Diagnostic Potential of Urine Proteome: A Broken Mirror of Renal Diseases

Michael S. Goligorsky, Francesco Addabbo, and Edmond O’Riordan
Departments of Medicine and Pharmacology and Renal Research Institute, New York Medical College, Valhalla, New York

ABSTRACT
This brief overview of studies into the urine proteome illustrates its potential value for diagnostic, prognostic, and pathophysiologic discovery. Hypothesis-targeted investigations of individual proteins as well as proteome-wide searches for urinary biomarkers of various diseases and their progression are reviewed. The majority of urine proteins appear as cleavage products that are found not only as free solutes but also in secreted membrane vesicles called exosomes. Described are several recent examples of important diagnostic findings using urine proteomics along with the idea that signature profiles of injury to individual nephron segments can be measured by this technology. Shared are some thoughts on the most challenging step: Integration of seemingly unrelated findings of various protein fragments into a rational pathogenetic pathway(s). The future chance that the centuries-old technique of uroscopy will reveal its secrets using modern proteomic approaches makes gradual improvement.


Proteinuria, a cardinal symptom of renal disease, has long been considered as a potential “black box” for diagnostic and even prognostic information. Urine proteins arise from various sources, including filtration of plasma proteins; impaired reabsorption of filtered proteins; and appearance of proteins that originate from injured glomeruli, tubules, infiltrating inflammatory cells, or connective tissue as well as those that enter the urine in the urinary tract below the kidney.

Not surprising, many traditional studies have focused on the excretion of (1) individual proteins such as enzymes or albumin, (2) selectivity of proteinuria, and (3) attempts at global characterization of urine proteins using two-dimensional electrophoresis. Insufficient resolution of many techniques, the lack of understanding of the physiologic and pathologic underpinnings of proteinuria, and frequent appearance of protein fragments that are undetectable using standard approaches all hampered early attempts to extract information from this ubiquitous resource. Recent progress in our understanding of molecular mechanisms of renal handling of proteins, especially albumin,1,2 combined with the ongoing revolution in the technological tools for peptide detection, quantification, and identification,3–7 have re-energized attempts to obtain proteomic footprints of renal disease. In this brief review, we recount the most important findings made so far in the field of urine biomarker proteins, reflect on the capabilities and limitations of these biomarkers—frequently protein fragments—in diagnosing the disease, and discuss the tortuous intellectual routes that lead to the faithful reconstruction of pathogenetic mechanisms of the disease from these seemingly unrelated protein fragments (integrating the image from pieces of the broken mirror).

FROM FOOT PROCESSES TO FOOTPRINTS OF DISEASE
Although the actual amount of protein and albumin filtered by glomeruli remains a highly controversial issue,1 a complex map of the slit diaphragm and podocyte with their protein makeup is rapidly emerging. This map should help in the understanding of the breadth of glomerular proteins that are contributed to the tubular fluid after glomerular injury. Since the discovery of nephrin, a host of protein components of the filtration barrier are known.8 Nephrinuria, for example, has been recognized as a feature of diabetes and diabetic nephropathy.9,10 Among 40 normoalbuminuric patients with type 1 diabetes, nephrinuria, as judged by immunorecognition of protein bands with molecular weight of 18-, 32-, 40-, 60-, 75-, and full-length 185-kD proteins, was detected in 30% of cases, whereas nephrin was undetectable by Western blot analysis in the urine of healthy individuals. Is it possible that nephrinuria may serve as an early warning of impending nephropathy, and studies to address this question are in progress in Holthofer’s laboratory.

In parallel, a more detailed picture of the machinery for tubular protein reabsorption is emerging. Nonspecific receptors megalin and cubilin and their internalization, lysosomal degradation of protein cargo, and recycling to the luminal plasma membrane in the proximal epithelium provided important insights into the origins of tubular proteinuria.11 In fact, it has become increasingly clear that most pro...
Proteins are excreted as proteolytic fragments rather than intact molecules and, therefore, are poorly detectable using today’s standard approaches. For instance, 90% of albumin is degraded to small, <10-kD fragments. Similar degradation has been documented for other proteins, including transferrin, apolipoprotein, IgG, glucose oxidase, and lactate dehydrogenase.

