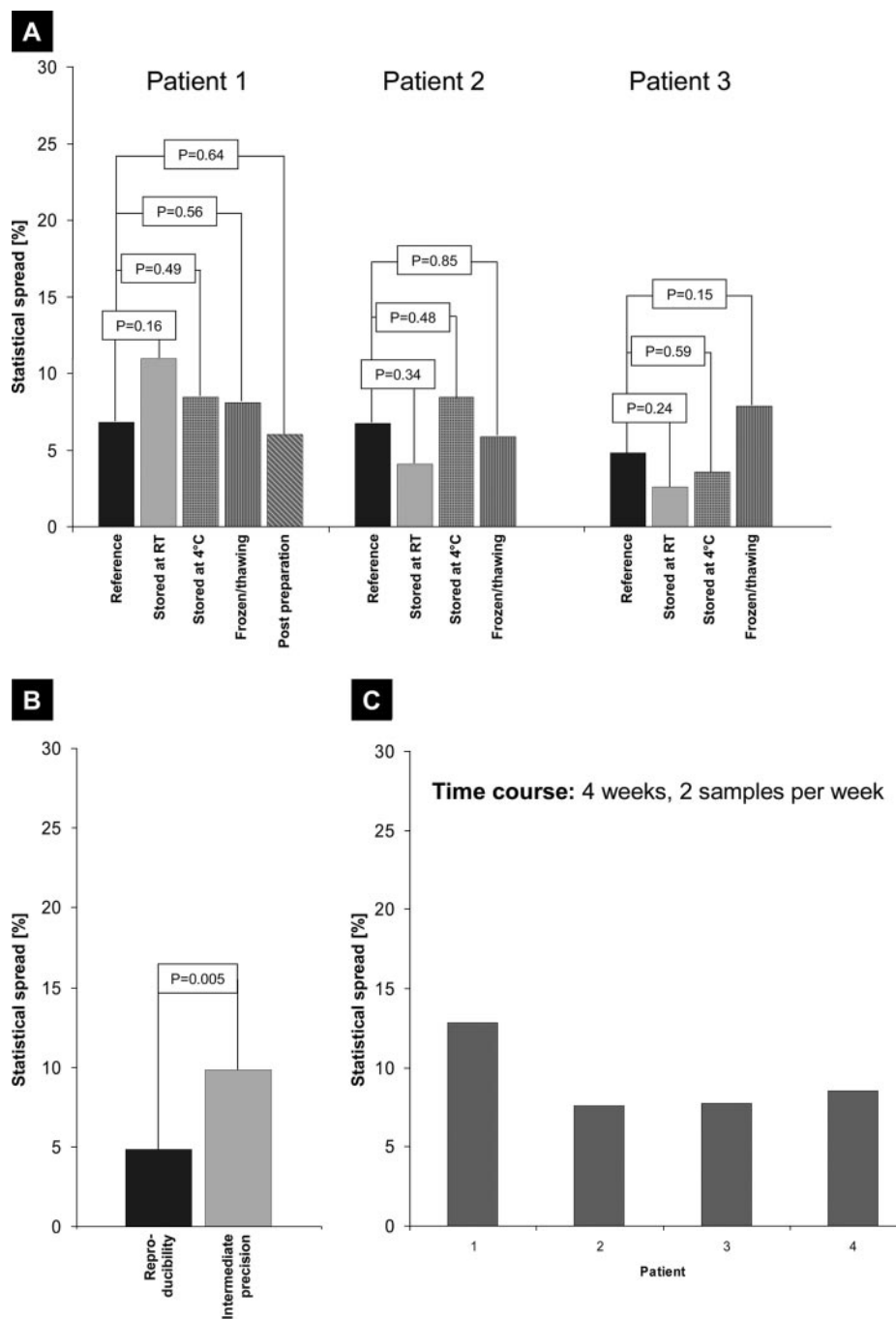
FIG. 3. **Analytical variability of CE-MS-based CKD platform.** *A*, statistical spread of temperature stability, freeze/thaw stability, and postpreparation stability are shown. Reproducibility and intermediate precision (*B*) and time course (*C*) are plotted against their statistical spread. *p* values are depicted in *boxes*.

the definition of biomarkers, is completely missing. Moreover, all of these studies used tryptic digests of urinary proteins. The sequences from these peptides allowed the investigators to tentatively assign a particular protein to the obtained sequence with variable confidence. Unfortunately, because of the *in vitro* manipulation of samples by digestion, it is not possible to define which of these peptides are natural constituents of human urine. The information that is essentially required in clinical proteomics is restricted to only naturally occurring urinary peptides. Furthermore, in cases where bi-

omarkers must be defined with a high level of confidence, information on their relative abundance is critically required.

