Discovery of Urinary Biomarkers*

Trairak Pisitkun‡, Rose Johnstone§, and Mark A. Knepper‡¶

A myriad of proteins and peptides can be identified in normal human urine. These are derived from a variety of sources including glomerular filtration of blood plasma, cell sloughing, apoptosis, proteolytic cleavage of cell surface glycosylphosphatidylinositol-linked proteins, and secretion of exosomes by epithelial cells. Mass spectrometry-based approaches to urinary protein and peptide profiling can, in principle, reveal changes in excretion rates of specific proteins/peptides that can have predictive value in the clinical arena, e.g. in the early diagnosis of disease, in classification of disease with regard to likely therapeutic responses, in assessment of prognosis, and in monitoring response to therapy. These approaches have potential value, not only in diseases of the kidney and urinary tract but also in systemic diseases that are associated with circulating small protein and peptide markers that can pass the glomerular filter. Most large scale biomarker discovery studies reported thus far have used one of two approaches to identify proteins and peptides whose excretion in urine changes in specific disease states: 1) two-dimensional electrophoresis with mass spectrometric and/or immunochemical identification of proteins and 2) top-down mass spectrometric methods (SELDI-TOF-MS and capillary electrophoresis-MS). These studies have been chiefly in the areas of nephrology, urology, and oncology. We review these applications, focusing on two areas of progress, viz. in bladder cancer and in acute rejection of renal transplants. Progress has been limited so far. However, with the advent of powerful LC-MS/MS methods along with methods for quantifying LC-MS/MS output, there is hope for an accelerated discovery and validation of disease biomarkers in urine. Molecular & Cellular Proteomics 5:1760–1771, 2006.

The use of urinary biomarkers to diagnose disease is a long-standing practice. Ancient clinicians detected glucose in the urine by tasting it or observing whether it attracted ants. The presence of albumin in the urine has been measured as an indicator of renal disease for centuries and in early times could be detected via the so-called “foam test” to determine whether albumin was present in the urine in large amounts. Even today clinicians frequently shake a urine sample to determine whether it develops a froth, prima facie evidence for a high level of protein, which often is indicative of glomerular disease. In that tradition, studies to identify biomarkers of disease in the urine have been an underlying component of investigative medicine throughout the 20th century and the early 21st century. These studies have been based on knowledge of the pathophysiology of disease to identify putative biomarkers that could be tested in clinical trials.

The advent of protein mass spectrometry has enabled a new approach to biomarker discovery, an open ended approach in which putative protein biomarkers can be identified by large scale profiling of the protein complement of urine (“proteomics” profiling). Large scale proteomics profiling of normal human urine samples has revealed the presence of at least 1000 different protein gene products and many more peptide fragments of larger proteins (1–11). There is hope, therefore, for discovery of urinary protein excretion profiles that can be used clinically for tasks such as early detection of disease, classification of disease, choice of therapeutic agents, assessment of prognosis, and monitoring of a particular therapeutic regimen. Although at face value this task seems straightforward, there are important barriers that must be overcome for such approaches to succeed. The purpose of this review is to discuss these barriers and the work that has been reported so far to overcome these barriers.

THE DISCOVERY-VALIDATION-IMPLEMENTATION PARADIGM

A proposed workflow approach for the development of clinically applicable protein biomarker assays is shown in Fig. 1 (12). In this paradigm, there are three fundamental stages in the developmental process: discovery of presumptive biomarkers, validation of these biomarkers with regard to their abilities to make useful predictions in patient populations, and implementation (development of a clinical assay, regulatory approval, etc.). We discuss these in turn.

Discovery of Presumptive Protein or Peptide Biomarkers—The first stage of biomarker development, the discovery stage, is essentially a hypothesis generation step. Traditionally hypotheses about potentially useful urinary biomarkers have been derived deductively from an understanding of pathophysiology. However, in the past few years, more and more investigators are turning to the use of protein mass spectrometry to discover presumptive biomarkers. Even when protein mass spectrometry is used, the discovery process can benefit from a consideration of the likely types and origins of proteins involved in the pathophysiology of the disease under consideration. For example, the choice of whether to isolate whole cells, exosomes, or the soluble pro-
were actually true positives. Thresholds for success must be generally expressed as a percentage of positive readouts that measure of the ability to eliminate false positives and is generally expressed as a percentage expressing the number of true positives detected relative to the total number of patients with the index disease or disease characteristic.