**GENERIC AND DISEASE-SPECIFIC MARKERS OF INJURY**

Commonality in the excretion patterns of certain proteins or their fragments in various diseases offers critical insights into the default mechanisms of pathologic processes. For instance, there are well-established markers of tubular injury (e.g., α2-microglobulin, prealbumin, lactate aminopeptidase) or infectious processes, such as the release of defensins. Mega-lin itself is normally processed via intramembrane proteolysis characterized by the protein kinase C–regulated metalloprotease-mediated shedding of its large ectodomain, and it remains to be elucidated how this process may be altered in kidney disease.

One of the techniques used to distinguish between generic and the site- or disease-specific excretory proteins during proteome-wide analyses is by stratifying overlapping occurrence of the proteins in the former category among patients with diverse diseases and contrasting it with unique occurrences in the latter. This type of analysis can be represented by a Venn diagram, as shown in Figure 1. The figure demonstrates a representative analysis performed by Gene@work-based software of protein peaks characteristic of interstitial fibrosis, tubular injury/atrophy, and vasculopathy in patients with biopsy-graded chronic allograft nephropathy (CAN). The presentation is segregated into classes of protein peaks that are characteristic for the individual histologic feature, peaks that are shared by two histologic features, and peaks that are common to all three major histologic presentations of CAN. The presence of peptides in the first group suggests that there is a theoretical possibility that each morphologic abnormality of CAN is definable, whereas the other two categories provide a basis for believing that certain proteomic markers are common to all patients with CAN.

**INTUITIVE SELECTION OF URINARY CANDIDATE BIOMARKER PROTEINS**

Enzymuria

Enzymuria has long been considered as a potentially informative marker for the detection of tubular injury on the basis of the prediction that the enzymes should leak into the urine from damaged tubular epithelia. Increased urinary excretion of a lysosomal enzyme N-acetyl-glucosaminidase has been reported in contrast nephropathy. Brush border enzymes, such as alkaline phosphatase, γ-glutamyltranspeptidase, and ala(leu-gly)-aminopeptidase, are elevated in acute renal injury. Westhuyzen et al. demonstrated that alkaline phosphatase and π-glutathione S-transferase markers of the brush border and distal tubular epithelia, respectively, have predictive value for development of acute renal failure (sensitivity 100%, specificity 91%) in patients who are admitted to a general intensive care unit. The major limiting factor for these markers consists of the high sensitivity of release by tubular epithelia even in mild injury, which does not necessarily portend development of acute renal failure. This is exemplified by a study of patients who underwent coronary bypass surgery and almost uniformly had postoperative increases in urine excretion of N-acetyl-glucosaminidase but did not succumb to acute renal failure.

Devarajan’s group screened 71 patients who were undergoing cardiopulmonary bypass (a carefully selected group of patients without additional confounding problems) for the urinary excretion of neutrophil gelatinase–associated lipocalin.
(NGAL). NGAL excretion 2 h after cardiopulmonary bypass was a powerful independent predictor of acute kidney injury, diagnosed in this study as a 50% rise in plasma creatinine from baseline, showing the sensitivity of 100% and specificity of 98% for a cutoff value of 50 μg/L (only one of 51 patients without acute kidney injury had urinary NGAL level >50 μg/L). This successful strategy capitalized on the previous unbiased experimental findings obtained through cDNA microarray screening and put forward a role for infiltrating inflammatory cells as important contributors to biomarker profiles from urine. Another inflammatory mediator, IL-18, at levels >100 pg/ml in the urine of 52 patients with acute respiratory distress syndrome was associated with a 6.5-fold increased odds for acute kidney injury in the next 24 h. Importantly, elevated IL-18 had urinary NGAL level 4 to 8 h after renal ischemia. Revealed an early upregulation of a secreted growth factor–inducible immediate early cysteine-rich protein 61 (CYR61), a marker of this acute kidney injury.