We have therefore attempted to obtain such substantial information: the naturally occurring peptides are defined by molecular mass and migration time, and the relative abundance is defined by ion counting, which is normalized by the use of "internal standards" (27, 33). The latter resemble peptides that are present in almost every sample and that do not change significantly in their ion counts in all samples and disease groups investigated to date. Although this approach

TABLE III

Source proteins and peptide distribution of CKD biomarkers

Protein	Swiss-Prot name	Number of fragments
Collagen α-1 (I) chain	CO1A1_HUMAN	126
Collagen α-1 (III) chain	CO3A1_HUMAN	55
α_1-Antitrypsin	A1AT_HUMAN	18
Collagen α-2 (I) chain	CO1A2_HUMAN	15
Uromodulin	UROM_HUMAN	11
Serum albumin	ALBU_HUMAN	9
Fibrinogen α chain	FIBA_HUMAN	5
Polymeric immunoglobulin receptor	PIGR_HUMAN	4
α_2-HS-glycoprotein	FETUA_HUMAN	3
Clusterin	CLUS_HUMAN	2
Collagen α-1 (II) chain	CO2A1_HUMAN	2
Membrane-associated progesterone receptor component 1	PGRC1_HUMAN	2
Osteopontin	OSTP_HUMAN	2
Sodium/potassium-transporting ATPase γ chain	ATNG_HUMAN	2
Transthyretin	TTHY_HUMAN	2
α_{1B}-Glycoprotein	A1BG_HUMAN	1
Antithrombin-III	ANT3_HUMAN	1
Apolipoprotein A-I	APOA1_HUMAN	1
β_2-Microglobulin	B2MG_HUMAN	1
CD99 antigen	CD99_HUMAN	1
Collagen α-1 (V)	CO5A1_HUMAN	1
Collagen α-1 (XVII) chain	COHA1_HUMAN	1
Collagen α-1 (XVIII) chain	COIA1_HUMAN	1
Collagen α-2 (VIII) chain	CO8A2_HUMAN	1
Cystatin-B	CYTB_HUMAN	1
Ig λ chain C regions	LAC_HUMAN	1
Neurosecretory protein VGF	VGF_HUMAN	1
Pro-SAAS	PCSK1_HUMAN	1
Prostaglandin-H$_2$ D-isomerase	PTGDS_HUMAN	1
Psoriasis susceptibility 1 candidate gene 2 protein	PS1C2_HUMAN	1

does not allow unambiguous identification that can be realized only by amino acid sequence analysis, it does permit a tentative identification based on the exact mass and migration time. Sequencing was performed in a second step but was hindered by several obstacles associated with MS sequencing of naturally occurring peptides (tryptic digests cannot be utilized because of a loss of connectivity to the original identification parameters (55)). Major obstacles are suboptimal use of proteomics search machines (like MASCOT or the open mass spectrometry search algorithm) for naturally occurring peptides (5, 34) as well as the chemical nature of the peptides that prevents successful sequencing (42). For an assignment of the sequences to the correct label (protein ID in our database), there is a small chance of error. Assignment of two different peptides to one label and assignment of one peptide to two labels are likely to occur. At the same time and based on our MS/MS data, the results indicate that not every peptide that may theoretically be found was in fact detected. The pool of identical peptides present in urine is generally limited, reducing the challenge of correct assignment (because of a limited number of features present). Furthermore, errors in assignment will result in higher variability; thus, such peptides will likely be excluded as potential biomarkers.

CE-MS analysis of urine enables a tentative identification of biomarkers for a variety of diseases of the kidney and the urogenital tract (21, 22, 24, 25, 27, 56), although the high biological variability of peptides represents a serious methodological impediment. Therefore, it appears imperative to evaluate clinical disease conditions not on the basis of single peptide markers but rather on the basis of a biomarker set consisting of distinct and clearly defined discriminating molecules. A panel of biomarkers will tolerate changes in individual analytes without jeopardizing the diagnostic precision, which means that variability will not result in gross changes of the diagnostic result.