In many studies reported thus far, presumptive biomarkers have been reported mainly as lists of mass-to-charge ratios of urinary polypeptides rather than as specific proteins or peptides (13–22). In theory, if mass spectrometry is used for validation and implementation, this list of mass-to-charge ratios can be sufficient to make useful diagnostic or prognostic predictions. However, it is potentially more fruitful to identify the putative biomarker proteins and peptides at an amino acid sequence level. There are at least two reasons for the preference for explicit protein/peptide identification. 1) Validation and implementation phases may be more cost-effective if they use technologies that do not require mass spectrometers, e.g., antibody-based microarrays or ELISAs. These alternative approaches require an explicit protein identification. 2) Identification of specific proteins may shed light on the pathophysiology of the disease process under consideration. Thus, in our view, the goal in most cases should be to identify proteins and peptides explicitly.

Beyond explicit identification is a strong need for quantification of protein abundance or relative abundance. Thus far, most biomarker discovery studies have used a qualitative approach, looking for proteins or peptides that either appear or disappear in the disease samples. However, the ability to discriminate adequately between some patient groups can benefit from quantification, i.e., assessment of changes in amounts of specific proteins rather than absolute presence or absence. We deal with quantitative approaches below under “Quantification in the Discovery Phase.”

Validation of Predictive Capabilities of Protein Biomarker Measurements—Once a presumptive biomarker or a presumptive set of biomarkers is identified in the discovery phase of research, the ability of these biomarkers to specify diagnosis, prognosis, optimal treatment regimen, etc. must be determined through validation studies. In general, such studies yield two parameters that quantify the predictive capabilities of the method, viz. sensitivity and specificity. Sensitivity is a measure of the ability to detect true positives and is generally expressed as a percentage expressing the number of true positives detected relative to the total number of patients with the index disease or disease characteristic. Specificity is a measure of the ability to eliminate false positives and is generally expressed as a percentage of positive readouts that were actually true positives. Thresholds for success must be assigned based on the natural history of the disease under consideration and potential consequences of false positives and false negatives.

In general, validation studies require much larger patient populations than do discovery experiments. These populations should be independent of the populations used for discovery. The assay that is used, whether based on mass spectrometry, antibody-based quantification, or other means, should be the same as that proposed for implementation. Validation studies are clinical trials that require the same sorts of regulatory certification as other clinical trials.

Implementation of Biomarker Assays—Prior to initiation of discovery or validation studies, it is important to assess how validated biomarkers can be brought to the clinical arena. In particular, a proposed assay or set of assays must be developed prior to the start of validation studies so that data from the validation trial can be used in the approval process (e.g., by the United States Food and Drug Administration). As discussed, it is theoretically possible to use mass spectrometry for implementation of a clinical assay. However, in our view, protein binding assays such as antibody arrays and ELISAs present a superior means of implementation chiefly because these methods will more readily benefit from the “economy of scale” in which price falls with increased volume. Ideally many biomarker assays can be incorporated on a single antibody array chip to allow personalized medicine to be carried out at costs that will allow everyone to benefit irrespective of socioeconomic status.

A significant concern in the use of mass spectrometers for clinical measurement of biomarkers is that, along with the benefit the biomarker assays bring, the popularization of the approach will engender increased costs to society as high end mass spectrometers and associated instrumentation are acquired by clinical laboratories. On the other hand, the development of reliable assays for disease biomarkers in urine has the potential to contribute to a net reduction in aggregate costs. For example, early diagnosis of glomerular disease could allow therapeutic measures that substantially delay or obviate the development of end stage renal disease, thus reducing the demand for dialysis and transplantation. Likewise early diagnosis of bladder cancer and other forms of malignancy could allow curative surgical treatment. Net reduction in overall costs requires that the biomarker assays that are implemented are low cost and available worldwide, a goal that probably can be more readily achieved with immunochemical assays rather than mass spectrometry at points of care.
An important step in planning implementation is to model how the assay would be used in the clinical setting. For example, if a biomarker assay for a urinary tract malignancy was to be developed that had a 1% false positive rate and if this assay was to be used to screen for the malignancy in every member of the United States population, three million individuals would show false positives. Depending on the diagnostic workup that would be required to reveal these false positives, a decision might be made to require a very stringent level of specificity to pass the validation stage.

