



HAL
open science

Urine in clinical proteomics.

Stéphane Decramer, Anne Gonzalez de Peredo, Benjamin Breuil, Harald Mischak, Bernard Monsarrat, Jean-Loup Bascands, Joost P Schanstra

► **To cite this version:**

Stéphane Decramer, Anne Gonzalez de Peredo, Benjamin Breuil, Harald Mischak, Bernard Monsarrat, et al.. Urine in clinical proteomics.. *Molecular and Cellular Proteomics*, 2008, 7 (10), pp.1850-62. 10.1074/mcp.R800001-MCP200 . hal-00360866

HAL Id: hal-00360866

<https://hal.science/hal-00360866>

Submitted on 12 Feb 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Urine in clinical proteomics

Stéphane Decramer^{1, 2, 3}, Anne Gonzalez de Peredo⁴, Benjamin Breuil^{1,2}, Harald Mischak⁵, Bernard Monsarrat⁴, Jean-Loup Bascands^{1, 2}, Joost P. Schanstra^{1,2*}

¹ Inserm, U858/I2MR, Department of Renal and Cardiac Remodeling, Team #5, 1 Avenue Jean Poulhès, BP 84225, 31432 Toulouse Cedex 4, France

² Université Toulouse III Paul Sabatier, Institut de Médecine Moléculaire de Rangueil, Toulouse, F-31000 France

³ Department of Pediatric Nephrology, Hôpital des Enfants, Toulouse, France. Centre de Référence du Sud Ouest des Maladies Rénales Rares.

⁴ Laboratoire de Protéomique et Spectrométrie de Masse des Biomolécules, Institut de Pharmacologie et de Biologie Structurale, CNRS UMR 5089, 205 route de Narbonne, 31077, Toulouse, France.

⁵ Mosaiques Diagnostics & Therapeutics AG, Hannover, Germany.

* Corresponding author:

Joost P. Schanstra

Inserm, U858/I2MR, Equipe 5

BP 84225

31432 TOULOUSE Cedex 4

France

E-mail: joost-peter.schanstra@toulouse.inserm.fr

Abstract

Urine has become one of the most attractive biofluids in clinical proteomics as it can be obtained non-invasively, in large quantities and is stable compared to other biofluids. The urinary proteome has been studied by almost any proteomic technology, but mass-spectrometry-based urinary protein and peptide profiling has emerged as most suitable for clinical application. After a period of descriptive urinary proteomics the field is moving out of the discovery phase into an era of validation of urinary biomarkers in larger prospective studies. Although, mainly due to the site of production of urine, the majority of these studies apply to the kidney and the urinary tract, recent data shows that analysis of the urinary proteome can also be highly informative on non urogenital diseases and be used in their classification. Despite this progress in urinary biomarker discovery, the contribution of urinary proteomics to the understanding of the pathophysiology of disease upon analysis of the urinary proteome is still modest, mainly due to problems associated to sequence identification of the biomarkers. Until now, research has focused on the highly abundant urinary proteins and peptides but analysis of the less abundant, and naturally existing urinary proteins and peptides still remains a challenge. In conclusion, urine has evolved as one of the most attractive bodyfluids in clinical proteomics with, potentially, a rapid application in the clinic.

Table of contents

1. Urine

- 1.1. Production
- 1.2. Urinary protein content
- 1.3. Urine as a source for biomarkers
 - 1.3.1. Advantages*
 - 1.3.2. Disadvantages*

2. Urine in clinical proteomics: techniques and prerequisites

- 2.1. Techniques
 - 2.1.1. SELDI-TOF*
 - 2.1.2. CE-MS*
 - 2.1.3. LC-MS*
 - 2.1.4. 2DE-MS*
- 2.2. The need for standards and the new trend
 - 2.2.1. Standards*
 - 2.2.2. New trend: from single markers to panels*

3. The use of urine in clinical proteomics

- 3.1. Urogenital disease: Non cancer.
 - 3.1.1. Kidney Transplantation*
 - 3.1.2. Chronic kidney disease*
 - 3.1.3. Diabetic nephropathy*
 - 3.1.4. Obstructive nephropathy*
- 3.2. Urogenital disease: Cancer.
 - 3.2.1. Renal cell carcinoma*
 - 3.2.2. Bladder cancer*
 - 3.2.3. Prostate cancer*
- 3.3. Application of urinary proteome analysis to non-urogenital diseases

4. From biomarkers to pathophysiology

5. Chasing low abundance urinary proteins

6. Conclusions

1. Urine

1.1. Production

Human urine plays an important role in clinical diagnostics. Physicians have examined urinary samples from patients to diagnose various disorders for centuries. The philosopher Hermogenes (5th century BCE) already described the color and other attributes of urine as indicators of certain diseases (1). Urine is produced by the kidney and allows the human body to eliminate waste products from blood. The kidney also maintains whole body homeostasis and produces hormones including renin and erythropoietin (2). The human kidney (**Figure 1**) is composed of 1 million functional units called nephrons, which can be divided in two functional parts: the glomerulus, which filters the plasma yielding the, so-called, “primitive” urine; and the renal tubule, which reabsorbs most of the primitive urine. In 24 hours, about 900 liters of plasma flows through the kidneys of which 150-180 L is filtered. However, more than 99% of this primitive urine is reabsorbed. The remainder (the “final” urine) exits the kidney *via* the ureter into the bladder (**Figure 1**). Therefore urine may contain information not only from the kidney and the urinary tract, but also from more distant organs via plasma, obtained by glomerular filtration. In healthy individuals, 70% of the urinary proteome originates from the kidney and the urinary tract while the remaining 30% represents proteins filtered by the glomerulus (3). The analysis of the urinary proteome might therefore allow the identification of biomarkers of both urogenital and systemic diseases.

1.2. Urinary protein content

Urine from a healthy individual contains a significant amount of peptides and proteins. The number of proteins and peptides identified in urine is still increasing. One of the first attempts to define the urinary proteome was published in 2001 (4). Using LC-MS, tryptic peptides of pooled urine samples were analyzed and 124 proteins were identified. While this study was not designed to define urinary biomarkers for disease, it showed the information potentially hidden in the urinary proteome and also indicated a possible approach towards its mining. In 2004, this number increased to 1400 distinct spots on two-dimensional electrophoresis gels, of which about 420 identified spots yielded 150 unique protein annotations (5). This number of identified urinary proteins increased significantly to around 1500 in 2006 by combining one-dimensional gel electrophoresis and reverse phase liquid chromatography coupled to (Orbitrap) mass-spectrometry (6), further underlining the complexity of the human urinary proteome. In a very recent study (2008), we determined that the human urinary proteome apparently contains over 100000 different peptides, at least 5000 with high frequency (observed in > 40% of individuals examined in different studies) (7). It is therefore “save to state” that urine is indeed a rich non-invasive source of potential biomarkers of disease that awaits exploration.

1.3. Urine as a source for biomarkers

1.3.1. Advantages

Compared to other body fluids, urine has several characteristics that make it a preferred choice for biomarker discovery: *First*, urine can be obtained in large quantities using non-invasive procedures. This allows repeated sampling of the same individual for disease surveillance. The availability of urine also allows easy

assessment of reproducibility or improvement in sample preparation protocols. *Second*, urinary peptides and lower molecular mass proteins are generally soluble. Therefore solubilization of these low molecular weight proteins and peptides, a process with a major influence on the proteomic analysis of cells or tissues, generally is no issue. Further, these lower-molecular-weight compounds (< 30 kDa) can be analyzed in a mass spectrometer without additional manipulation (e.g., tryptic digests). *Third*, in general, the urinary protein content is relatively stable, probably due to the fact that urine “stagnates” for hours in the bladder, hence proteolytic degradation by endogenous proteases may be essentially complete by the time of voiding. This is in sharp contrast to blood, for which activation of proteases (and consequently generation of an array of proteolytic breakdown products) is inevitably associated with its collection (8, 9). Two laboratories independently showed that the urinary proteome did not change significantly when urine was stored up to 3 days at 4°C or up to 6 hours at room temperature (10, 11). In addition, urine can be stored for several years at -20°C without significant alteration of its proteome. However, these considerations may not apply to specialized applications, such as the recently described urinary exosomes that may be less stable (12). *Finally*, as described above, not only the changes in the kidney and genitourinary tract are reflected by changes in the urinary proteome, but also changes at more distant sites. This will be developed in more detail below.

