Published in final edited form as:

J Urol. 2009 September ; 182(3): 835–843. doi:10.1016/j.juro.2009.05.001.

Proteomics and Opportunities for Clinical Translation in Urologic Disease

Ali R. Vaezzadeh^{1,2}, Hanno Steen^{2,3}, Michael R. Freeman¹, and Richard S. Lee^{1,2}

¹Department of Urology, Children's Hospital Boston, Harvard Medical School, Boston, MA, USA

²Proteomics Center at Children's Hospital Boston, Boston, MA, USA

³Department of Pathology, Children's Hospital Boston, Harvard Medical School, Boston, MA, USA

Abstract

Purpose—Proteomics is a rapidly growing new discipline that has the potential to increase and improve the understanding of protein function and interaction in the context of systems biology. As a translational science it has the potential to identify many new therapeutic targets and diagnostic and prognostic biomarkers of disease. Proteomics approaches consist of a combination of powerful technologies: protein/peptide separation, identification and bioinformatic detection and quantitation based on powerful computational data processing tools. In this article, we present an overview of current proteomics technologies, a review of proteomics applications in urology and a perspective on the future of proteomics in clinical medicine.

Materials and Methods—A literature search was performed on the basic concepts of proteomics and technologies commonly used in this field. Advantages, challenges and limitations of current proteomics approaches were discussed. Proteomics applications in the field of urology were presented.

Results—The proteomic approaches to answer clinical questions have only recently been introduced. Many different technologies have been employed in this field, which is moving from simple description into quantitative clinical applications.

Conclusions—Proteomics offers new approaches to the study of diseases of the genitourinary tract and the potential to identify clinically relevant biomarkers and new therapeutic targets.

Keywords

Mass Spectrometry; Liquid Chromatography; Biomarker; Urinary Proteomics

Proteomics: A new era

Proteomics is broadly defined as a collection of scientific approaches and technologies to characterize the protein content and their modifications within cells, tissues, body fluids and whole organisms at a particular stage or condition. Our level of understanding the

complexities of protein expression, function and detection has matured with advancing technologies and improved bioinformatic processing of protein pathways and networks¹. In principle, “-omics” technologies are aimed at profiling the entire pattern of information of a particular system. Proteomics is a logical continuation of the widely used, but technically unrelated, methodology of transcriptional profiling. The ultimate goal of proteomics is to analyze all the proteins in a particular sample, including their splice variants and modifications, and to then determine their biological and clinical relevance.

There is a great difference in complexity between proteome analysis and genome analysis. The human genome encodes approximately 20,500 genes². Because of variations arising from RNA splicing, most genes encode more than one protein, and these variants can be further diversified by Post-Translational Modifications (PTM), such as phosphorylation, glycosylation and ubiquitination. This diversity may give rise to a several fold increase in the potential number of expressed proteins at each gene locus. Although the dynamic nature of the proteome provides ample justification for studying diseases directly at the protein level, addressing this complexity is technically challenging. First, the estimated dynamic range of mammalian proteomes spans over 8–10 orders of magnitude and even more in body fluids such as blood and/or serum. Second, current mass spectrometry based proteomics approaches can only cover a dynamic range of up to 4 orders of magnitude. Third, highly abundant proteins often mask less abundant proteins in the sample; because no protein equivalent of DNA amplification exists; thus making less abundant species, such as most important cellular regulators, more difficult to detect. Finally, the proteome may vary because of external factors that arise from the environment, such as diet and medicines.

Even with these challenges, the discipline of proteomics has evolved dramatically since its early days, and is showing exponential growth in terms of numbers of publications and technology advancement (Figure 1). The remarkable progress of proteomic research over the last two decades has been catalyzed by several developments. First, the wealth of DNA sequence information generated by large scale genome sequencing projects, and the development of computational gene prediction and annotation tools, have produced sequence databases that are expected to contain most coding gene regions. Second, technological improvements in peptide/protein separation techniques and mass spectrometry allow rapid and sensitive protein identification from minute amounts of complex biological samples. Third, the development of computational/bioinformatics tools, such as the assignment and statistical validation of tandem mass (MS/MS) spectra to peptide sequences, permit the consistent analysis of large datasets with minimal human intervention. With the combination of powerful bioinformatics resources and high-throughput mass spectrometry, proteomics has the potential to identify, characterize and quantify thousands of proteins directly from complex biological samples in different disease states.