**Sources of Urinary Proteins**

Urinary proteins include soluble proteins and protein components of solid phase elements of urine (Table I). Solid phase elements consist of “sediments” that can be precipitated at low centrifugation speeds and “exosomes” that are of very low density and sediment only with ultracentrifugation. Prefractionation of these components can be useful as a means of enriching for markers of particular types of disease. A study of urine collected from normal human adult subjects indicated that, of the total urinary protein excreted, ~48% was contained in sediments, 49% was soluble, and the remaining 3% was in exosomes (23).

**Soluble Proteins**—The soluble proteins in urine are derived largely from glomerular filtration. The glomerular filter effectively retards passage of high molecular weight proteins. However, even with very low sieving coefficients, proteins that are abundant in the blood plasma such as albumin and various globulins can pass the glomerular filter in substantial amounts to enter the lumen of the nephron. Beyond this, peptides and small proteins (<10 kDa) are freely filtered by the glomerulus. Most of the proteins and peptides that pass the glomerular filter are scavenged and proteolyzed in the proximal tubule by highly specialized apical uptake processes that involves receptor-like recognition of the polypeptide molecules (24, 25). Thus, a change in the amount of a given soluble protein that reaches the final urine can result from a change in its concentration in the blood plasma, a change in the function of the glomerular filter, or an alteration in the proximal tubule scavenging system. Based on these mechanisms, changes in excretion rate of specific urinary proteins can be indicative of systemic disease, glomerular disease, or diseases affecting the proximal tubule, respectively.

Some of the soluble proteins in urine originate as membrane-bound proteins that are proteolytically cleaved from their membrane attachments. One of these is Tamm-Horsfall protein (uromodulin), an abundant soluble urinary protein that is secreted by the thick ascending limb of Henle loop, a nephron segment downstream from the proximal tubule (26). It originates as a glycosylphosphatidylinositol-linked protein present in the apical plasma membrane that can be cleaved from its cell attachment proteolytically. Tamm-Horsfall protein is generally the most abundant protein in urine (Fig. 2A), and its presence can interfere with the detection of other proteins if it is not removed (5). It is a member of the zona pellucida (ZP) protein family that forms large networks of fibers, typified by the zona pellucida of fertilized ova (27). Tamm-Horsfall protein forms similar networks of fiber (Fig. 2B), which can constrain membrane elements in the urine and interfere with fraction-
tion procedures. Tamm-Horsfall protein can be depolymerized by use of reducing agents and warming (5).

**Solid Phase Components**—Urine typically contains relatively high density sediments consisting chiefly of sloughed epithelial cells and casts, both of which can be isolated using a low speed centrifugation. Increased numbers of whole cells or casts can be indicative of renal disease, and the microscopy-based classification of cells and casts can give important diagnostic information. Beyond this, there are other solid phase components of urine that constitute a source of urinary proteins and a pool of potential disease biomarkers. Some of these components are small fragments of membrane that are delivered to the urinary space by shedding of microvilli or by apoptosis. These components can be isolated by centrifugation at moderate speed, typically $17,000 \times g$. Additionally urine contains numerous exosomes, derived from virtually every epithelial cell type facing the urine including glomerular podocytes, renal tu-

---

**Fig. 2.** A, two-dimensional gel of urinary low density membranes prepared without prior Tamm-Horsfall protein (THP) removal process (data from Pitskun et al. (5)). Horizontal separation was by means of isoelectric focusing (Ettan IPGphor, Amersham Biosciences). Vertical separation was by SDS-PAGE with an Ettan DALTwelve separation unit (Amersham Biosciences). The gel was stained with SYPRO Ruby (Molecular Probes). The abundant protein at ~90 kDa was identified as Tamm-Horsfall protein by MALDI-TOF MS (Ettan MALDI-TOF Pro, Amersham Biosciences). The low pl of this abundant protein is consistent with the low theoretical pl of Tamm-Horsfall protein (4.97). In an additional experiment, Tamm-Horsfall protein was removed from the final pellet by reduction with DTT and reultracentrifugation. The appearance of Tamm-Horsfall protein in the soluble fraction after ultracentrifugation at $200,000 \times g$ in this experiment is consistent with reports that most Tamm-Horsfall protein in urine has been released from the membrane-bound state (76). B, Tamm-Horsfall protein networks and constrained vesicles in urine demonstrated by immunogold electron microscopy of urinary low density membranes probed with antibody to Tamm-Horsfall protein (a kind gift from Dr. John R. Hoyer). The inset shows the magnified image within the dashed box; the arrowheads indicate gold particles (12 nm in diameter), and the arrow indicates a vesicle.
Exosomes can be isolated from urine by either high speed centrifugation or ultrafiltration. (See “Proteomics of Urinary Exosomes” below for more detail on these structures and their role in disease diagnosis.) The native proteins and peptides in urine are delivered to the mass spectrometer without prior trypsinization (Fig. 3). De- 