1.3.2. Disadvantages

Urine has the disadvantage that it widely varies in protein and peptide concentrations mostly due to differences in the daily intake of fluid. However, this shortcoming can be countered by standardization based on creatinine (13) or peptides generally

present in urine (14). In addition, definition of disease-specific biomarkers in urine, and most likely in other compartments, is complicated by significant changes in the proteome during the day. These changes are likely caused by variations in the diet, metabolic or catabolic processes, circadian rhythms, exercise, as well as circulatory levels of various hormones (15). The reproducibility of any analysis is reduced by these physiological changes, even if the analytical method shows high reproducibility. However, these variations appear mostly limited to a fraction of the urinary proteome; a large portion remains unaffected by these processes (16).

2. Urine in clinical proteomics: techniques and prerequisites

2.1. Techniques

Almost any known mass-spectrometry technique has been used for the analysis of the urinary proteome including two dimensional gel-electrophoresis followed by mass spectrometry (2DE-MS), liquid chromatography coupled to mass spectrometry (LC-MS), surface enhanced laser desorption/ionisation coupled to mass spectrometry (SELDI-TOF) and capillary electrophoresis coupled to mass spectrometry (CE-MS). (17). The ideal sequence for biomarker discovery would be mass-spectrometry based discovery, followed by Elisa-based validation and clinical application. This is easier said than done, and to our knowledge, no examples of this ideal “sequence” are available for the discovery of urinary biomarkers, yet. However, over the last few years profiling approaches which allow the use of mass-spectrometry-based techniques in the discovery/validation/clinical phase for the analysis of the urinary proteome have emerged: SELDI-TOF and CE-MS (17). The technical details of both techniques can be found in detail elsewhere (17, 18), but we will describe briefly the advantages and disadvantages of both SELDI-TOF and CE-MS (**Table**). In addition we will describe LC-MS and 2DE-MS, which, until now, may be used for biomarker discovery, but not for clinical applications.

2.1.1. SELDI-TOF

Advantages of SELDI-TOF include the capacity to analyze multiple samples in a short time and its ease of use. Therefore SELDI-TOF has been used in numerous studies aiming the definition of biomarkers (19). Although the technology is easy to use, it is, unfortunately, also prone to generating artifacts (20, 21). This may be due,

in part, to difficulties with calibration and lack of precision of the determined molecular masses of the analytes. Furthermore, only a very small fraction of all proteins in a sample binds to the chip surface. Therefore, only a fraction of the information contained in a biological sample can be exploited for the presence of biomarkers, even if there are a number of different chip-surfaces available. In addition, binding to the different chip surfaces varies depending on samples concentration, pH, salt content, and the presence of interfering compounds. Finally the SELDI-TOF instrument cannot be directly interfaced with MS/MS instruments for sequencing.

2.1.2. CE-MS

CE-MS (18) provides relatively fast analysis (1h) and high resolution, it is rather robust and compatible with most buffers and analytes and it provides a stable constant flow avoiding elution gradients that may interfere with MS detection. A disadvantage of CE is that the analysis is restricted to low molecular weight proteins as the larger proteins tend to precipitate at the low pH generally used in the running buffer. This might be seen as a drawback, but is of little consideration for the analysis of urine, as the urinary proteome contains a high percentage of low molecular weight proteins (7). Another potential drawback of CE-MS is that only a relatively small amount of volume can be loaded onto the capillary, leading to a potentially lower sensitivity of detection. However, improvement of both, coupling and the detection limits of mass spectrometers enable detection in the amol range, making this issue less relevant (22). Sequencing of CE-MS defined biomarkers can be performed (although with limited success due to low amount of sample volume that can be

loaded) by direct interfacing of a CE with MS/MS instruments (23), or by subsequent targeted sequencing (24).

2.1.3. LC-MS

Other techniques, such as LC-MS or 2DE-MS, have been used to study the urinary proteome, but have, in general, only been applied on a reduced number of individuals without subsequent blinded validation.

The majority of the approaches based on LC-MS rely on digestion of the sample with trypsin, and separation of the resulting tryptic peptides by nanoLC-MS. Using this method, sequencing of tryptic peptides by tandem mass spectrometry (nanoLC-MS/MS) can be automatically triggered, providing sequence information on the peptides detected and identification of the proteins from which they derive. NanoLC-MS/MS has proved efficient for qualitative description of the urine proteome (6, 25-28). However, this approach suffers currently from, at least, two drawbacks for sample profiling and biomarker discovery: i) as the sample is digested with trypsin, the complexity of the resulting mixture is much higher than that of the starting material. This leads to MS/MS undersampling, resulting in incomplete analytic coverage of the digest (29). This would be less of a issue if differential studies could be performed based on the MS signal of tryptic peptides and not only on comparison of protein lists identified by MS/MS. This might soon be the case with the evolution of mass spectrometers toward high mass accuracy and resolution, and the development of new bioinformatic software for peptide patterns alignment across multiple runs (30-34). ii) even with modern instruments, nanoLC-MS profiling of highly complex tryptic mixtures will probably remain a quite laborious process that may be applicable to only a limited number of patients. This problem might be circumvented by the use of targeted mass spectrometry approaches like MRM (Multiple Reaction

Monitoring) to validate potential biomarkers previously identified by nanoLC-MS/MS in the discovery phase on a reduced number of patients. MRM based approaches allow to quantify with high sensitivity and selectivity peptides in complex mixtures and may be applied at high throughput to screen simultaneously several biomarkers on large cohorts of patients (35, 36). This method has been applied recently to detect proteins at very low levels in plasma (37). The use of such hypothesis-driven approaches generated by nanoLC-MS/MS on a limited number of samples may also prove to be useful in the future for the clinical validation of candidate markers in urine.

2.1.4. 2DE-MS

2DE-MS is still the most accessible technique, allows to study large molecules and has been used on numerous occasions for the description of the urinary proteome (38). However, it has the drawback that the reproducibility is low, time of analysis is long and the technique is difficult to automate. The recently introduced concept of 2D difference gel electrophoresis (2D-DIGE) using fluorescence dyes and internal standards provides better reproducibility and more accurate quantification (39). This allows the satisfactory comparison of two samples, but the comparison of several different experiments remains a challenge. However, recently, the first studies showing the use of 2D-DIGE to compare the urinary proteomes of multiple healthy individuals and different disease states (40, 41) were published. Another limitation is that it is impossible to study peptides (in general: peptides/proteins < 10 kDa) by 2DE.

2.2. The need for standards and the new trend

2.2.1. Standards

Appropriate techniques are not the only prerequisites for clinical proteomics. Basic principles should be applied to increase the chances of clinical application of the identified biomarkers. This issue has been discussed in detail in a number of recent papers to which we refer the reader (38, 42, 43), and we only summarize the guidelines that need to be respected in any clinical proteomic study.

The technical platform must be well characterized (standard operating procedures, known technical variability of the platform) and allow appropriately precise measurements. To reduce the inevitable biological variability to a minimum, standard protocols for urine collection and preparation, as outlined recently (38) are highly advisable, together with high numbers of comparable datasets. Standardization of protocols will allow exchanging resources and data between laboratories and increase the potential application of urinary biomarkers. World-wide (Human Kidney and Urine Proteome Project (HKUPP) (44)) and European initiatives (European Kidney and Urinary Proteomics (EuroKUP) (45)) are on-going for standardization of kidney and urine proteomics. In addition, several publications describing detailed urine sample preparation protocols for specific proteomics applications were recently published (38, 46-48).

It is imperative that proper statistical methods are being used combined with a precise clinical question or hypothesis (42). A Student T-test is insufficient. Correction for multiple testing (e.g. adjustment according to Bonferroni (49), or similar like (50-52)) must be applied. As several of the underlying hypotheses for statistical evaluation (e.g. even distribution of data, comparability of datasets, absence of bias, etc.) are generally not fulfilled and statistics does not enable assessment of correct

classification rate, the results must be validated using an independent blinded set of samples.