In comparison with LC-MS, CE-MS seems to have several advantages when analyzing urine samples. Of high relevance are the reproducibility and insensitivity toward interfering compounds and the fact that no "flow-through fraction" is generated. Small peptides that are highly charged generally do not bind to the typically used reversed phase columns and, thus, will be lost for LC-MS analysis. However, these small peptides would be detected by CE-MS analysis. Large molecules ($>$5 kDa) frequently do not elute efficiently off the LC column, and their relative abundance thus cannot be assessed with sufficient accuracy. These molecules are, however, easily and reproducibly detectable by CE-MS. A further advantage is the relative insensitivity of CE-MS to precipitates

that interfere substantially with LC separation. Furthermore, such precipitates can be removed effectively using NaOH, which is applied routinely after each CE-MS analysis (enabling reproducible reconditioning). A similar approach is impossible in LC-MS. These and additional advantages are outlined in more detail in several recent reviews (3, 34, 57). Furthermore, CE-MS analysis could be viewed as a very reproducible proteomics approach for the diagnosis of CKD. The analysis of samples stored under different conditions showed no increased variance, indicating that these urinary peptides are stable biomarkers, which may be also an advantage of using multiple biomarkers (biomarker pattern) instead of a single biomarker. Variances observed for postpreparation stability suggested sufficient resistance of the CKD biomarkers to oxidation, postpreparation precipitation, and degradation. In summary, these data underline the analytical robustness of the CE-MS-based diagnostic platform and represent the foundation for its routine clinical application for the diagnosis of CKD as already demonstrated for other indications, such as graft-*versus*-host disease (26).

We have shown that several peptides were differentially excreted in the urine of patients with different forms of CKD compared with healthy individuals (22, 25). These studies were pilot studies aiming toward the differential diagnosis of certain types of CKD and were compromised by low statistical power for marker definition because of the small numbers of patients and relatively poor mass resolution. Furthermore, the frequently observed high content of proteins in samples from patients with CKD resulted in precipitation during separation even in the case of CE-MS, thereby limiting accuracy and rendering comparison of multiple data sets challenging or even impossible. Therefore, the protocol for sample preparation has been optimized by the removal of proteins above 25 kDa without significant loss of low molecular mass urinary components (23). This improved method, consisting of an additional ultrafiltration step in the presence of SDS and urea to prevent the interaction of proteins and peptides (23), was used for all samples reported here and resulted in a higher consistency of obtained data sets compared with those from the pilot studies mentioned above. Nevertheless, these changes in preanalytical handling clearly disqualify old data sets for storage in the urinary database and proteomics comparison on the highest level for data consistency.

Recently, using the new sample preparation protocol, we showed that urinary biomarkers enable differential diagnosis of specific single chronic renal diseases (IgA nephropathy, diabetic nephropathy, and ANCA-associated vasculitis) with sufficient sensitivity and specificity in blinded data sets (28–30). Here, we present the applications of the urinary peptidome database for the non-invasive diagnosis of CKD in general.

To assess the value of an optimized human urinary peptidome database, we utilized the data for the diagnosis of CKD as a representative example of urogenital tract diseases. Using this peptidome database, 273 biomarkers were defined. The complexity of diseases and the pathological processes involved suggest that the concept of a single biomarker indicating not only a reliable diagnosis but also a stage in the pathological process and prognosis appears questionable. Combining multiple independent biomarkers into a diagnostic or predictive pattern (Fig. 1) may better address this problem. Many algorithms utilizing the available information on multiple biomarkers have emerged (39) with hierarchical decision tree-based classification methods (58), support vector machines (59), and Gaussian processes (60) among them (for more detailed information, see Ref. 39).

We hypothesized that a combination of the 273 CKD-specific biomarkers in a CKD-specific biomarker pattern would more accurately discriminate CKD patients from unaffected individuals. Although some of the classifiers of these biomarkers might appear as unnecessary to increase their specificity and sensitivity, an analysis investigating different algorithms for identification of biomarkers and establishment of classifiers demonstrated that an increase in the number of biomarkers produced a more robust model.[4] Therefore, it appears appropriate to include a large number of biomarkers in a multidimensional model, and this is the reason why we did not reduce the number of biomarkers in the model. Due to the excellent performance of the classifier, it has not been further optimized because further optimization will be quite challenging, likely requiring additional data on >1,000 subjects.

Another issue is the heterogeneity of the cohort: CKD is a conglomerate of diseases, and this fact is also reflected by the character of the biomarkers. Although certain biomarkers may serve well as classifiers on their own, their combination does not necessarily enhance classification performance. On the other hand, a moderately well performing individual biomarker may in fact increase performance of the classification algorithm, *i.e.* when a particular biomarker reflects disease for only a subset of samples. Therefore, we built an SVM classification model with 273 peptides to distinguish healthy subjects from individuals with biopsy-proven kidney disease in the training set. Subsequently, we applied the model/panel to the test set to test the model. Under such circumstances, overfitting to or "memorizing" the training set and thus poor classification of blinded data sets can be minimized. The classification of a blinded test cohort of 34 healthy controls and 110 patients with CKD resulted in a sensitivity of 85.5% and a specificity of 100.0%. We further examined the specificity of the identified markers using additional disease controls in the human urinary peptidome database. The overall specificity of 97.8% reflects the relatively high 10–20% prevalence of CKD worldwide (61, 62), especially manifested in older patients with diabetes mellitus type II and arterial hypertension (see Table I).