Proteomics technologies

The choice of proteomics technique will mostly be governed by the biological question asked. Some of the most commonly used proteomics methods are shown in Figure 2. Because of the notably different properties of proteins and peptides, two major approaches have evolved for proteome analysis. The approach differs depending on the stage at which

the intact protein or protein fragments are separated and identified. In the “top-down” approach, fractionation of the proteins occurs at a very early stage³. In this method proteins are directly introduced into the mass spectrometer, providing a high specificity for studying site-specific mutations and post-translational modifications. “Bottom-up” proteomics analysis involves proteolytic digestion of the proteins immediately after their isolation, which is also called the “shotgun” approach⁴. Protein digestion is performed using specific enzymes (proteases), such as trypsin.

Today, mass spectrometry (MS) based proteomics is a universal technique for analyzing complex protein samples¹. A mass spectrometer consists of three main components: an ion source (to generate ionized species), a mass analyzer (to measure mass-to-charge ratio (m/z)) and a detector (to count the number of ions at each m/z). To make separations possible, electric or magnetic fields are generated inside the instrument. The effects of an electric field on the ion movement are inversely proportional to the mass of the ion and directly proportional to its electrical charge. The mass spectrum of a compound consists of a plot of ion abundance versus its m/z ratio, from which the MW can be deduced.

The two most commonly used “soft” ionization techniques are Electrospray Ionization (ESI)⁵ and Matrix-Assisted Laser Desorption Ionization (MALDI)⁶. ESI ionizes the sample from a liquid solution and is frequently combined with Liquid Chromatography (LC) based separations. MALDI sublimates and ionizes a sample that has been crystallized with a matrix. Comparison of ESI and MALDI has revealed that these two strategies are complementary and their combination results in increased proteome coverage⁷.

Ions from either source are then analyzed by a single or most often a combination of mass analyzers such as Time Of Flight (TOF), Ion Trap (IT), Quadrupole (Q), Fourier Transform Ion Cyclotron Resonance (FT-ICR), and Orbitrap. Since these instruments differ considerably in sensitivity and mass accuracy (Figure 3), the choice of the mass spectrometer depends on the biological question. For an extensive review see the article by Glish⁸.

If more detailed amino acid sequence or the site/type of a post-translational modification (PTM) is needed, tandem mass spectrometry (MS^n) is performed. In these instruments, two mass analyzers are used in series. The first usually selects the ion of interest, which is then fragmented, and the second analyzes the fragments. MS instruments that use different types of analyzers for the first and second stages of mass analysis in MS/MS experiments are often called “hybrid” instruments (e.g. Orbitrap).

Surface Enhanced Laser Desorption Ionization (SELDI)-TOF MS is a variant of MALDI-MS, which employs ProteinChip arrays to selectively enrich subset of proteins prior to direct profiling of the bound proteins. Although the technology is relatively easy to use, SELDI is not a quantitative approach and is not capable of direct identification of proteins⁹. In addition, although many proteins readily bind to the chip’s surface, only small proteins (< 20kDa) are efficiently detected.

Protein / Peptide Separation

Due to the complexity of the protein/peptide mixtures, upfront separation is critical to achieve appropriate analytical depth into the proteome. Proteins or peptides are separated using their different physicochemical properties. For example, the molecular volume is used in Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis (SDS-PAGE); or the charge or Isoelectric Point (pI) is used in IsoElectric Focusing (IEF). Liquid chromatography is performed as a function of hydrophobicity, ionic charge, affinity and size.

Typically, one dimension of separation is not sufficient to decrease the proteome complexity. Therefore, two or more independent properties are employed to separate the mixtures into individual components. When these properties are truly independent, the separation method is considered “orthogonal” and the peak capacity is approximately equal to the product of the individual peak capacities of each dimension. One of the common orthogonal methods of separation is two-dimensional gel electrophoresis (2-DE), which combines IEF and SDS-PAGE methods¹⁰. Differential in-gel electrophoresis (2D-DIGE) is a recent improvement of the 2-DE technology¹⁰. In this method, proteins from different disease states are separately labeled with different fluorescent cyanine dyes, and an internal pooled standard is labeled with another dye. The labeled samples are then combined and subjected to 2-DE. The gel is scanned at different emission wavelengths, generating multiple images that can be overlaid and compared to identify differences between samples. Features of interest on the gel are then excised, digested and analyzed by MS. Overall; this method improves reproducibility, minimizes alignment variation and potentially allows quantitative comparison between samples.