fig. 3. Schematic of top-down versus bottom-up approaches. The top-down approach is normally applied to small proteins and/or peptides without preceding digestion. First stage mass analysis of precursor ions (MS1) in this approach yields the list of m/z that can be used for the pattern comparison between control and disease conditions. In contrast, the bottom-up approach requires an initial digestion step with proteases to break proteins into peptides, allowing protein identification based on peptide mass fingerprinting from MS1 or on peptide fragmentation spectra from MS2.

**Top-down versus bottom-up approaches**

Analytical protein mass spectrometry generally uses a so-called “bottom-up” approach in which protein identification depends on initial treatment with a protease (usually trypsin) to break proteins into relatively small peptides whose m/z can be accurately determined by mass spectrometry (Fig. 3). This allows reliable identification based on the so-called “peptide mass fingerprinting” or using peptide fragmentation spectra in a tandem mass spectrometer (2–8, 11, 29–35).

A different approach has become popular in biomarker discovery experiments, viz. a so-called top-down approach in which the native proteins and peptides are delivered to the mass spectrometer without prior trypsinization (Fig. 3). Depending on the mass accuracy of the mass spectrometer being used, peptides and small proteins up to 10–20 kDa can often be recognized. Top-down strategies have been exploited using both capillary electrophoresis (CE)-MS (13–17, 36–39) and SELDI-TOF-MS (18–22, 40–43) systems. Absolute identification of proteins is more difficult with top-down approaches than with bottom-up approaches (m/z ratios are often reported rather than protein identifications). However, top-down approaches have the major advantage of being able to recognize native small peptides (e.g., proteolytic fragments) that may prove to be excellent markers of a given pathophysiological entity. This may be of particular value in cancer proteomics where proteases that are secreted by tumors may reflect the presence of and/or the degree of aggressiveness of the malignancy. These proteases may be more readily recognized by the pattern of proteolytic fragments generated than by direct identification of the protease proteins themselves (44). Conceivably if these proteolytic fragments are present in urine, they may be easier to detect therein than in blood.

**Quantification in the discovery phase**

Quantification of protein and peptide abundances is an important prerequisite for success in biomarker discovery studies. Most of the top-down discovery studies reported so far provide an indication of the relative abundance of specific proteins and peptides (control versus patient urine samples) based simply on measurement of peak height in spectra from TOF mass analyzers run in the linear mode. Bottom-up approaches allow more facile protein identification, but quantification can be more challenging. Proteomics profiling based on two-dimensional electrophoresis can give accurate quantification based on spot density comparisons either by matching parallel gels for control and patient samples or through use of fluorescence labeling methods such as DIGE (45) that allow comparison of two samples in the same gel. Shotgun proteomics using LC-MS/MS is enjoying increased use in proteomics investigations in general and is beginning to be used for biomarker discovery studies. A number of methods for quantification in the setting of LC-MS/MS have become available in recent years. These methods, not yet extensively used in reported biomarker discovery studies, can be divided into two categories, viz. non-labeling methods and labeling methods.

Non-labeling methods applied to LC-MS/MS resolve the two samples in separate LC-MS/MS runs and quantify appropriate peptides by integrating the area under the pseudodimensional electrophoresis chromatograms constructed of the MS1 peak heights for the appropriate peptide ions (46) (Fig. 4). The peptides are identified using the corresponding MS2 spectra. There are significant cost advantages of such an approach vis-à-vis labeling methods (see below) because there is no need to purchase expen-

1 The abbreviations used are: CE, capillary electrophoresis; iTRAQ, isobaric tags for relative and absolute quantitation; MVB, multivesicular body; VPS, vacuolar protein sorting; TCC, transitional cell carcinoma; 2-D, two-dimensional.
sive derivatizing reagents. Success with the method depends on the quality of the mass spectrometer particularly with regard to mass resolution. False positive identification of peaks in the MS\(^1\) spectra (due to insufficient mass resolution) during construction of the pseudochromatogram can result in errors in quantification, and the analyses must be monitored for such errors.\(^2\)