Most of the biomarker peptides in the urine are a product of proteolytic activity. Extracellular proteases may reflect the presence of the disease and its progression (63). Disease-induced changes in protease activities may be more readily recognized by focusing the analysis on the proteolytic fragments rather than on the specific protease itself (63). CE-MS analysis may be suitable to indicate changes in the tightly regulated activity of proteases and protease inhibitors by displaying relative abundance of their potential cleavage products.

Collagen fragments, especially fragments of collagen α-1 (I) chain, appear to be the major constituents of urinary peptides (see Table III). These peptides likely reflect normal physiological turnover of the extracellular matrix (64). In addition to CKD, collagen fragments are also the source of identified biomarkers for the diagnosis of coronary artery disease (CAD) (65, 66). The main difference of the biomarkers for the diagnosis of CAD and of CKD is the direction of their regulation. Whereas most of the collagen-derived biomarkers for CAD showed increased urinary excretion, the collagen-derived biomarkers indicated CKD by their relative paucity (see supplemental Table I). This differential regulation may arise from differences in the activity of collagenases. Increased levels of circulating collagenases have been reported in patients with stable angiographic coronary atherosclerosis or intermittent claudication (65). Elevated matrix metalloprotease-9 activity has been found in unstable plaques, suggesting a crucial role in plaque rupture. In contrast, decreased activity of collagenases was observed in CKD patients (29). Regardless of the primary etiology, CKD in its severe presentation is characterized by tubular atrophy, interstitial fibrosis, and glomerulosclerosis. Hence, it has been assumed that diminished activity of matrix metalloproteases may be responsible for the accumulation of proteins in the extracellular matrix and collagens that typify the fibrotic kidney (67). This histology may be interpreted to indicate increased levels of inhibitors of matrix metalloproteases, *i.e.* tissue inhibitors of matrix metalloproteinase. Furthermore, accumulation of extracellular matrix as predominantly observed in diabetic nephropathy was recently shown to be associated with decreased urinary excretion of several specific collagen fragments (29). These explanations are speculative but are fitting at a conceptual level and have been mentioned here to highlight the significance of the human urinary database in a way that opens a door for further investigation whereby researchers of all disciplines are invited to participate.

α_1-Antitrypsin and its fragments were reported to be upregulated in several types of CKD (57, 63, 68, 69). Thus, increased urinary excretion of a fragment of this highly abundant plasma protein, together with fragments derived from serum albumin and fibrinogen, may reflect chronic renal damage. Uromodulin has been reported to be downregulated in diabetic nephropathy (29, 70, 71), and reduced excretion of specific uromodulin fragments also has been observed in other forms of CKD (30). Decreased urinary levels of these collagen- and uromodulin-derived peptide fragments may serve as indicators of (patho)physiological changes in CKD.

As we learn to better appreciate the individual differences in the responses to therapy, objective methods to measure these changes will become of prime importance. This need will pave the way for proteomics-based personalized diagnosis that could be used to tailor therapy to an individual patient. Non-invasive urinary proteomics has the advantage of real time monitoring, and adjustments to therapy can thus be made accordingly. This vision appears within reach, but its realization depends heavily on the establishment of tools that allow a quick and robust comparison of the patient's proteomic profiles against those of healthy controls and patients with other diseases. Thus, we posit that the urinary peptidome and its application in CKD as presented here are a major step forward in this direction. We anticipate that the availability of such tools in the future will significantly improve the options for patients with respect to diagnosis and therapy.

The present study with several hundred patients reveals the potential of distinct biomarker panels in clinical diagnosis of complex chronic diseases. Thorough platform standardization is needed to address the intrinsic/biological variability of the human peptidome. The human urinary peptidome database presented here is an important step toward proteomics becoming a diagnostic tool in the clinical settings as recently demonstrated for graft-*versus*-host disease (26).