Despite the longstanding success of 2-DE, there are several fundamental limitations, including the difficulty of identifying low-abundance proteins, membrane proteins, and proteins with extremes in isoelectric point and molecular weight. These challenges led to the development of alternative, “gel-free” approaches. Yates and colleagues developed an approach called Multi-Dimensional Protein Identification Technology (MudPIT)⁴. This was the first description of analyzing a complex peptide mixture by combining Strong Cation Exchange (SCX), which separates peptides based upon their charge, and Reverse Phase (RP) chromatography, along with tandem mass spectrometry (SCX/RP/MS/MS). In comparison to gel-based separation methods, sample handling and preparation are minimal in LC. Other methods for separation at the protein or peptide level that may be used in combination with a RP step prior to MS include: high pH RP, Ion Pair (IP), Hydrophobic Interaction Chromatography (HIC), Weak Anion Exchange (WAX), Strong Anion Exchange (SAX), and immobilized pH gradient (IPG)-IEF.

Although integration of extensive multidimensional fractionation/separations with MS greatly increases the overall analytical dynamic range of proteome coverage, extensive fractionation results in decreased throughput. To date, most reports involving extensive fractionation have been limited to small scale studies of one or two pooled clinical samples rather than larger scale comparative quantitative studies.

Protein identification

Identifying proteins by measuring the masses of their proteolytic peptides by MS is called Peptide Mass Fingerprinting (PMF)¹¹. Protein identification is performed by correlating the measured mass values with theoretical masses computed from protein sequences stored in databases. Closely related to PMF, is the Peptide Fragmentation Fingerprint (PFF) which allows high-confident protein identification. MS/MS provides additional experimental sequence information for each identified peptide, which is then compared to an *in-silico* model MS/MS database of theoretical peptide fragments. The degree of similarity between the experimental spectra and the modeled ones is measured and a probability score that represents the statistical likelihood of a correct match is assigned¹².

Recent improvements in the resolving power of microcapillary separations and the improved sensitivity of mass spectrometers are such that a single MudPIT analysis can produce 10⁴ MS/MS spectra. Since manual validation of the spectrum-to-peptide assignments is not feasible, strategies have been developed to validate the data. One common approach is to estimate the rate of false positive peptide identification by employing a randomized (reversed or reshuffled) sequence database, also known as “decoy” database. This method estimates the number of false-positive identifications by counting the number of peptide matches to a randomized sequence database using the same search criteria. Another method is to use several identification tools to choose the common assignments. Finally, a statistical method utilizes machine learning techniques to calculate a “correct assignment” probability based on peptides properties and identification score¹³. For an extensive review, see the article by Hernandez¹².

Quantitative proteomics

In many aspects of proteomics research, such as biomarker discovery, protein identification must be accompanied by quantitative information. Recently, significant progress has been made in the development of absolute and relative protein quantitation methods. Most methods of quantitation involve labeling the study sample with either a heavy isotope or chemical tag, and then comparing that to a control sample that is labeled with a natural isotope or different chemical tag. Samples are mixed in equal proportions; the relative signal for each peptide can be measured by MS; and then the peak intensities between the two isotopes or tags can be quantified. There are several possibilities for sample labeling:

I. Metabolic labeling

Stable Isotope Labeling with Amino acids in Cell culture (SILAC) is a robust quantitation strategy primarily used in cell culture¹⁴. In SILAC, heavy amino acids (¹⁵N or ¹³C atoms), such as Leu, Arg, and Lys are incorporated during cell culture and label virtually every peptide of the study sample, where as the natural light amino acid isotope labels the control. This technique is only applicable to cultured cells and small animal models (through feeding) and requires a careful setup of the model system and the growth media.