Sequencing and exact definition of the detected native proteins or peptides upon profiling represents a frequently underestimated issue. Using profiling methods such as SELDI-TOF or CE-MS, biomarkers are defined by several physical characteristics (e.g. mass and affinity to a certain surface chip for SELDI-TOF, accurate mass and retention time for CE-MS). The use of a bottom-up analysis for sequence determination (i.e. MS/MS analysis after tryptic digestion and database search) is difficult to implement because the biomarkers have to be previously purified. In the case of SELDI-TOF, this is done usually by using the chromatographic matrix related to the surface chip on which the biomarker was identified. In the case of CE-MS studies, the use of preparative CE to isolate a specific fraction containing the marker of interest has also been described, although with limited success (53). However, upon tryptic digestion of a biomarker, which often represents itself a fragment of a larger protein, connectivity to the mass of the biomarker is lost. Moreover, the bottom-up approach usually does not take posttranslational modifications (PTM) into account. PTM are a major, sometimes even the most important part of biomarker definition and failure to identify them may subsequently result in failure of the validation process (18, 42, 43). Consequently, it may be more appropriate and accurate to define a potential biomarker via several physical characteristics (e.g. precise mass and retention time). Recent advances in the field of Fourier Transform Ion Cyclotron Resonance (FT-ICR) and Orbitrap instruments and the introduction of Electron Transfer Dissociation (ETD) now allow sequencing of molecules > 10 kDa. While these approaches are not routine methods yet, they

clearly show the path towards sequence determination that does enable accurate definition of biomarkers.

As outlined recently in greater detail (43), currently, the privileged pathway of biomarker discovery with the largest chances to be applicable in the clinic consists of: a clear clinical question → many samples obtained in a standardized fashion → analysis by instrumentation allowing relatively high throughput and high reproducibility → appropriate statistical analysis for these type of large sample numbers → validation of the potential biomarkers in a blinded study → sequencing of these biomarkers.

2.2.2. New trend: from single markers to panels

The potential of a protein or peptide to serve as a biomarker depends on how selective and sensitive it enables disease assessment. Most of the analytes currently used in the clinical laboratory for screening and diagnostic purposes have been identified based on knowledge of the underlying disease gathered over a long period of time. This tedious and laborious procedure often resulted in the identification of single markers with frequently only moderate diagnostic value, mostly due to low specificity. For example, prostate specific antigen (PSA) is currently widely used as a marker for prostate cancer. Its prognostic relevance, however, is the subject of ongoing debates due to a lack of specificity when PSA levels are only moderately increased (4-10 ng/mL). This uncertainty not only results in unnecessary biopsies, but also in higher rates of false positive prostate cancer diagnosis (54). Another example is the use of microalbuminuria as an early non-invasive marker of renal damage. Microalbuminuria can be present in diabetic patients before apparent damage to glomerular function or increased serum creatinine levels (55). However,

microalbuminuria is also found in apparently healthy individuals, and cannot be utilized as a predictive marker of renal disease (56). These two examples underline the need for more accurate biomarkers.

Can a single marker fulfill the requirements to reliably detect a disease as early as possible, to unambiguously distinguish it from other pathological conditions, and to monitor the efficacy of therapy? Probably not. An alternative strategy is identification of several markers, which as a stand-alone marker do not present high specificity and sensitivity. But, as a panel (or pattern) the markers work in concert (18). The general criteria that are applied onto a biomarker to be used for clinical assessment (*e.g.* known identity, reproducible detection, known deviation) also apply for the single biomarkers that makeup the multi-marker panel (43).

3. The use of urine in clinical proteomics

The field of biomarker identification using urinary proteomics is moving towards application phase. Most of the studies described below showed potential for clinical application and adhered to the recommendations for biomarker discovery described above. A number of older studies will also be cited to give more complete overview of the attempts to identify biomarkers. As the majority of the urinary proteins and peptides have been found to originate from the kidney and the urinary tract (3), most of the completed studies have focused on these organs and tissues.

3.1. Urogenital disease: Non cancer.

3.1.1. Kidney Transplantation

One of the main areas of research with the aim to identify urinary biomarkers has been the evaluation of kidney transplant-associated complications. Acute rejection is one of the key factors that determines long-term graft function and survival in renal transplant patients (57). SELDI was used by three independent laboratories to detect potential biomarkers for acute allograft rejection in kidney transplant patients (58-60). Clusters of urinary proteins correctly classified between 30 and 50 patients (depending on the study) with high sensitivity and specificity. However these results were only obtained on training sets and an independent validation on a separate cohort is still lacking. One follow-up study describes the identification of two proteins that were used in the above prediction and the use of one of them, β -Defensin-1 (a host defense peptide), in an immunoassay to predict acute transplant rejection (61). The use this single biomarker allowed the prediction of acute rejection, but with a significantly lower sensitivity and specificity. This further indicates that the use of

several urinary proteins or peptides yields higher diagnostic specificity and sensitivity. Although the three different laboratories all studied acute renal allograft rejection, completely different biomarkers were defined. This gives the impression that the results of SELDI are erratic, but can be explained by the use of different chip surfaces, progress in chip surface preparation and might also originate from different instrument settings such as the signal to noise ratio or mass calibration.

CE-MS was used on urinary samples from patients with different grades of subclinical or clinical acute transplant rejection, patients with urinary tract infection and patients without evidence of rejection or infection (62). Substantial differences were found between patients with transplanted kidneys and patients with native kidneys, most likely due to treatment with cyclosporin A, a calcineurin-inhibitor immunosuppressant. In addition, a distinct urinary polypeptide pattern identified 16 of the 17 patients with acute tubulointerstitial rejection. Potentially confounding variables, such as acute tubular lesions, tubular atrophy, tubulointerstitial fibrosis, calcineurin inhibitor toxicity, proteinuria, hematuria, allograft function, and different immunosuppressive regimens did not affect the results. To enable differentiation between infection and acute rejection, an additional biomarker pattern was developed. These polypeptide patterns were validated in a blinded assessment of nine acute rejection patients, seven patients with urinary tract infection and ten controls. Most samples were correctly classified using these biomarkers (62). This suggests that urinary proteome analysis can be used for the non-invasive monitoring of renal transplant patients although awaits validation in larger cohorts.

3.1.2. Chronic kidney disease

Chronic kidney disease is becoming a global health problem as the number of individuals with chronic kidney disease is steadily increasing. This is mainly due to the increased life-expectancy and the increasing incidence of type II diabetes (63). Early detection of chronic kidney disease is mandatory to reduce the number of patients requiring renal replacement therapy. Currently, chronic kidney disease is detected at a late stage when renal function has already significantly deteriorated, mainly due to the absence of non-invasive biomarkers. Therefore the selection of urinary polypeptide biomarkers for chronic kidney diseases is of utmost importance.

One of the first reports was the analysis of urinary polypeptide markers of membranous glomerulonephritis by SELDI and CE-MS (64). Using identical urine samples, three potential biomarkers were defined using SELDI analysis compared to 200 potential biomarkers from the CE-MS analysis. Additional work, using CE-MS, on urine samples from patients with other chronic renal diseases suggested that panels of 20 to 50 urinary polypeptide markers allow to discriminate (differential diagnosis) between different kidney diseases such as IgA nephropathy, focal-segmental glomerulosclerosis, membranous glomerulonephritis, minimal-change disease, and diabetic nephropathy (14, 16, 65). A recent study, using CE-MS on 3600 samples obtained from 20 different centers (Europe, America, and Australia), allowed the establishment of a database of more than 5000 urinary peptides (7). This database was used to define biomarkers of chronic kidney disease in general, but also for the differential diagnosis of for example focal and segmental glomerulosclerosis (FSGS) and membranous glomerulonephritis (MGN). The validation of these biomarkers in a blinded and independent heterogeneous cohort (as encountered in everyday life in the clinic) of 134 individuals allowed sorting out the 89 of the 101 chronic renal disease patients, yielding an 88% sensitivity and 100% specificity. Furthermore, the 3

patients with FSGS and 3 of the 4 patients with MGN among the 134 individuals were identified in this population (Good et al., submitted). This study shows the potential of urinary biomarkers to identify patients with chronic kidney disease in a heterogeneous clinical setting.

3.1.3. Diabetic nephropathy

In addition to disease-specific biomarkers, stage-specific urinary polypeptide markers can be defined. This will be exemplified below for the selection of urinary markers of diabetic nephropathy. Stage-specific biomarkers for diabetic nephropathy in patients with diabetes mellitus were defined (66, 67). In these two studies, the individual data sets of healthy volunteers (n=9 and 39, respectively), patients with diabetes type I or II without macroalbuminuria (n=28 and 46, respectively), and with intermittent or persistent macroalbuminuria (n=16 and 66, respectively) were combined to create typical polypeptide patterns. In patients with type II diabetes mellitus and a normal albumin excretion rate, the detected polypeptide pattern differed significantly from that in patients with advanced albuminuria. Comparable results were obtained for patients with diabetes type I. A recent study on a larger cohort, including 300 patients and controls (68), further confirms the initial findings on type I diabetic patients based on a standardized sample preparation protocol (7). This study defined, and in a blinded assessment subsequently validated, biomarkers for diabetes, diabetic nephropathy and biomarkers that enabled differentiation of diabetic nephropathy from other chronic renal diseases. In addition, these biomarkers could also be used to predict microalbuminuric patients at increased risk of progressing towards diabetic nephropathy over a 3 year period (68). The validity of this approach was subsequently confirmed using an independent set of samples from the Coronary

Artery Calcification in Type 1 (CACTI) Diabetes study with similar results for both, detection of diabetes and diabetes nephropathy (Snell Bergeon et al., manuscript submitted). These data indicate that the urinary biomarkers can be used not only for detection of diabetes and diabetic nephropathy, but also to predict disease progression. The use of urinary biomarkers in the prediction of disease progression is confirmed by the studies described in the next paragraph.