REFERENCES

1. Alebiosu, C. O., and Ayodele, O. E. (2005) The global burden of chronic kidney disease and the way forward. *Ethn. Dis.* **15,** 418–423
2. Good, D. M., Thongboonkerd, V., Novak, J., Bascands, J. L., Schanstra, J. P., Coon, J. J., Dominiczak, A., and Mischak, H. (2007) Body fluid proteomics for biomarker discovery: lessons from the past hold the key to success in the future. *J. Proteome Res.* **6,** 4549–4555
3. Kolch, W., Neusüss, C., Pelzing, M., and Mischak, H. (2005) Capillary electrophoresis-mass spectrometry as a powerful tool in clinical diagnosis and biomarker discovery. *Mass Spectrom. Rev.* **24,** 959–977
4. González-Buitrago, J. M., Ferreira, L., and Lorenzo, I. (2007) Urinary proteomics. *Clin. Chim. Acta* **375,** 49–56
5. Fliser, D., Novak, J., Thongboonkerd, V., Argilés, A., Jankowski, V., Girolami, M. A., Jankowski, J., and Mischak, H. (2007) Advances in urinary proteome analysis and biomarker discovery. *J. Am. Soc. Nephrol.* **18,** 1057–1071
6. Vidal, B. C., Bonventre, J. V., and I-Hong Hsu, S. (2005) Towards the application of proteomics in renal disease diagnosis. *Clin. Sci.* **109,** 421–430
7. Pieper, R., Gatlin, C. L., McGrath, A. M., Makusky, A. J., Mondal, M., Seonarain, M., Field, E., Schatz, C. R., Estock, M. A., Ahmed, N., Anderson, N. G., and Steiner, S. (2004) 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* **4,** 1159–1174
8. Thongboonkerd, V., McLeish, K. R., Arthur, J. M., and Klein, J. B. (2002) Proteomic analysis of normal human urinary proteins isolated by acetone precipitation or ultracentrifugation. *Kidney Int.* **62,** 1461–1469
9. Pisitkun, T., Shen, R. F., and Knepper, M. A. (2004) Identification and proteomic profiling of exosomes in human urine. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 13368–13373
10. Marshall, T., and Williams, K. (1996) Two-dimensional electrophoresis of human urinary proteins following concentration by dye precipitation. *Electrophoresis* **17,** 1265–1272
11. Shihabi, Z. K., Konen, J. C., and O'Connor, M. L. (1991) Albuminuria vs urinary total protein for detecting chronic renal disorders. *Clin. Chem.* **37,** 621–624
12. Yudkin, J. S., Forrest, R. D., and Jackson, C. A. (1988) Microalbuminuria as predictor of vascular disease in non-diabetic subjects. Islington Diabetes Survey. *Lancet* **2,** 530–533
13. Schaub, S., Rush, D., Wilkins, J., Gibson, I. W., Weiler, T., Sangster, K., Nicolle, L., Karpinski, M., Jeffery, J., and Nickerson, P. (2004) Proteomic-based detection of urine proteins associated with acute renal allograft rejection. *J. Am. Soc. Nephrol.* **15,** 219–227
14. Büeler, M. R., Wiederkehr, F., and Vonderschmitt, D. J. (1995) Electrophoretic, chromatographic and immunological studies of human urinary proteins. *Electrophoresis* **16,** 124–134
15. Adachi, J., Kumar, C., Zhang, Y., Olsen, J. V., and Mann, M. (2006) The human urinary proteome contains more than 1500 proteins including a large proportion of membranes proteins. *Genome Biol.* **7,** R80.1–R80.16
16. Ngai, H. H., Sit, W. H., Jiang, P. P., Xu, R. J., Wan, J. M., and Thongboonkerd, V. (2006) Serial changes in urinary proteome profile of membranous nephropathy: implications for pathophysiology and biomarker discovery. *J. Proteome. Res.* **5,** 3038–3047
17. Oh, J., Pyo, J. H., Jo, E. H., Hwang, S. I., Kang, S. C., Jung, J. H., Park, E. K., Kim, S. Y., Choi, J. Y., and Lim, J. (2004) Establishment of a near-standard two-dimensional human urine proteomic map. *Proteomics* **4,** 3485–3497