Labeling methods include ICAT, iTRAQ, and end labeling approaches. ICAT uses matched derivatizing agents containing isotopic substitutions (H\(^2\) versus H\(^1\) or C\(^{13}\) versus C\(^{12}\)) to label cysteines on two samples to be quantitatively compared (47). iTRAQ uses different isotopic tags that are indistinguishable by precursor mass (isobaric) but yield different masses upon fragmentation by collision-induced dissociation (m/z from 114 to 117), allowing relative quantification of the same peptide from different samples (48). End labeling can be done in several ways, the simplest of which is to carry out the trypsinization alternatively in \(^{18}\)O-labeled water or unlabeled water, adding an \(^{18}\)O or \(^{16}\)O atom to the C terminus of each tryptic peptide (49). Another effective end labeling method is IVICAT (in vacuo isotope-coded alkylation technique) in which the N termini of tryptic peptides are trimethylated using either methyl iodide (CH\(_3\)I) or deuterium-substituted methyl iodide (CD\(_3\)I) giving a mass difference to 9 amu (50). The quantification for all labeling techniques is generally done through construction of a pseudochromatogram from the peak heights of identical precursor ion masses, expected to represent identical peptides, along the corresponding LC elution period.

\(^2\) T. Pisitkun and M. A. Knepper, unpublished observations.
All of these quantitative approaches require reliable control samples. In general, control samples can be constructed by pooling samples from a large number of normal subjects. The effects of factors such as age, gender, and diet on the urinary proteome have not yet been reported, and care should be taken to match these variables as much as is practical with those of the patient group being evaluated.

For quantitative comparisons, it is desirable to measure the rate of excretion of putative biomarkers and ask whether the excretion rates differ between patients and controls. Measurements of biomarker concentrations alone are insufficient because normal physiological variations in water excretion can dilute or concentrate urinary proteins. For absolute quantification, a 24-h urine collection would be desirable. However, it is probably impractical to carry out quantitative collections for most biomarker discovery and validation studies because of low patient compliance and a history of unreliable results (51). An appropriate alternative is to collect spot urine samples and normalize the biomarker concentration by the creatinine concentration. Creatinine is excreted in the urine at relatively constant rates allowing it to be used to normalize urinary excretion of a particular protein in lieu of time. In practice, it might prove most effective to determine the creatinine concentration for each sample and adjust the volume of test sample to obtain the same amount of total creatinine for each sample. Normalization or loading on the basis of total protein amount is generally unsatisfactory because total protein excretion can vary broadly among various pathophysiological states.

Our own studies have focused on urinary exosome proteomics because we believe that exosomes may be the best source of biomarkers for renal tubulopathies, i.e. disorders that affect the function of renal tubule epithelia. Because exosomes only account for around 3% of the total protein in normal human urine samples, their isolation can result in a very large enrichment of urinary proteins that are derived from renal tubule epithelial cells.

Exosomes were discovered by and named by one of us (R. Johnstone) in studies of erythrocyte maturation (28). They have been the subject of recent review articles (52–54). Since the initial description of exosomes, many cell types have been shown to secrete them including B cells, T cells, dendritic cells, reticulocytes, mastocytes, enterocytes, and platelets. Erythroleukemia cells and other tumor cells also secrete exosomes, and there may be some advantage in the examination of exosomes in cancer proteomics studies. In addition to their abundance in urine (5), exosomes have been found in blood plasma (55–58) and potentially could be utilized to enrich biomarkers derived from maturing erythrocytes, leukocytes, endothelial cells, or metastatic cancer cells.

Exosomes are the internal vesicles of multivesicular bodies (MVBs) that are delivered to the extracellular fluid by fusion of the outer membrane of MVB with the plasma membrane. The mechanism of urinary exosome secretion is summarized in Fig. 5 (see figure legend for full description). Urinary exosomes
are derived from all cell types that face the urinary space including glomerular podocytes, renal tubule cells, and the cells lining the urinary drainage system (5). It remains untested whether renal or urinary drainage tract malignancies are associated with excretion of exosomes derived from cancer cells.

The proteomics analysis of urinary exosomes from normal human subjects has been carried out using LC-MS/MS coupled with upstream one-dimensional SDS-PAGE separation (5). The experiments identified a number of proteins involved in exosome biogenesis such as class E vacuolar protein sorting (VPS) proteins (ESCRT (endosomal sorting complex required for transport) I-II-III, VPS4, and ALIX). The study identified 295 proteins in urinary exosomes, including at least 19 proteins already known to be associated with various renal diseases and hypertension (Table II). The full list of urinary exosome proteins is available on the World Wide Web (Urinary Exosome Protein Database, dir.nhlbi.nih.gov/papers/lkem/exosome/index.htm).