Recently it was demonstrated that AQP2 is also excreted in the urine as part of secreted exosomes. Exosomes are small vesicles that are derived indirectly from the apical endosomal system. Among 295 proteins identified, AQP2 (potential marker of nephrogenic diabetes insipidus), polycystin-1 (autosomal dominant polycystic kidney disease type 1), podocin (autosomal recessive steroid-resistant nephrotic syndrome), nonmuscle myosin II (potential marker of Fehner syndrome and Epstein syndrome), angiotensin-converting enzyme (hypertension), Na+/K+/2Cl− co-transporter (Bartter syndrome type 1), thiazide-sensitive Na+/Cl− co-transporter (Gitelman syndrome), and epithelial sodium channel (Liddle syndrome and autosomal recessive pseudohypoaldosteronism type 1) are excreted on exosomes. The studied exosomes were obtained by ultracentrifugation of a large volume of urine (400 ml), and proteins were initially separated by electrophoresis followed by in-gel trypsin digestion. The tryptic peptides were analyzed by nanospray liquid chromatography–tandem mass spectrometry (MS). A modification of this technique was recently used to detect more than a 50-fold increase in fenbufen A excretion in experimental animals and patients with acute kidney injury.

**TECHNOLOGICAL TOOLS FOR PROTEOME-WIDE RESEARCH**

Technological platforms that are used in the discovery phase of proteome-wide research have been reviewed elsewhere, and we only recount them here. There are several types of mass spectrometers that use the assortment of electrical, radio frequency, and magnetic fields: Time-of-flight, quadrupole, ion trap, and Fourier transform ion cyclotron resonance. Mass spectrometers lack the ability to quantify the protein detected, the drawback that can be circumvented using isotope-coded affinity tagging. Because of the competition between proteins and peptides for capture of charges, the detection of low-abundance species may be compromised. Therefore, additional separation strategies are required.

**Two-Dimensional Electrophoresis.** The cartesian position of a protein is a product of its pl and molecular mass. This technique has the ability to be quantitative either through comparison of spot size between two gels or by a differential in-gel electrophoresis. The shortcomings of two-dimensional electrophoresis include that more than one protein per spot can exist and that multiple spots can contain the same posttranslationally modified protein, at times making abundance profiling problematic. Furthermore, very small or large proteins and very acidic or basic proteins are not visualized.

**HPLC.** HPLC is applied to the urine and includes size exclusion, reverse phase, strong and weak cation binding and affinity binding (i.e., Ig adsorbing to protein A). One of the disadvantages of this system is that a buffer dilutes the proteins, thus requiring repeated re-concentration of the sample and making quantification problematic.

**Capillary Electrophoresis.** Capillary electrophoresis is a powerful separation technology, but it lacks the ability to quantify absolutely the proteome.

**Protein Array Technologies**

This concept is similar to gene microarray with antibodies or tissue sections robotically placed on a glass slide. Techniques for protein–protein (including antigen–antibody), protein–DNA, protein–lipid, and
protein–drug interactions exist. Limitations include accurate quantification and dependence on the availability of antibodies. This technology has yet to be applied to renal proteomics. A variation of this technique is Luminex multiplex analysis, which is based on the array of polystyrene microspheres with two spectrally distinct fluorochromes. Using the precise ratio of these fluorochromes, 100 bead sets have been created, each with its unique color-coded signature. Each signature bead is conjugated to an analyte-specific antibody and combined in a single assay to measure up to 100 analytes in up to 96 samples simultaneously. The assays are based on the conventional two-site sandwich method. After conjugation reaction, a mixture of beads is analyzed using a dual wavelength laser flow cytometer–like apparatus. One laser beam detects color-coded beads, and another quantifies the reporter signal on each bead. These techniques may be better suited for a more targeted analysis of isolated proteins.

Bioinformatics

Several databases are used to identify protein fragments by peptide mass fingerprinting after trypsin digestion or after MS/MS sequencing. Databases compare the size of the fragments recorded by the mass spectrometer with the translated DNA sequence; search for trypsin digestion sites, comparing the theoretical with the measured values; and calculate the probability of a correct match. Logistic analysis of the findings is conducted using the following:

1. Tree-based technologies such as RandomForest:
   As the samples are classified according to the proteins that most accurately classify the whole population, an inverted tree is formed.

2. AdaBoost: A method for combining weak classifiers to create a summary and stronger classifier: The basic principle is that after selection of the variable that is most likely to predict correctly the class of a sample, the samples are reweighted with increasing weight applied to the misclassified samples; the next best classifier variable is then selected. This process is then repeated with summation of the classifiers to create a robust and accurate classifier.