18. Spahr, C. S., Davis, M. T., McGinley, M. D., Robinson, J. H., Bures, E. J., Beierle, J., Mort, J., Courchesne, P. L., Chen, K., Wahl, R. C., Yu, W., Luethy, R., and Patterson, S. D. (2001) Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. *Proteomics* **1,** 93–107
19. Cadieux, P. A., Beiko, D. T., Watterson, J. D., Burton, J. P., Howard, J. C., Knudsen, B. E., Gan, B. S., McCormick, J. K., Chambers, A. F., Denstedt, J. D., and Reid, G. (2004) 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.* **18,** 170–175
20. Schiffer, E., Mischak, H., and Novak, J. (2006) High resolution proteome/peptidome analysis of body fluids by capillary electrophoresis coupled with MS. *Proteomics* **6,** 5615–5627
21. Decramer, S., Wittke, S., Mischak, H., Zürbig, P., Walden, M., Bouissou, F., Bascands, J. L., and Schanstra, J. P. (2006) Predicting the clinical outcome of congenital unilateral ureteropelvic junction obstruction in newborn by urinary proteome analysis. *Nat. Med.* **12,** 398–400
22. Haubitz, M., Wittke, S., Weissinger, E. M., Walden, M., Rupprecht, H. D., Floege, J., Haller, H., and Mischak, H. (2005) Urine protein patterns can serve as diagnostic tools in patients with IgA nephropathy. *Kidney Int.* **67,** 2313–2320
23. Theodorescu, D., Fliser, D., Wittke, S., Mischak, H., Krebs, R., Walden, M., Ross, M., Eltze, E., Bettendorf, O., Wulfing, C., and Semjonow, A. (2005) Pilot study of capillary electrophoresis coupled to mass spectrometry as a tool to define potential prostate cancer biomarkers in urine. *Electrophoresis* **26,** 2797–2808
24. Rossing, K., Mischak, H., Parving, H. H., Christensen, P. K., Walden, M., Hillmann, M., and Kaiser, T. (2005) Impact of diabetic nephropathy and angiotensin II receptor blockade on urinary polypeptide patterns. *Kidney Int.* **68,** 193–205
25. Weissinger, E. M., Wittke, S., Kaiser, T., Haller, H., Bartel, S., Krebs, R., Golovko, I., Rupprecht, H. D., Haubitz, M., Hecker, H., Mischak, H., and Fliser, D. (2004) Proteomic patterns established with capillary electrophoresis and mass spectrometry for diagnostic purposes. *Kidney Int.* **65,** 2426–2434
26. Weissinger, E. M., Schiffer, E., Hertenstein, B., Ferrara, J. L., Holler, E., Stadler, M., Kolb, H. J., Zander, A., Zürbig, P., Kellmann, M., and Ganser, A. (2007) Proteomic patterns predict acute graft-versus-host disease after allogeneic hematopoietic stem cell transplantation. *Blood* **109,** 5511–5519
27. Theodorescu, D., Wittke, S., Ross, M. M., Walden, M., Conaway, M., Just, I., Mischak, H., and Frierson, H. F. (2006) Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. *Lancet Oncol.* **7,** 230–240
28. Julian, B. A., Wittke, S., Novak, J., Good, D. M., Coon, J. J., Kellmann, M., Zürbig, P., Schiffer, E., Haubitz, M., Moldoveanu, Z., Calcatera, S. M., Wyatt, R. J., Sýkora, J., Sládková, E., Hes, O., Mischak, H., and McGuire, B. M. (2007) Electrophoretic methods for analysis of urinary polypeptides in IgA-associated renal diseases. *Electrophoresis* **28,** 4469–4483
29. Rossing, K., Mischak, H., Dakna, M., Zürbig, P., Novak, J., Julian, B. A., Good, D. M., Coon, J. J., Tarnow, L., and Rossing, P. (2008) Urinary Proteomics in Diabetes and CKD. *J. Am. Soc. Nephrol.* **19,** 1283–1290
30. Haubitz, M., Good, D. M., Woywodt, A., Haller H, Rupprecht, H., Theodorescu, D., Dakna, M., Coon, J. J., and Mischak, H. (2009) Identification and validation of urinary biomarkers for differential diagnosis and evaluation of therapeutic intervention in anti-neutrophil cytoplasmic antibody-associated vasculitis. *Mol. Cell. Proteomics* **8,** 2296–2307
31. Frommberger, M., Zürbig, P., Jantos, J., Krahn, T., Mischak, M., Pich, A., Just, I., Schmitt-Kopplin, P., and Schiffer, E. (2007) Peptidomic analysis