### Table II

<table>
<thead>
<tr>
<th>Diseases or syndromes</th>
<th>Identified proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autosomal dominant and autosomal recessive nephrogenic diabetes insipidus</td>
<td>Aquaporin-2</td>
</tr>
<tr>
<td>Antenatal Bartter syndrome type 1</td>
<td>Sodium potassium chloride cotransporter-2</td>
</tr>
<tr>
<td>Gitelman syndrome</td>
<td>Thiazide-sensitive Na-Cl cotransporter</td>
</tr>
<tr>
<td>Autosomal recessive pseudohypoaldosteronism type 1</td>
<td>Epithelial sodium channel α, β, γ</td>
</tr>
<tr>
<td>Liddle syndrome</td>
<td>Epithelial sodium channel β, γ</td>
</tr>
<tr>
<td>Familial renal hypomagnesemia</td>
<td>FXYD domain-containing ion transport regulator-2</td>
</tr>
<tr>
<td>Autosomal recessive syndrome of osteopetrosis with renal tubular acidosis</td>
<td>Carbonic anhydrase II</td>
</tr>
<tr>
<td>Proximal renal tubular acidosis</td>
<td>Carbonic anhydrase IV</td>
</tr>
<tr>
<td>Autosomal dominant polycystic kidney disease type 1</td>
<td>Polycystin-1</td>
</tr>
<tr>
<td>Medullary cystic kidney disease 2 and familial juvenile hyperuricemic nephropathy</td>
<td>Tamm-Horsfall protein (uromodulin)</td>
</tr>
<tr>
<td>Autoimmune polyendocrinopathy with enteropathy and susceptibility to infections</td>
<td>Podocin</td>
</tr>
<tr>
<td>Fehnert syndrome and Epstein syndrome</td>
<td>Nonmuscle myosin heavy chain IIA</td>
</tr>
<tr>
<td>2,8-Dihydroxyadenine urolithiasis</td>
<td>Adenine phosphoribosyltransferase</td>
</tr>
<tr>
<td>Hypertension</td>
<td>Angiotensin I-converting enzyme isoform-1, aminopeptidase A, aminopeptidase N, aminopeptidase P, nephrilysin, 15-hydroxyprostaglandin dehydrogenase (NAD), dimethylarginine dimethylaminohydrolase-1</td>
</tr>
</tbody>
</table>

**HANDLING AND STORAGE OF URINARY SAMPLES**

There is a critical need for development of standard protocols for collection and storage of urine samples. Two studies, published recently, have addressed these issues in the context of mass spectrometric analysis of soluble components of urine (59) and the analysis of urinary exosomes (23). However, more work is needed in this area.

Schaub et al. (59) published a careful evaluation of urine collection and storage procedures in the context of top-down studies of unfractionated urine using SELDI-TOF-MS technology. Differences were noted in spectra obtained from first void and midstream urine samples in female subjects indicating that the collection method can be important. Storage at 4 °C resulted in altered spectra over a 3-day period reflecting sample degradation. Samples stored at −70 °C, however, gave identical spectra to those from fresh urine samples. Freezing and thawing had little effect on the spectra for up to four freeze-thaw cycles, although evidence of degradation began to be apparent after five freeze-thaw cycles.

Zhou et al. (23) also carried out a thorough evaluation of sample handling and storage conditions with regard to preservation of urinary exosomes using gel electrophoresis and immunoblotting to evaluate protein yield and integrity. The experimental results demonstrated the importance of the use of protease inhibitors (and bacterial inhibitors) to prevent degradation in samples collected in a clinical setting. They also demonstrated that there is normally no significant in vivo degradation of proteins in the first morning urine due to overnight retention in the urinary bladder. The studies also confirmed the importance of freezing and mixing of thawed samples after retrieval from low temperature storage. Freezing at −70 °C gave a better recovery of exosomes than did freezing at −20 °C perhaps because slow freezing aggregates the exosomes or results in the formation of large networks of polymerized Tamm-Horsfall protein prior to complete solidification (Fig. 2B). Such networks would be expected to entrain and precipitate exosomes at a lower centrifugation speed than would be needed for unpolimerized samples. Recommended protocols for isolation and storage of exosomes have been published on the World Wide Web (intramural.niddk.nih.gov/research/uroprot/).