3. Genetic algorithms, neural networks, and unified maximum separability analysis: These are alternative approaches that have also proved to be useful in classification that is based on spectral data.

In summary, it is critical to appreciate that there is no single proteomic or informatics technique to fit the diverse requirements of analyses; therefore, combination of several approaches offers the optimal solution to the problem.

PROTEOME-WIDE SEARCH AND SELECTION OF BIOMARKERS

Using the arsenal of technological and bioinformatics tools already discussed, the following findings have been reported.

Nephropathies

In a study of 57 control urine samples compared with samples from patients with minimal-change disease (n = 16), membranous nephropathy (n = 18), and FSGS (n = 10), a group of 690 polypeptides were present in >50% of all normal samples. Plots of >500 polypeptides typical of each disease were compiled. The rates of correct classification were 71.4% for minimal-change disease/FSGS and 92.9% for membranous nephropathy.

Woroniecki et al. studied steroid-resistant nephrotic syndrome in a pediatric population with idiopathic nephritic syndrome. A protein of mass 4144 Da was identified as the single most important marker for distinguishing steroid-sensitive and steroid-resistant patients with a high level of confidence.

Distinct polypeptide signatures also seem to be associated with IgA nephropathy. In a study of 45 patients—including those whose total urinary protein levels were within normal ranges—the urinary peptide patterns had a sensitivity of 100% and a specificity of 90%. Three of the most promising polypeptides were sequenced and shown to be albumin fragments. IgA nephropathy could be differentiated from membranous nephropathy with a sensitivity of 77% and a specificity of 100% and from minimal-change disease, FSGS, and diabetic nephropathy with a sensitivity and a specificity of 100%.

Cutillas et al. applied three different techniques to examine the Dent disease proteome both qualitatively and quantitatively. They found that carrier proteins, complement components, and bioactive peptides were excreted at higher concentrations in patients with Dent disease.

Proteome of the Transplanted Kidney

Three investigative groups have reported proteomic diagnosis of acute renal allograft rejection. Clarke et al. found that proteins of 6.5, 6.6, 6.7, 7.1, and 13.4 kDa performed well as biomarkers of acute rejection. Proteins in the mass ranges 5270 to 5550 and 10,530 to 11,000 Da were reported to be good biomarkers by Schaub et al. and a subsequent report identified these proteins as β-2 microglobulin and its fragments. O’Riordan et al. identified urine proteins with masses of 4756.3, 25,665.7, and 19,018.8 Da as candidate markers of acute kidney transplant rejection compared with recipients with stable transplants. Multiple protein peaks provided a more accurate assessment than relying on only single biomarkers. In a more recent study, O’Riordan et al. chemically identified β1-defensin and anti-chymotrypsin as valuable candidate biomarkers of acute rejection. We also obtained preliminary data suggesting that chronic allograft nephropathy is associated with the increased degradation of perlecan and urinary excretion of its fragment endorepellin (unpublished observations, M.S.G. and E.O.).

Diabetic Nephropathy

When 29 healthy individuals were compared with 112 patients with type 2 diabetes, a distinct “no albuminuria/diabetic pattern” was detected. Another distinct pattern was observed in patients with albuminuria >100 mg/L. Found in 35% of patients with elevated urinary albumin excretion rates and only 4% of healthy volunteers, this pattern identified individuals who were more likely to have retinopathy. The characteristic polypeptides were insulin-like peptide 3, uromodulin, and an albumin fragment. The putative role of nephrinuria as a potential biomarker of diabetic nephropathy has been mentioned.
Renal Cancer
Rogers et al.47 analyzed urinary proteome in patients with clear cell renal carcinoma and compared it with that of healthy volunteers and patients with other urogenital diseases. In another study, kininogen levels were found to be elevated in the urine of a patient with renal cancer; the concentration fell after nephrectomy.7 Urinary proteomic analysis has identified several biomarkers of bladder cancer: γ-Synuclein, a soluble isofrom of catechol-O-methyltransferase, and calreticulin, which, when tested prospectively, were found to have a combined sensitivity of 76.8% and a specificity of 77.4%.48

DYNAMICS OF DISEASE AND METAMORPHOSIS OF MARKERS
It would be unrealistic to expect that each disease process is identifiable by unique signature proteins in the urine, which have been obtained during a previous validating snapshot analysis. Different stages of disease, variations in mechanisms, and other comorbidities are likely to modify urine proteome. An example of the dynamics among signature proteins in the urine is presented in Figure 2, where the results of proteomic analyses are plotted against the Banff-defined stages of CAN. Knowledge of protein dynamics in the course of disease processes may be helpful in their staging and in monitoring response to therapy.