of rat urine using capillary electrophoresis coupled to mass spectrometry. *Proteomics Clin. Appl.* **1,** 650–660

32. Neuhoff, N., Kaiser, T., Wittke, S., Krebs, R., Pitt, A., Burchard, A., Sundmacher, A., Schlegelberger, B., Kolch, W., and Mischak, H. (2004) 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.* **18,** 149–156

33. Jantos-Siwy, J., Schiffer, E., Brand, K., Schumann, G., Rossing, K., Delles, C., Mischak, H., and Metzger, J. (2009) Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. *J. Proteome. Res.* **8,** 268–281

34. Mischak, H., Julian, B. A., and Novak, J. (2007) High-resolution proteome/peptidome analysis of peptides and low-molecular-weight proteins in urine. *Proteomics Clin. Appl.* **1,** 792–804

35. DeLeo, J. M. (1993) Receiver operating characteristic laboratory (ROCLAB): software for developing decision strategies that account for uncertainty, in *Second International Symposium on Uncertainty Modeling and Analysis, College Park, April 25–28, 1993*, pp. 318–325, IEEE, Washington, D. C.

36. Tusher, V. G., Tibshirani, R., and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 5116–5121

37. Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J. Y., and Zhang, J. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5,** R80.1–R80.16

38. R Development Core Team (2008) *R: a Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria

39. Girolami, M., Mischak, H., and Krebs, R. (2006) Analysis of complex, multidimensional datasets. *Drug Discov. Today Technol.* **3,** 12–19

40. Yang, Z. R. (2004) Biological applications of support vector machines. *Brief. Bioinform.* **5,** 328–338

41. Yang, Z. R., and Chou, K. C. (2004) Bio-support vector machines for computational proteomics. *Bioinformatics* **20,** 735–741

42. Zürbig, P., Renfrow, M. B., Schiffer, E., Novak, J., Walden, M., Wittke S., Just, I., Pelzing, M., Neusüss, C., Theodorescu, D., Root, K. E., Ross, M. M., and Mischak, H. (2006) Biomarker discovery by CE-MS enables sequence analysis via MS/MS with platform-independent separation. *Electrophoresis* **27,** 2111–2125

43. Coon, J. J., Shabanowitz, J., Hunt, D. F., and Syka, J. E. (2005) Electron transfer dissociation of peptide anions. *J. Am. Soc. Mass Spectrom.* **16,** 880–882

44. Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 9528–9533

45. Good, D. M., and Coon, J. J. (2006) Advancing proteomics with ion/ion chemistry. *BioTechniques* **40,** 783–789

46. Good, D. M., Wirtala, M., McAlister, G. C., and Coon, J. J. (2007) Performance characteristics of electron transfer dissociation mass spectrometry. *Mol. Cell. Proteomics* **6,** 1942–1951

47. Good, D. M., Wenger, C. D., McAlister, G. C., Bai, D. L., Hunt, D. F., and Coon, J. J. (2009) Post-acquisition ETD spectral processing for increased peptide identifications. *J. Am. Soc. Mass Spectrom.* **20,** 1435–1440

48. Good, D. M., Wenger, C. D., and Coon, J. J. (2010) The effect of interfering ions on search algorithm performance for electron-transfer dissociation data. *Proteomics* **10,** 164–167

49. Williams, B. J., Russell, W. K., and Russell, D. H. (2007) Utility of CE-MS data in protein identification. *Anal. Chem.* **79,** 3850–3855

50. Taylor, C. F., Paton, N. W., Lilley, K. S., Binz, P. A., Julian, R. K., Jr., Jones, A. R., Zhu, W., Apweiler, R., Aebersold, R., Deutsch, E. W., Dunn, M. J., Heck, A. J., Leitner, A., Macht, M., Mann, M., Martens, L., Neubert, T. A., Patterson, S. D., Ping, P., Seymour, S. L., Souda, P., Tsugita, A., Vandekerckhove, J., Vondriska, T. M., Whitelegge, J. P., Wilkins, M. R., Xenarios, I., Yates, J. R., 3rd, and Hermjakob, H. (2007) The minimum information about a proteomics experiment (MIAPE). *Nat. Biotechnol.* **25,** 887–893

51. Dudoit, S., and van der Laan, M. J. (2007) *Multiple Testing Procedures with Applications to Genomics*, Springer, Berlin

52. Mischak, H., Apweiler, R., Banks, R. E., Conaway, M., Coon, J. J., Dominizak, A., Ehrich, J. H., Fliser, D., Girolami, M., Hermjakob, H., Hochstrasser, D. F., Jankowski, V., Julian, B. A., Kolch, W., Massy, Z., Neususs, C., Novak, J., Peter, K., Rossing, K., Schanstra, J. P., Semmes, O. J., Theodorescu, D., Thongboonkerd, V., Weissinger, E. M., Van Eyk, J. E., and Yamamoto, T. (2007) Clinical proteomics: a need to define the field and to begin to set adequate standards. *Proteomics Clin. Appl.* **1,** 148–156

53. Woodcock, J. (2007) The prospects for "personalized medicine" in drug development and drug therapy. *Clin. Pharmacol. Ther.* **81,** 164–169

54. Castagna, A., Cecconi, D., Sennels, L., Rappsilber, J., Guerrier, L., Fortis, F., Boschetti, E., Lomas, L., and Righetti, P. G. (2005) Exploring the hidden human urinary proteome via ligand library beads. *J. Proteome. Res.* **4,** 1917–1930

55. Chalmers, M. J., Mackay, C. L., Hendrickson, C. L., Wittke, S., Walden, M., Mischak, H., Fliser, D., Just, I., and Marshall, A. G. (2005) Combined top-down and bottom-up mass spectrometric approach to characterization of biomarkers for renal disease. *Anal. Chem.* **77,** 7163–7171