**URINARY PROTEOMICS IN CANCER DIAGNOSIS**

Early cancer detection is an important goal in clinical medicine. In this regard, urinary proteomics offers an attractive approach to cancer biomarker discovery both for kidney and
urological malignancies (19, 32, 34, 35, 40, 60–63) but also for systemic malignancies (33, 43). The latter may be detectable as a result of generation of proteolytic peptides in the blood plasma that can pass the glomerular filter and be excreted in the urine (see “Top-Down Versus Bottom-Up Approaches” above). Perhaps the greatest progress has been made in the area of bladder cancer.

Most bladder cancers are transitional cell carcinomas (TCCs) that arise from the epithelium lining the urinary drainage system. Squamous cell carcinoma is much less common. Celis et al. (64) have performed thorough 2-D gel analysis of tumor tissue from both types of bladder cancer to create a reference database useful in biomarker discovery endeavors (proteomics.cancer.dk/). The differentially expressed proteins described may be considered candidates for identification of possible urinary biomarkers.

Although transitional cell carcinoma has been the chief target of biomarker discovery and validation studies, some work using 2-D electrophoresis of tissue samples has focused on the rarer squamous cell carcinoma (60, 65). The experiments have identified psoriasin as a potential biomarker protein that is present in urine from squamous cell carcinoma patients but is less frequently detectable in urine from transitional cell carcinoma patients (60). In the remainder of this section, we focus on the more common transitional cell carcinoma of the bladder.

Mass spectrometry-based biomarker discovery studies of urine focusing on TCC of the bladder are being carried out based on the prior availability of a number of potentially useful urinary biomarker assays (hemoglobin, nuclear matrix protein-22, complement factor-H, hyaluronic acid-hyaluronidase, cytokeratins 8 and 18, tumor-associated trypsin inhibitor, soluble E-cadherin, prothymosin-α, etc.). An excellent review of these assays has been presented by van Rhijn et al. (66). Some of these are already being used clinically. Existing assays are used largely to provide a means to monitor patients for recurrence after treatment for transitional cell carcinoma. In general, these assays are not sufficiently specific to be used to screen the general population for bladder cancer.

Urinary cytology has been used for decades but is not sufficiently sensitive to detect bladder cancer at an early stage in most cases. New assays using whole cells isolated from urine and developed from pathophysiologically based considerations (fluorescence in situ hybridization, survivin, and telomerase) have led to some improvement in sensitivity of cytological methods.

None of the pre-existing biomarker approaches provide the specificity and sensitivity characteristics necessary for use in screening the general population for TCC. One goal of recent mass spectrometry-based proteomics studies of urine is to discover and validate new biomarkers or ensembles of biomarkers that have better specificity and sensitivity characteristics than existing biomarker assays or cytology. Vlahou et al. (40) used SELDI-TOF-MS to identify polypeptides in urine samples from TCC patients. Mass-to-charge ratios for five potential TCC biomarkers were identified. One of these was subsequently identified as a member of the defensin family, a group of endogenous antibiotic peptides that are locally secreted in response to inflammation. Another SELDI-TOF-MS study (61) demonstrated an additional set of polypeptides (reported as m/z ratios) associated with bladder cancer that when used in a machine learning-based computer algorithm gave 80% sensitivity and 90–97% specificity in a validation set of patient samples. Proteomics investigation of TCC tissue using 2-D electrophoresis and immunoblotting led to the recognition that adipocyte-type fatty acid-binding protein is expressed in markedly decreased amounts in TCC compared with normal transitional epithelium, and it has been proposed as an indicator of the degree of progression of the tumor (67, 68), although adipocyte-type fatty acid-binding protein has not been taken forward as a potential urinary biomarker for clinical use. Subsequent studies using a similar 2-D electrophoresis approach identified orosomucoid and zinc-α₂-glycoprotein as potential tumor markers; both are increased in urine of patients with TCC (35). Interestingly the abundance of these markers in urine increased with the stage of the tumors. In another study (34), the authors carried out 2-D electrophoresis of urine samples from TCC patients and control samples. These samples had been treated with gelatin beads to affinity purify the gelatin binding domain-containing proteins and to remove high abundance proteins that hinder the detection of clinically valuable low abundance proteins in the patient samples. These studies demonstrated that the amounts of several proteins increased in TCC including matrix metalloproteinase-2, metalloproteinase-9, and fibronectin. Interestingly in addition to changes in amounts of these proteins, their distributions on the 2-D gels were altered presumably due to differences in both glycosylation and proteolytic cleavage. Such data support the idea that changes in these and other post-translational modifications may provide useful information for identification of patients with early stage (superficial) tumors. Two other studies (62, 63) using 2-D electrophoresis and MALDI-TOF analysis of tissue identified three other potential TCC biomarkers, viz. calreticulin, γ-synuclein, and a soluble isoform of catechol O-methyltransferase, all of which were subsequently identified in urine samples immunochemically.