FROM DEDUCING PROTEIN IDENTITY TO INTEGRATING THE FINDINGS INTO THE PATHOGENETIC PATHWAY: A LONG WAY TO GO
Let’s consider an optimistic scenario: With multiple databases developed for assistance in chemical identification of proteins, the investigator, after completing cross-sectional and prospective analyses, eventually confirms the diagnostic value of an ensemble of biomarker peptides/proteins. As important as it is by itself, the actual understanding of the defined markers can be attained only when the mechanisms of their appearance/disappearance in the urine become elucidated. Reconstruction of molecular pathways that are involved in these processes and their integration into systems biology are aided in silico by the growing collection of publications dealing with components of individual pathways, as well as by the emergence of several software suits for pathway analysis, yet the connectivity of diverse pathways through multiple components of an individual pathway results in a complex arborized structure that requires actual analysis of multiple components to profile the correct one, as illustrated in Figure 3. Taking into account that urinary proteins reflect not only the pathway(s) but also the mechanism(s) of the appearance or disappearance of a certain peptide/protein component of a pathway in the excretory compartment, this analysis becomes even more complicated. For instance, we recently identified a group of matricellular proteins—all products of the enzymatic cleavage of larger molecules—that appear in the urine of patients with CAN; importantly, the same enzyme is responsible for the activation of the latent TGF-β (unpublished observation, M.S.G. and E.O.), thus potentially linking several pathways leading to fibrosis. Integration of proteomic findings into pathogenetically rational pathways requires substantial investiture. Integration can be assisted by the analysis of cDNA microarrays (e.g., detection of increased expression of defensin and anti-chymotrypsin mRNA corroborated the findings made in acute renal allograft rejection and highlighted the importance of pro- and anti-inflammatory pathways).45 Tissue analysis of proteins in question could be best met using immunohistochemical analyses of multiple intermediates and/or through the recently described MS analysis of tissue sections. The latter permits the acquisition of protein profiles from individual nephron segments under microscopic guidance,53 an excellent example of fusion between histology and proteomics.

FUTURE DIAGNOSTIC ALGORITHMS
This brief overview was intended to proselytize for proteomics approaches to renal disease and intentionally overlook the abounding problems—these are addressed in several recent reviews.49–52 In short, many technical and bioinformatics issues await resolution, yet initial important findings are emerging, and the whole field of inquiry is undergoing exponential growth. Through accumulation of candidate biomarkers and their validation in larger patient populations, various signature combinations of proteins and their fragments should eventually become available for many diseases and their stages. Studies of the urinary proteome are inseparable from in-depth morphologic analysis of the kidney, proteomic mapping of different nephron...
Figure 3. Blueprints for integration of protein/peptide findings into metabolic pathways: Systems biology approach. Proteins/peptides that belong to diverse pathways (depicted as separate planes) may be linked via cross-talk between the pathways (blue lines) or promoter elements (entries into each pathway), common enzymes (brown semicircles) acting on individual elements from different pathways, or comprising feedbacks between pathways (red lines). Proof of such links may require additional targeted studies to reconstitute connectivity, enzymes, or regulatory elements (blue). In addition, tissue analysis could unveil whether changes in the urinary proteome are a result of parallel or reciprocal changes in the abundance of biomarker proteins.

REFERENCES
These studies were supported in part by National Institutes of Health grants DK45462, DK54602, and DK52783 (M.S.G.) and the Westchester Artificial Kidney Foundation. E.O.R. was a recipient of the Kevin J. and Gloria B. Kiely National Kidney Foundation Fellowship.

DISCLOSURES
None.