56. Mischak, H., Kaiser, T., Walden, M., Hillmann, M., Wittke, S., Herrmann, A., Knueppel, S., Haller, H., and Fliser, D. (2004) Proteomic analysis for the assessment of diabetic renal damage in humans. *Clin. Sci.* **107,** 485–495

57. Varghese, S. A., Powell, T. B., Budisavljevic, M. N., Oates, J. C., Raymond, J. R., Almeida, J. S., and Arthur, J. M. (2007) Urine biomarkers predict the cause of glomerular disease. *J. Am. Soc. Nephrol.* **18,** 913–922

58. Girolami, M., and Rogers, S. (2006) Variational Bayesian multinomial probit regression with Gaussian process priors. *Neural Comput.* **18,** 1790–1817

59. Burges, C. J. C. (1998) A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* **2,** 121–167

60. Rasmussen, C. E., and Williams, C. K. I. (2006) *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, MA

61. Wen, C. P., Cheng, T. Y., Tsai, M. K., Chang, Y. C., Chan, H. T., Tsai, S. P., Chiang, P. H., Hsu, C. C., Sung, P. K., Hsu, Y. H., and Wen, S. F. (2008) All-cause mortality attributable to chronic kidney disease: a prospective cohort study based on 462 293 adults in Taiwan. *Lancet* **371,** 2173–2182

62. Xu, R., Zhang, L., Zhang, P., Wang, F., Zuo, L., and Wang, H. (2009) Comparison of the prevalence of chronic kidney disease among different ethnicities: Beijing CKD survey and American NHANES. *Nephrol. Dial. Transplant.* **24,** 1220–1226

63. Candiano, G., Musante, L., Bruschi, M., Petretto, A., Santucci, L., Del Boccio, P., Pavone, B., Perfumo, F., Urbani, A., Scolari, F., and Ghiggeri, G. M. (2006) Repetitive fragmentation products of albumin and alpha1-antitrypsin in glomerular diseases associated with nephrotic syndrome. *J. Am. Soc. Nephrol.* **17,** 3139–3148

64. Rossing, K., Mischak, H., Rossing, P., Schanstra, J. P., Wiseman, A., and Maahs, D. M. (2008) The urinary proteome in diabetes and diabetes-associated complications: new ways to assess disease progression and evaluate therapy. *Proteomics Clin. Appl.* **2,** 997–1007

65. Zimmerli, L. U., Schiffer, E., Zürbig, P., Good, D. M., Kellmann, M., Mouls, L., Pitt, A. R., Coon, J. J., Schmieder, R. E., Peter, K. H., Mischak, H., Kolch, W., Delles, C., and Dominiczak, A. F. (2008) Urinary proteomics biomarkers in coronary artery disease. *Mol. Cell. Proteomics* **7,** 290–298

66. von Zur Muhlen C., Schiffer, E., Zuerbig, P., Kellmann, M., Brasse, M., Meert, N., Vanholder, R. C., Dominiczak, A. F., Chen, Y. C., Mischak, H., Bode, C., and Peter, K. (2009) Evaluation of urine proteome pattern analysis for its potential to reflect coronary artery atherosclerosis in symptomatic patients. *J. Proteome. Res.* **8,** 335–345

67. Cheng, S., Pollock, A. S., Mahimkar, R., Olson, J. L., and Lovett, D. H. (2006) Matrix metalloproteinase 2 and basement membrane integrity: a unifying mechanism for progressive renal injury. *FASEB J.* **20,** 1898–1900

68. Kwak, N. J., Heo, Y., Jin, D. C., Cha, J. H., Lee, K. H., Yang, C. W., Kang, C. S., Choi, Y. J., and Wang, E. H. (2007) Proteomic analysis of alpha-1-antitrypsin in immunoglobulin A nephropathy. *Proteomics Clin. Appl.* **1,** 420–428

69. Argiles, A., Mourad, G., Mion, C., Atkins, R. C., and Haiech, J. (1990) Two-dimensional gel electrophoresis of urinary proteins in kidney diseases. *Contrib. Nephrol.* **83,** 1–8

70. Chakraborty, J., Below, A. A., and Solaiman, D. (2004) Tamm-Horsfall protein in patients with kidney damage and diabetes. *Urol. Res.* **32,** 79–83

71. Sejdiu, I., and Torffvit, O. (2008) Decreased urinary concentration of Tamm-Horsfall protein is associated with development of renal failure and cardiovascular death within 20 years in type 1 but not in type 2 diabetic patients. *Scand. J. Urol. Nephrol.* **42,** 168–174