3.1.4. Obstructive nephropathy

Antenatal screening detects fetal hydronephrosis (dilation of the kidney due to urine accumulation) in around 1 out of 100 births with about 20% of the cases being clinically significant. Ureteropelvic junction (UPJ) obstruction (**Figure 2A**) is found in 40-50 % of these clinically significant cases (69). UPJ obstruction is thus a frequently encountered clinical situation. UPJ obstruction is functionally defined as a restriction to the urinary outflow that, when left untreated, will cause progressive renal deterioration. Alternatively, obstruction has been more generally defined as a condition that hampers optimal renal development (70). Since hydronephrosis is not always synonymous with obstruction, the differentiation between a dilated obstructed and dilated non obstructed kidney is often a difficult problem (**Figure 2A**). No reference standards are available to correctly identify obstruction. Further, diagnosis, based on arbitrary threshold values (in the absence of reference standards), is usually achieved through repeating the various radiologic investigations. These radiologic investigations expose these infants to radiation and may need injection of radiocontrast or radioisotope material. We have studied the urinary proteome of the newborns with UPJ obstruction to identify biomarkers of obstruction that can be used to predict whether a neonate with UPJ obstruction evolves towards spontaneous

resolution or surgery (**Figure 2A**, (71, 72)). We used CE-MS-based urinary proteome analysis to define specific biomarker patterns for different grades of ureteropelvic junction obstruction. In a blinded prospective study on 36 UPJ obstruction patients, these patterns predicted with 95% accuracy the clinical outcome of the newborns nine months in advance (**Figure 2B**, (72)). After 15 months of follow up, the accuracy of the prediction increased to 97% as one of the newborns with UPJ obstruction had to be operated at a late stage, as predicted by the urinary proteome analysis (**Figure 2B**, (71)). A multi-center prospective study on 358 UPJ patients is ongoing for validation.

These data and the recent study on the urinary proteome-based prediction of the progression of microalbuminuric diabetic patients (68) strongly suggest the possibility to predict the progression of disease by urinary proteome analysis.

3.2. Urogenital disease: Cancer.

3.2.1. Renal cell carcinoma

One of the first applications of urinary proteome analysis for a clinically relevant question aimed to define renal cell carcinoma (RCC)-specific biomarkers (73). Samples of 218 individuals were analyzed by SELDI-TOF. Samples from patients before nephrectomy for RCC (n=48), normal healthy volunteers (n=38), and outpatients with benign diseases of the urogenital tract (n=20) were used as a training set for biomarker definition. The defined markers were subsequently validated in two blinded assessments with an initial "blind" group of 32 samples (12 patients with RCC, 11 healthy controls, and 9 patients as disease controls) and a second group of 80 samples (36 patients with RCC, 31 healthy volunteers, and 13 patients with benign urological conditions). While in the first round sensitivities and

specificities of 81.8-83.3% were achieved, the values significantly declined, ranging from 41.0% to 76.6%, for the second set of samples collected 10 months later. The authors analyzed possible contributing factors including sample stability, changing laser performance, and chip variability to assess a long-term robustness of the approach. One of the main conclusions from this study was the need for rigorous evaluation of such variables that may influence stability/robustness.

3.2.2. Bladder cancer

Bladder cancer (BCa) is among the five most common malignancies worldwide. Urothelial (transitional cell) carcinoma (TCC) constitutes 95% of all these BCa cases in the Western countries. 80% of the BCa patients have superficial carcinomas that can be treated, but these patients must be closely screened for reoccurrence. This requires cytological examination of urine which lacks sensitivity especially for lower stage tumors. Cystoscopy is more sensitive, but invasive. This underscores the need for novel, non-invasive, biomarkers of BCa. A number of studies have been performed with the aim to identify urinary biomarkers of BCa. Many potential biomarkers were identified including psoriasin (S100A7, (74)), metalloproteinases MMP-2, -9, fibronectin (75), orosomucoid and zinc- α_2 -glycoprotein (76), but without subsequent validation of these biomarkers in prospective studies. These studies on biomarkers for BCa will therefore not be described in more detail. However some examples of identification *and* validation in prospective studies of urinary biomarkers of BCa by proteomic analysis were published, these are outlined below.

One of the first studies using 2D-DIGE for the analysis of the urinary proteome aimed to identify biomarkers of BCa (41). 2D-DIGE was used to analyze 7 different sets of patients and healthy controls yielding 12 clearly differentially expressed spots.

One of the differentially expressed proteins was regenerating protein-1 (Reg-1). Reg-1 is proposed to act as an inhibitor of apoptosis leading to Reg-1 activated proliferative activity. Reg-1 expression in BCa biopsies was found to be associated with tumor progression and clinical outcome. In the next step an immunoassay was developed to study Reg-1 expression in urine. In a prospective analysis on 80 individuals, containing 32 BCa patients (stage Ta to T2), this Reg-1 immunoassay allowed to discriminate between BCa patients and healthy controls with a specificity and sensitivity of 81.3% of 81.3%, respectively (41).

SELDI-TOF profiling was used by several laboratories in detecting BCa in blinded sets of samples: sensitivity ranged from 71.7%-93.3% and specificity from 62.5%-87% (77) to discriminate BCa patients from healthy controls. As described above, in SELDI-TOF comparability of the datasets is not easy to achieve due to differences in chip surfaces and conditions in the different studies.

CE-MS was also used for the detection and validation of biomarkers of TCC (11). A BCa-specific biomarker pattern was established by initial definition in a training set composed of 46 patients with TCC and 33 healthy subjects and further refinement using CE-MS spectra of 366 urine samples from healthy volunteers and patients with malignant and non-malignant genitourinary diseases. By this two-step biomarker discovery approach, the authors were able to establish a prediction model composed of 22 urinary peptides, which, when applied to a blinded test set containing 31 TCC patients, 11 healthy individuals and 138 non-malignant genitourinary disease patients, correctly classified all TCC patients and all healthy controls. Differentiation between bladder cancer and other malignant and non-malignant diseases (such as renal nephrolithiasis) was accomplished with at least 86% - 100% sensitivity.

Urinary proteome studies allowed to identify biomarkers that distinguished between BCa and controls in prospective studies with variable sensitivities and specificities. This is promising. The next challenge is to find biomarkers that can predict tumor stage, recurrence, progression, and treatment response in patients with BCa.

3.2.3. Prostate cancer

In a pilot study (78), CE-MS was used to define potential urinary peptide biomarkers for prostate cancer (PCa). Urine samples from 47 patients who underwent prostate biopsy were analyzed. On the basis of prostate biopsy, 26 patients in this group were diagnosed as having PCa and 21 as having benign prostatic hyperplasia (BPH). The data indicated several polypeptides allowing prediction of PCa with 92% sensitivity and 96% specificity. However, these data could not be validated in a larger cohort (Mischak, unpublished), once more underlining the importance of validation in a blinded, independent test set. Upon more thorough analysis, first-void urine was found to contain potentially useful biomarkers for PCa, while the generally used mid-stream urine appeared not to contain significant PCa-related information (79). These results indicate that the biomarkers originate from secretions of the prostate into the urine, and also underline the importance of accurate sampling. After refinement of the PCa-specific biomarker pattern using urine samples from 54 PCa and 35 BPH patients, a model with 12 potential biomarkers resulted in the correct classification of 89% of the PCa and of 51% of the BPH patients in a second blinded cohort of 213 patient samples (79). Inclusion of age and free PSA increased the sensitivity and specificity to 91 and 69%, respectively.

3.3. Application of urinary proteome analysis to non-urogenital diseases

It has been estimated that 30% of the urinary proteins does not originate from the urogenital tract (**Figure 1**) and the first studies showing the identification and validation of urinary markers for other than urogenital diseases is emerging.