URINARY PROTEOMICS IN NEPHROLOGY AND RENAL TRANSPLANTATION

Nephrologists and urologists have begun to use mass spectrometry-based methods for discovery of potential disease biomarkers for a number of renal diseases including diabetic glomerulopathy (14, 16, 37, 69), acute renal injury (22), IgA nephropathy (15), Fanconi syndrome of various causes (29, 31, 70), idiopathic glomerulopathy (13), congenital obstructive nephropathy (39, 71), vesicoureteral reflux (72), interstitial cystitis (73), and urinary bladder reconstruction (74).
A particularly interesting finding was that three pathologically different forms of idiopathic glomerular disease (minimal change disease, membranous nephropathy, and focal segmental glomerulosclerosis) can be discriminated from each other based on the pattern of small proteins and peptides in the urine using a CE-MS approach (13). In general, these applications to nephrology and urology represent relatively early stages of the discovery process and herald future progress as specific biomarker candidates are taken forward for validation studies.

The detection of acute renal allograft rejection has been an important target for investigation (17, 18, 20, 21, 75). Biomarker discovery studies have been conducted with the view toward identification and validation of biomarkers of acute rejection that could detect impending rejection prior to a rise in serum creatinine and obviate or reduce the use of needle biopsy in diagnosis of acute rejection. The earliest such study used the SELDI-TOF-MS approach to identify a number of polypeptide peaks that either increased or decreased in association with acute rejection (18). This was followed by an additional SELDI-TOF-MS study that demonstrated a different set of m/z peaks that are associated with acute rejection (21). Another study that utilized SELDI-TOF-MS identified an additional set of polypeptide peaks that correlated with the presence of acute allograft rejection (20). These peaks were later identified using microcapillary LC-MALDI-MS/MS as proteolytic products of β2-microglobulin, presumably reflective of damaged renal proximal tubule cells (75). It remains uncertain whether these cleaved products are specific for acute rejection or might be seen under other circumstances associated with proximal tubule damage such as perioperative ischemic acute renal injury, injury associated with calcineurin inhibitors, or poliovirus type BK nephropathy. Recent studies using CE-MS have also found clear-cut differences in urinary polypeptide profiles in acute rejection compared with control transplant patients without rejection (17).

In general, mass spectrometry-based biomarker discovery studies have yet to yield a set of biomarkers that have found practical use in early detection of acute allograft rejection. However, this area holds great promise for progress as emphasized in all of the top-down studies reported thus far. The field would benefit from an attempt to identify additional specific proteins whose excretion rates are altered in association with acute allograft rejection perhaps using quantitative bottom-up approaches involving either 2-D electrophoresis or quantitative LC-MS/MS.

**CONCLUSION**

A major point of this brief review is that, although there are many barriers to success with mass spectrometry-based urinary biomarker discovery and validation, the field holds great promise. Investigators have a growing set of tools that can be used to identify holoproteins in urine as well as proteolytically derived peptides that are especially relevant in cancer diagnosis. A key area of current progress is the development of better means of quantification using LC-MS/MS that can be coupled to identification of specific proteins. Although there are several areas of need for better urinary biomarkers, some examples of particular areas of promise are: 1) discovery and validation of urinary biomarkers that can be used to classify patients with essential hypertension with regard to prediction of effective drug combinations, 2) discovery and validation of urinary biomarkers that will allow facile diagnosis in patients with salt-losing nephropathies, 3) discovery and validation of urinary biomarkers that will discriminate whether neonates with congenital obstructive nephropathy require surgical intervention or will resolve the obstruction spontaneously, and 4) discovery and validation of urinary biomarkers that can be used in oncology both with regard to cancers of the urinary tract as well as malignancies elsewhere in the body.

*The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.*

¶ To whom correspondence should be addressed: National Institutes of Health, Bldg. 10, Rm. 6N260, 10 Center Dr., MSC 1603, Bethesda, MD 20892-1603. Tel.: 301-496-3064; Fax: 301-402-1443; E-mail: knep@helix.nih.gov.

**REFERENCES**


Discovery of Urinary Biomarkers

37. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K.,