A first example is the clinical follow-up of patients after allogeneic hematopoietic stem cell transplantation (HSCT) (80) (81). Urine samples from 40 patients after HSCT (35 allogeneic, 5 autologous) and 5 patients with sepsis were collected during a period of 100 days (a maximum of 10 samples per patient) for CE-MS analysis. A pattern consisting of 16 differentially excreted polypeptides indicated early graft-versus-host-disease (GVHD), enabling discrimination of patients with early GVHD from patients without complications with 82% specificity and 100% sensitivity. A subsequent blinded multicenter validation study of 100 patients with more than 600 samples collected prospectively confirmed the results, although with reduced specificity and sensitivity (82). First, preliminary data on patients that received preemptive therapy of GvHD based on urinary proteome analysis indicate a clear benefit: reduction of both, occurrence of GvHD and lethality (Weissinger et al., unpublished). These preliminary data are currently further substantiated in a multicenter prospective trial.

Zimmerli et al. were able to define and validate biomarkers for coronary artery disease (CAD) in urine (83). Urine from 88 CAD patients and 282 controls was examined by CE-MS. This resulted in the identification of 15 peptides that defined a characteristic CAD signature panel. In a second step this panel was evaluated in a blinded study on 47 CAD patients and 12 healthy individuals. CAD patients were identified with greater than 90% sensitivity and specificity. In addition, the polypeptide CAD signature panel significantly changed towards the healthy polypeptide signature

after therapeutic intervention. These data were further substantiated in a study by von zur Muehlen et al., (submitted), where patients with CAD could be distinguished from patients presenting symptoms of CAD, but without clinical evidence in the coronary angiography. The prospective value of the data could further be validated in prospectively collected samples from patients with type I diabetes (Snell-Bergeon et al., manuscript submitted). In this blinded study the value of urinary proteome analysis in the prediction of future CAD events could be demonstrated.

Although still limited in number, these examples show that urine can also be a source of biomarkers for more distant organs.

4. From biomarkers to pathophysiology

The field of urinary proteomics has advanced and is now entering the era of validation of the selected urinary biomarkers for a number of urogenital and systemic pathologies. Most of us also entered this biomarker research with the hope to find clues to better understand the pathophysiology of disease: this necessitates identification of the biomarkers. As described above, urinary protein profiling enables selection and validation of biomarkers for disease. But the identification of these biomarkers remains challenging, especially of proteins >10 kDa. Nevertheless, several studies reported a number of sequenced biomarkers. While we refrain from listing the different proteins or peptides identified for each specific disease discovered by urinary proteome analysis, we aim towards discussing the major conclusions drawn from biomarker sequences.

Most of the currently identified urinary biomarkers for disease are i) abundant plasma proteins or fragments thereof (*i.e.* albumin, β 2-macroglobulin, α 1 antitrypsin, etc...) due to leakage in the pathological state, and ii) abundant kidney and structural proteins (*i.e.* collagens, uromodullin) (84). These proteins or peptides were identified using various mass spectrometry based proteomic techniques. Although useful as a biomarker, these abundant urinary proteins are at first sight of little information on the underlying pathology. However, the specific fragments of these abundant proteins might give some clues on the underlying physiopathology of disease (7). For example, renal disease without albuminuria still exhibits disease-specific changes in urinary polypeptides (16), including specific fragments of albumine (85). This strongly suggests that these peptides contain clues about the pathogenesis and are not simple degradation products of abundant urinary proteins. It is tempting to speculate that the disease-specific peptides may be indirect indicators of the activity of disease-

specific proteases (65). This hypothesis is further strengthened by work (86), in which the presence of specific collagen fragments correlated with the disease-specific activity of matrix metalloproteases. Moreover, a similar process has been described in the case of some cancer biomarkers identified in plasma, shown to be fragments of abundant plasma proteins specifically cleaved by proteases released from cancer cells (87). While the evidence is still scarce, it is an attractive hypothesis that urinary peptides of diagnostic value are not merely degradation products of abundant larger proteins, but a result of distinct, disease-specific processes, in many cases due to significant changes in the activity of proteases. This assumption is supported by sometimes apparently unrelated findings; for example, the increase of collagen and extracellular matrix in patients with diabetes and DN has been established by a variety of methods. Our recent findings that collagen fragments are significantly reduced in diabetic urine (68) fit in this scenario and further supports the hypothesis that both, reduced activity of proteases and protection of the extracellular matrix from proteolysis by advanced glycosylation end products may be key pathological changes in diabetes mellitus (84). A similar scenario may be applicable to albuminuria. Consequently, an albumin-derived biomarker is not simply “an albumin fragment”, but rather a specific fragment, defined by its specific N- and C-terminus. Similar observations, the presence of specific urinary fragments of albumin and alpha-1-antitrypsin associated with nephrotic syndrome in chronic kidney disease has recently also been described (88). Unfortunately, such essential detailed information about protein processing by proteases is difficult to obtain, both from the “protein side”, as nanoLC-MS/MS approaches often identify proteins based on the sequencing of a few tryptic peptides which do not necessarily map the cleavage sites (6), and from the “peptide side”, as profiling approaches like CE-MS and SELDI-TOF

does not provide the sequence of the detected biomarker peptides. A thorough examination of the sequences of the urinary peptides and comparison with protease specificities may strengthen the above hypothesis and leads to a better insight into regulation and pathophysiological role of specific proteases in many diseases.

Another attractive hypothesis is that the urinary peptidome displays to a large degree the turnover of the extracellular matrix. This hypothesis has been generated as a result of the observation that the most abundant urinary peptides (based on ion counting) are not, as expected, the “usual suspects” like albumin or uromodulin, but specific collagen degradation products (7). Further, several distinct collagen peptides are significantly reduced in diseases where an increase of ECM has been reported (84). Consequently, these peptides may be derived from ECM turnover. Changes in this turnover also result in indicative changes in urinary peptides, which may serve as very specific indicator for such a change, which in turn is likely to be disease specific. Such changes in the ECM turnover may be due to e.g. invasion of tumors (ECM needs to be “dissolved in order to make room for the growing tumor), fibrosis (reduced ECM degradation), increased arterial stiffness (change in ECM composition), changes in endothelium, etc... Therefore mapping of the collagen cleavage sites might incriminate specific proteases in ECM turnover not previously identified and define their activities under pathophysiological conditions.

5. Chasing low abundance urinary proteins

The dynamic range of protein concentrations in body fluids often spans several orders of magnitude (35, 89, 90). A major challenge is thus to identify low abundant components in complex protein mixtures with high dynamic range. Immunodepletion of abundant urinary proteins might help to unmask the low abundance proteins as was shown by 2D-DIGE analysis of urine of patients with diabetic nephropathy (40). A major drawback of such immunosubtraction method appears to be co-depletion. For example, depletion of plasma for human serum albumin, co-depleted another 815 species (not including albumin) (91). When capturing IgGs, another 2091 species (not including IgG) were co-depleted. These IgG co-depleted proteins contained 56% sequences coding for antibodies and 44% of low abundance cytokines or related proteins. Interestingly, fewer proteins could be detected in albumin- and IgG-depleted plasma sample, than in the samples destined to be discarded (91).

Recent developments might help to uncover the underexplored urinary proteome. A novel and very efficient approach has been described for capturing the "hidden proteome", rare proteins that constitute the vast majority in any cell or tissue lysate and in biological fluids (92-94). It is based on a combinatorial library of hexameric peptide ligands bound to porous polyacrylate beads named "Proteominer" (formerly called "equalizer beads", **Figure 3**). Each bead contains billions of copies of a unique hexapeptide ligand distributed throughout its porous structure, and each bead potentially has a different ligand from every other bead. With a population of millions of individual peptide ligands obtained by combinatorial chemistry, any protein present in the starting material could theoretically interact with one or a few particular beads. Once the most abundant protein species have saturated their binding sites,

the remaining molecules are washed away in the flowthrough, while minor protein species get progressively enriched on their corresponding beads. Thus, instead of simplifying the complex mixture into fractions or partitioning away the most abundant proteins, this approach captures the species present in solution up to the saturation of the solid phase ligand library. The protein mixture is thus “equalized” and the dynamic range of protein concentrations strongly reduced. This ligand library has been efficiently used for capturing and revealing a very large population of previously undetected proteins from serum (95), platelets (96), or red blood cells (Roux-Dalvai et al, unpublished results). It has also been applied to urine (27), and analysis of the sample by both 1D gels, 2D gels and SELDI-TOF revealed that the treatment induced a strong decrease in the levels of the most abundant proteins, notably albumin, while it allowed detection of numerous previously undetected species. Moreover, nanoLC-MS/MS analysis of the treated urine by high-resolution, fast-sequencing Fourier Transform mass spectrometry resulted in the identification of more than 300 protein species in only one analytical run of about 1 hour, to be compared to identification of 134 proteins in non-treated urine. Thus, application of the Proteominer technology may allow extending protein profiling towards lower abundant species. However, treatment with peptide ligand libraries will modify the abundances of proteins in the treated sample. It needs yet to be assessed if this approach can be used for differential proteomic studies, *i.e.* if proteins found differentially expressed in samples to be compared are still found with the same differential expression ratio in untreated samples. Test experiments performed by spiking standard proteins in cell lysates indicated that, if the protein does not saturate the beads, the relative quantitative information is well conserved and the method is reproducible (Roux-Dalvai et al, unpublished results). Although further validation is

needed, this technology may represent a useful way to detect and quantify low abundant proteins in urine.

6. Conclusions

Urinary proteome analysis is emerging as a powerful diagnostic and prognostic tool not only in kidney disease, but also in diseases of more distant organs. While urinary proteome analysis is far from becoming a routine tool in the clinical setting, studies on larger cohorts of patients reveal its potential in clinical diagnosis. Efforts have to be made to validate these panels of biomarkers on even larger- and, probably more importantly, heterogeneous-cohorts to move away from the bench paradigm “disease versus control”, to the bedside.

The contribution of urinary proteomics to the understanding of the pathophysiology of disease upon analysis of the urinary proteome is still modest, however. The evolution of mass spectrometers toward high mass accuracy and resolution, new ways to explore low-abundance proteins and peptides, and new bioinformatics tools should help to sequence more biomarkers in the near future and thus learn more about the pathophysiology of the underlying disease.

Acknowledgements

The work of JPS was supported by Inserm, the “Direction Régional Clinique” (CHU de Toulouse, France) under the Interface program and by the Fondation pour la Recherche Médicale. The work of SD was sponsored by the Inserm Interface program. The work of BM and AGP was supported by the CNRS and the Génopole Toulouse Midi-Pyrénées. SD, BM, AGP and JPS acknowledge financial support from the Agence Nationale pour la Recherche (ANR-07-PHYSIO-004-01), the Fondation pour la Recherche Médicale “Grands Equipements pour la Recherche Biomédicale” and the CPER2007-2013 programme. HM was supported in part by EUROTRANS-BIO grant ETB-2006-016 and EU Funding through InGenious HyperCare (LSHM-C7-2006-037093) and PREDICTIONS (1272568).

Table : advantages and disadvantages of each mass spectrometry-based proteomic technique for use clinical applications.

Technology	Advantages	Disadvantages
SELDI-TOF	Easy-to-use, high throughput, automation, low sample volume	Restricted to selected proteins, low resolution MS, lack of comparability, sensitive toward interfering compounds.
CE-MS	Automation, high sensitivity, fast, low sample volume, multidimensional	Generally not suited for larger molecules (>20kDa)
LC-MS	Automation, multidimensional, high sensitivity, used for detection of large molecules (>20kDa) after tryptic digest, sequence determination of biomarkers provided by MS/MS	Reassembly of tryptic peptides into their precursor molecule can be problematic, time consuming, relatively sensitive toward interfering compounds, medium throughput.
2DE-MS	Detection of large molecules, Enables estimation of actual molecular weight of large molecules, sequencing of biomarkers easy to perform from 2D spots	Not applicable to molecules <10kDa, no automation, time consuming, medium throughput. moderate comparability

Figures

Figure 1: 70% of the urinary proteins and peptides originate for the kidney and the urinary tract while the remaining 30% originates for the circulation (3). About 900 liters of plasma flow through the kidneys each day. 150-180 liters of this plasma is filtered by the glomeruli (the “primitive urine”). However, more than 99% of this primitive urine is reabsorbed by the renal tubule. The remainder (“final urine”) exits the kidney *via* the ureter into the bladder.

Figure 2: A) UPJ obstruction, a stenosis at the intersection of the ureter and the pelvis of the kidney, induces urine accumulation in the kidney leading to hydronephrosis. A low number of newborns with severe UPJ obstruction need to be rapidly operated after birth while another, larger, group with low level obstruction evolves to spontaneous resolution of the obstruction. However the remaining group of neonates with “intermediate” obstruction needs medical surveillance to determine the evolution of the pathology. The common clinical practice to decide whether infants should undergo surgery depends on radiologic investigations, involving exposure of the newborns to radiation. **B)** Urinary proteomics has shown to be efficiently predict the progression of UPJ patients (71, 72). Urinary protein profiles from 36 patients from the “intermediate” UPJ obstruction group were classified using a hierarchic disease model based on the discriminating polypeptides between the healthy newborn, mild (spontaneous resolution) and severe (surgery) UPJ-obstruction. Each patient was scored with this model using support vector machines. This results in membership values between -1 and 1. A negative value suggests evolution towards surgery and a positive value suggests evolution towards the

spontaneous resolution of UPJ obstruction. This prediction was compared with the clinical evolution after 9 and 15 months of follow-up. This resulted in 34 out of 36 correct predictions (94%) at 9 months and 97% (35/36) at 15 months as one patient evolved to severe UPJ obstruction at a late stage (arrow). This figure was partly reproduced from Decramer et al. 2006, *Nat Med* **12**:398-400, with permission.

Figure 3: Sample « equalization » using ProteoMiner beads. Schematic representation of the reduction of the dynamic concentration range of a protein mixture, using combinatorial ligand libraries. Each bead (grey sphere) carries a single type of ligand and interacts with one protein species (blue, red, green, purple). Proteins in excess (blue and green), rapidly saturate the binding sites of the corresponding beads; they remain thus uncaptured and are eliminated during bead wash. Captured proteins are finally desorbed from the beads and collected.

References

1. Iorio, L., and Avagliano, F. (1999) Observations on the Liber medicine orinalibus by Hermogenes. *Am J Nephrol* 19, 185-188.
2. Moe, O. W., Berry, C. A., and Rector, F. C. (2000) *The Kidney*, W.B. Saunders, Philadelphia, USA.
3. Thongboonkerd, V., and Malasit, P. (2005) Renal and urinary proteomics: current applications and challenges. *Proteomics* 5, 1033-1042.
4. 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.
5. 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.
6. 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 membrane proteins. *Genome Biol* 7, R80.
7. Coon, J. J., Zurbig, P., Dakna, M., Dominiczak, A. F., Decramer, S., Fliser, D., Frommberger, M., Golovko, I., Good, D. M., Herget-Rosenthal, S., Jankowski, J., Julian, B. A., Kellmann, M., Kolch, W., Massy, Z., Novak, J., Rossing, K., Schanstra, J. P., Schiffer, E., Theodorescu, D., Vanholder, R., Weissinger, E. M., Mischak, H., and Schmitt-Kopplin, P. (2008) CE-MS analysis of the human urinary proteome for biomarker discovery and disease diagnostics. *Proteomics Clin Appl*, in press
8. Omenn, G. S., States, D. J., Adamski, M., Blackwell, T. W., Menon, R., Hermjakob, H., Apweiler, R., Haab, B. B., Simpson, R. J., Eddes, J. S., Kapp, E. A., Moritz, R. L., Chan, D. W., Rai, A. J., Admon, A., Aebersold, R., Eng, J., Hancock, W. S., Hefta, S. A., Meyer, H., Paik, Y. K., Yoo, J. S., Ping, P., Pounds, J., Adkins, J., Qian, X., Wang, R., Wasinger, V., Wu, C. Y., Zhao, X., Zeng, R., Archakov, A., Tsugita, A., Beer, I., Pandey, A., Pisano, M., Andrews, P., Tammen, H., Speicher, D. W., and Hanash, S. M. (2005) Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* 5, 3226-3245.
9. Kolch, W., Neuss, 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.
10. Schaub, S., Wilkins, J., Weiler, T., Sangster, K., Rush, D., and Nickerson, P. (2004) Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. *Kidney Int* 65, 323-332.
11. 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.
12. Zhou, H., Yuen, P. S., Pisitkun, T., Gonzales, P. A., Yasuda, H., Dear, J. W., Gross, P., Knepper, M. A., and Star, R. A. (2006) Collection, storage,

- preservation, and normalization of human urinary exosomes for biomarker discovery. *Kidney Int* 69, 1471-1476.
13. Vestergaard, P., and Leverett, R. (1958) Constancy of urinary creatinine excretion. *J Lab Clin Med* 51, 211-218.
 14. 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.
 15. Fliser, D., Wittke, S., and Mischak, H. (2005) Capillary electrophoresis coupled to mass spectrometry for clinical diagnostic purposes. *Electrophoresis* 26, 2708-2716.
 16. 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.
 17. Thongboonkerd, V. (2007) Recent progress in urinary proteomics. *Proteomics Clin Appl* 1, 780-791.
 18. Fliser, D., Novak, J., Thongboonkerd, V., Argiles, 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.
 19. Poon, T. C. (2007) Opportunities and limitations of SELDI-TOF-MS in biomedical research: practical advices. *Expert Rev Proteomics* 4, 51-65.
 20. Check, E. (2004) Proteomics and cancer: running before we can walk? *Nature* 429, 496-497.
 21. Kiehnopf, M., Siegmund, R., and Deufel, T. (2007) Use of SELDI-TOF mass spectrometry for identification of new biomarkers: potential and limitations. *Clin Chem Lab Med* 45, 1435-1449.
 22. Klampfl, C. W. (2006) Recent advances in the application of capillary electrophoresis with mass spectrometric detection. *Electrophoresis* 27, 3-34.
 23. Wittke, S., Mischak, H., Walden, M., Kolch, W., Radler, T., and Wiedemann, K. (2005) Discovery of biomarkers in human urine and cerebrospinal fluid by capillary electrophoresis coupled to mass spectrometry: towards new diagnostic and therapeutic approaches. *Electrophoresis* 26, 1476-1487.
 24. Zurbig, P., Renfrow, M. B., Schiffer, E., Novak, J., Walden, M., Wittke, S., Just, I., Pelzing, M., Neususs, 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.
 25. Sun, W., Li, F., Wu, S., Wang, X., Zheng, D., Wang, J., and Gao, Y. (2005) Human urine proteome analysis by three separation approaches. *Proteomics* 5, 4994-5001.
 26. Wang, L., Li, F., Sun, W., Wu, S., Wang, X., Zhang, L., Zheng, D., Wang, J., and Gao, Y. (2006) Concanavalin A-captured glycoproteins in healthy human urine. *Mol Cell Proteomics* 5, 560-562.
 27. 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.
 28. Ru, Q. C., Katenhusen, R. A., Zhu, L. A., Silberman, J., Yang, S., Orchard, T. J., Brzeski, H., Liebman, M., and Ellsworth, D. L. (2006) Proteomic profiling of human urine using multi-dimensional protein identification technology. *J Chromatogr A* 1111, 166-174.

29. de Godoy, L. M., Olsen, J. V., de Souza, G. A., Li, G., Mortensen, P., and Mann, M. (2006) Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system. *Genome Biol* 7, R50.
30. Mueller, M., Martens, L., and Apweiler, R. (2007) Annotating the human proteome: beyond establishing a parts list. *Biochim Biophys Acta* 1774, 175-191.
31. Mueller, M., Vizcaino, J. A., Jones, P., Cote, R., Thorneycroft, D., Apweiler, R., Hermjakob, H., and Martens, L. (2008) Analysis of the experimental detection of central nervous system-related genes in human brain and cerebrospinal fluid datasets. *Proteomics* 8, 1138-1148.
32. Listgarten, J., and Emili, A. (2005) Statistical and computational methods for comparative proteomic profiling using liquid chromatography-tandem mass spectrometry. *Mol Cell Proteomics* 4, 419-434.
33. Qian, W. J., Jacobs, J. M., Liu, T., Camp, D. G., 2nd, and Smith, R. D. (2006) Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Mol Cell Proteomics* 5, 1727-1744.
34. Page, J. S., Masselon, C. D., and Smith, R. D. (2004) FTICR mass spectrometry for qualitative and quantitative bioanalyses. *Curr Opin Biotechnol* 15, 3-11.
35. Rifai, N., Gillette, M. A., and Carr, S. A. (2006) Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol* 24, 971-983.
36. Domon, B., and Aebersold, R. (2006) Challenges and opportunities in proteomics data analysis. *Mol Cell Proteomics* 5, 1921-1926.
37. Keshishian, H., Addona, T., Burgess, M., Kuhn, E., and Carr, S. A. (2007) Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics* 6, 2212-2229.
38. Thongboonkerd, V. (2007) Practical Points in Urinary Proteomics. *J Proteome Res* 6, 3881-3890.
39. Hoorn, E. J., Hoffert, J. D., and Knepper, M. A. (2006) The application of DIGE-based proteomics to renal physiology. *Nephron Physiol* 104, p61-72.
40. Rao, P. V., Lu, X., Standley, M., Pattee, P., Neelima, G., Girisesh, G., Dakshinamurthy, K. V., Roberts, C. T., Jr., and Nagalla, S. R. (2007) Proteomic identification of urinary biomarkers of diabetic nephropathy. *Diabetes Care* 30, 629-637.
41. Orenes-Pinero, E., Corton, M., Gonzalez-Peramato, P., Algaba, F., Casal, I., Serrano, A., and Sanchez-Carbayo, M. (2007) Searching urinary tumor markers for bladder cancer using a two-dimensional differential gel electrophoresis (2D-DIGE) approach. *J Proteome Res* 6, 4440-4448.
42. 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.
43. Mischak, H. A., R.; Banks, R. E.; Conaway, M.; Coon, J. J.; Dominiczak, 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.; Yamamoto, T. (2007) Clinical proteomics: A need to define the field and to begin to set standards *Proteomics Clin Appl* 1, 148-156.

44. Yamamoto, T., Langham, R. G., Ronco, P., Knepper, M. A., and Thongboonkerd, V. (2008) Towards standard protocols and guidelines for urine proteomics: a report on the Human Kidney and Urine Proteome Project (HKUPP) symposium and workshop, 6 October 2007, Seoul, Korea and 1 November 2007, San Francisco, CA, USA. *Proteomics* 8, 2156-2159.
45. Vlahou, A., Schanstra, J., Frokiaer, J., El Nahas, M., Spasovski, G., Mischak, H., Domon, B., Allmaier, G., Bongcam-Rudloff, E., and Attwood, T. (2008) Establishment of a European Network for Urine and Kidney Proteomics. *J Proteomics* in press
46. Zerefos, P. G., and Vlahou, A. (2008) Urine sample preparation and protein profiling by two-dimensional electrophoresis and matrix-assisted laser desorption ionization time of flight mass spectroscopy. *Methods Mol Biol* 428, 141-157.
47. Fiedler, G. M., Baumann, S., Leichtle, A., Oltmann, A., Kase, J., Thiery, J., and Ceglarek, U. (2007) Standardized peptidome profiling of human urine by magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Clin Chem* 53, 421-428.
48. Thongboonkerd, V., and Saetun, P. (2007) Bacterial overgrowth affects urinary proteome analysis: recommendation for centrifugation, temperature, duration, and the use of preservatives during sample collection. *J Proteome Res* 6, 4173-4181.
49. Abdi, H. (2007) *Bonferroni and Sidak corrections for multiple comparisons*, Sage, Thousand Oaks, USA.
50. Westfall, P. H., and Young, S. S. (1993) *Resampling-based Multiple Testing: Examples and Methods for P-Value Adjustment.*, Wiley, New York.
51. Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society (Series B: Methodological)* 57, 289-300.
52. Reiner, A., Yekutieli, D., and Benjamini, Y. (2003) Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* 19, 368-375.
53. 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.
54. Hernandez, J., and Thompson, I. M. (2004) Prostate-specific antigen: a review of the validation of the most commonly used cancer biomarker. *Cancer* 101, 894-904.
55. Mogensen, C. E. (1994) Systemic blood pressure and glomerular leakage with particular reference to diabetes and hypertension. *J Intern Med* 235, 297-316.
56. Rossing, K. (2007) Progression and remission of nephropathy in type 2 diabetes: new strategies of treatment and monitoring. *Dan Med Bull* 54, 79-98.
57. Gwinner, W. (2007) Renal transplant rejection markers. *World J Urol* 25, 445-455.
58. Clarke, W., Silverman, B. C., Zhang, Z., Chan, D. W., Klein, A. S., and Molmenti, E. P. (2003) Characterization of renal allograft rejection by urinary proteomic analysis. *Ann Surg* 237, 660-664; discussion 664-665.
59. 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.
60. O'Riordan, E., Orlova, T. N., Mei, J. J., Butt, K., Chander, P. M., Rahman, S., Mya, M., Hu, R., Momin, J., Eng, E. W., Hampel, D. J., Hartman, B., Kretzler, M.,

- Delaney, V., and Goligorsky, M. S. (2004) Bioinformatic analysis of the urine proteome of acute allograft rejection. *J Am Soc Nephrol* 15, 3240-3248.
61. O'Riordan, E., Orlova, T. N., Podust, V. N., Chander, P. N., Yanagi, S., Nakazato, M., Hu, R., Butt, K., Delaney, V., and Goligorsky, M. S. (2007) Characterization of urinary peptide biomarkers of acute rejection in renal allografts. *Am J Transplant* 7, 930-940.
 62. Wittke, S., Haubitz, M., Walden, M., Rohde, F., Schwarz, A., Mengel, M., Mischak, H., Haller, H., and Gwinner, W. (2005) Detection of acute tubulointerstitial rejection by proteomic analysis of urinary samples in renal transplant recipients. *Am J Transplant* 5, 2479-2488.
 63. Meguid El Nahas, A., and Bello, A. K. (2005) Chronic kidney disease: the global challenge. *Lancet* 365, 331-340.
 64. 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.
 65. 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.
 66. Meier, M., Kaiser, T., Herrmann, A., Knueppel, S., Hillmann, M., Koester, P., Danne, T., Haller, H., Fliser, D., and Mischak, H. (2005) Identification of urinary protein pattern in type 1 diabetic adolescents with early diabetic nephropathy by a novel combined proteome analysis. *J Diabetes Complications* 19, 223-232.
 67. 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.
 68. Rossing, K., Mischak, H., Dakna, M., Zürgbig, P., Novak, J., Julian, B. A., Good, D. M., Coon, J. J., Tarnow, L., and Rossing, P. (2008) Proteomic discovery and evaluation of urinary biomarkers for diabetes and chronic renal disease. *J Am Soc Nephrol*, in press.
 69. Chang, C. P., McDill, B. W., Neilson, J. R., Joist, H. E., Epstein, J. A., Crabtree, G. R., and Chen, F. (2004) Calcineurin is required in urinary tract mesenchyme for the development of the pyeloureteral peristaltic machinery. *J Clin Invest* 113, 1051-1058.
 70. Csaicsich, D., Greenbaum, L. A., and Aufricht, C. (2004) Upper urinary tract: when is obstruction obstruction? *Curr Opin Urol* 14, 213-217.
 71. Decramer, S., Bascands, J. L., and Schanstra, J. P. (2007) Non-invasive markers of ureteropelvic junction obstruction. *World J Urol* 25, 457-465.
 72. Decramer, S., Wittke, S., Mischak, H., Zurbig, 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.
 73. Rogers, M. A., Clarke, P., Noble, J., Munro, N. P., Paul, A., Selby, P. J., and Banks, R. E. (2003) Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural-network analysis: identification of key issues affecting potential clinical utility. *Cancer Res* 63, 6971-6983.

74. Celis, J. E., Rasmussen, H. H., Vorum, H., Madsen, P., Honore, B., Wolf, H., and Orntoft, T. F. (1996) Bladder squamous cell carcinomas express psoriasin and externalize it to the urine. *J Urol* 155, 2105-2112.
75. Saito, M., Kimoto, M., Araki, T., Shimada, Y., Fujii, R., Oofusa, K., Hide, M., Usui, T., and Yoshizato, K. (2005) Proteome analysis of gelatin-bound urinary proteins from patients with bladder cancers. *Eur Urol* 48, 865-871.
76. Irmak, S., Tilki, D., Heukeshoven, J., Oliveira-Ferrer, L., Friedrich, M., Huland, H., and Ergun, S. (2005) Stage-dependent increase of orosomucoid and zinc-alpha2-glycoprotein in urinary bladder cancer. *Proteomics* 5, 4296-4304.
77. Schiffer, E., Mischak, H., Theodorescu, D., and Vlahou, A. (2008) Challenges of using mass spectrometry as a bladder cancer biomarker discovery platform. *World J Urol* 26, 67-74.
78. Theodorescu, D., Fliser, D., Wittke, S., Mischak, H., Krebs, R., Walden, M., Ross, M., Eltze, E., Bettendorf, O., Wulfig, 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.
79. Theodorescu, D., Eric Schiffer, E., Bauer, H. W., Douwes, F., Eichhorn, F., Polley, R., Schmidt, T., Schöfer, W., Zürgb, P., Good, D. M., Coon, J. J., and Mischak, H. (2008) Discovery and validation of urinary biomarkers for prostate cancer. *Proteomics Clin Appl* in press
80. Kaiser, T., Kamal, H., Rank, A., Kolb, H. J., Holler, E., Ganser, A., Hertenstein, B., Mischak, H., and Weissinger, E. M. (2004) Proteomics applied to the clinical follow-up of patients after allogeneic hematopoietic stem cell transplantation. *Blood* 104, 340-349.
81. Weissinger, E. M., Hertenstein, B., Mischak, H., and Ganser, A. (2005) Online coupling of capillary electrophoresis with mass spectrometry for the identification of biomarkers for clinical diagnosis. *Expert Rev Proteomics* 2, 639-647.
82. Weissinger, E. M., Schiffer, E., Hertenstein, B., Ferrara, J. L., Holler, E., Stadler, M., Kolb, H. J., Zander, A., Zurbig, 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.
83. Zimmerli, L. U., Schiffer, E., Zurbig, P., Good, D. M., Kellmann, M., Mouls, L., Pitt, A. R., Coon, J. J., Schmieder, R. E., Peter, K., Mischak, H., Kolch, W., Delles, C., and Dominiczak, A. F. (2007) Urinary proteomic biomarkers in coronary artery disease. *Mol Cell Proteomics* 7, 290-298.
84. 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 - clinical applications*, in press.
85. 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.
86. Nemirovskiy, O. V., Dufield, D. R., Sunyer, T., Aggarwal, P., Welsch, D. J., and Mathews, W. R. (2007) Discovery and development of a type II collagen neoepitope (TIINE) biomarker for matrix metalloproteinase activity: from in vitro to in vivo. *Anal Biochem* 361, 93-101.
87. Bachmeier, B. E., Nerlich, A. G., Lichtinghagen, R., and Sommerhoff, C. P. (2001) Matrix metalloproteinases (MMPs) in breast cancer cell lines of different tumorigenicity. *Anticancer Res* 21, 3821-3828.

88. 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.
89. Lescuyer, P., Hochstrasser, D., and Rabilloud, T. (2007) How shall we use the proteomics toolbox for biomarker discovery? *J Proteome Res* 6, 3371-3376.
90. Anderson, N. L., and Anderson, N. G. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 1, 845-867.
91. Shen, Y., Kim, J., Strittmatter, E. F., Jacobs, J. M., Camp, D. G., 2nd, Fang, R., Tolie, N., Moore, R. J., and Smith, R. D. (2005) Characterization of the human blood plasma proteome. *Proteomics* 5, 4034-4045.
92. Thulasiraman, V., Lin, S., Gheorghiu, L., Lathrop, J., Lomas, L., Hammond, D., and Boschetti, E. (2005) Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands. *Electrophoresis* 26, 3561-3571.
93. Guerrier, L., Thulasiraman, V., Castagna, A., Fortis, F., Lin, S., Lomas, L., Righetti, P. G., and Boschetti, E. (2006) Reducing protein concentration range of biological samples using solid-phase ligand libraries. *J Chromatogr B Analyt Technol Biomed Life Sci* 833, 33-40.
94. Boschetti, E., Monsarrat, B., and Righetti, P. G. (2007) The "Invisible Proteome": how to capture the low abundance proteins via combinatorial ligand libraries *Current Proteomics* 4, 198-208.
95. Sennels, L., Salek, M., Lomas, L., Boschetti, E., Righetti, P. G., and Rappsilber, J. (2007) Proteomic Analysis of Human Blood Serum Using Peptide Library Beads. *J Proteome Res* 6, 4055-4062.
96. Guerrier, L., Claverol, S., Fortis, F., Rinalducci, S., Timperio, A. M., Antonioli, P., Jandrot-Perrus, M., Boschetti, E., and Righetti, P. G. (2007) Exploring the platelet proteome via combinatorial, hexapeptide ligand libraries. *J Proteome Res* 6, 4290-4303.