II. Chemical labeling

Isotope Coded Affinity Tag (ICAT)¹⁵ was the first commercially available labeling reagent, but has been superseded by other strategies, such as iTRAQ¹⁶. In this method, as opposed to SILAC, labeling only occurs at the peptide level after digestion of the proteins has been completed. In iTRAQ, isotopic labels are attached to the primary amino group of every peptide through reactive chemical compounds. The advantage of iTRAQ is that up to eight different labels may be used, thereby allowing multiplexed quantitative comparisons between various samples. The major disadvantages are labeling efficiency, potential changes in chromatography, and side reactions creating unexpected products.

III. Enzymatic labeling

¹⁸O atoms are incorporated into the C-terminal side of cleaved peptides¹⁷. This approach has the advantage of being a sequence-independent labeling approach; allowing analysis of any peptide. Drawbacks include the small mass difference between labeled and unlabeled samples; and that full labeling is rarely achieved resulting in complicated data analysis.

IV. Label free

This method requires no additional experimental steps and utilizes only the measured data from the mass spectrometer. Relationships between measured data and sampling statistics, such as sequence coverage, peptide count, ion intensity, and spectral count, are used to calculate relative abundance¹⁸. In contrast to other methods, label free methods are fast, inexpensive, unlimited by experimentation, and do not result in complex spectra comparisons. However, label free methods may not be as accurate for lower abundant peptides as compared to labeling methods, and may have difficulty separating quantitation of multiple protein isoforms.

V. Absolute quantitation

Absolute quantitation of proteins (AQUA) is a method in which particular peptide elements of a sample are compared against a spiked-in isotopic labeled synthetic analogue¹⁹. This method focuses the quantitation to only one or a few particular proteins of interest. This approach has been refined by constructing synthetic genes that express concatenated standard peptides, which upon tryptic digestion either provide multiple peptides of the same protein for quantitation, or quantitation standards for a group of proteins of interest. The disadvantages are difficulty in determining the amount of labeled standard to use and the cost of AQUA.

Post-Translational Modifications

Although quantitative differences in protein identification are critical, further analysis of protein PTMs may be necessary to attain a comprehensive understanding of overall function. These modifications can change the solubility, stability, localization and biological activity of proteins. Phosphorylation is one of the most common PTMs and plays an important role in regulating cell physiology²⁰. Phosphorylated proteins and peptides are generally present at low levels; therefore, enrichment prior to mass analysis is necessary for identification. Immobilized-Metal Affinity Chromatography (IMAC) can preferentially isolate

phosphorylated proteins using a phosphate group's high affinity for a cation. Titanium dioxide (TiO₂) is an additional phosphopeptide enrichment methodology.

Glycosylation is another common form of PTM. It is estimated that more than 50% of all proteins are glycosylated. Glycosylation plays a role in protein sorting, receptor binding, immune recognition, inflammation, and pathogenicity. As in phosphoproteomics, glycosylated proteins often require enrichment steps to improve identification. Lectin affinity chromatography is frequently used to enrich for glycoproteins. Due to the enormous variety of free and protein-bound glycans expressed in cells, analysis of glycoproteins is difficult as compared to other modifications. For an extensive review on these two important PTMs, see the review by Temporini.

For an extensive review on these two important PTMs, see the review by Temporini²¹.

Proteomics and Biomarkers

Subtle changes in either the quantitative expression level or a PTM may reflect a pathological process. These changes suggest the potential role of proteins as biomarkers of disease. Biomarkers may be useful for accurate diagnosis and prognosis of disease, gauging the efficacy and safety of novel therapeutic agents, and as an aid in drug development. Although initially promising as a technology to discover biomarkers, large scale comparative proteomics studies have not yet produced markers of clinical significance. This is most probably due to the low concentration of relevant biomarkers compared to interfering, highly abundant proteins in samples, inadequate study design, limitations in quantitation, bioinformatic redundancy, and PTM variability. Advances in the field have improved on some of these limitations, such as the advent of the newest ion trap instrumentation for PTM identification. Additionally, new rigorous standards towards protein identification, study design, cohort analysis, and statistical rigor are now required for validation and publication of results²².

The identification of clinically relevant disease markers is difficult and requires exacting study design and statistical analysis. For these reasons, it is recommended that each proteomics study, especially for biomarker discovery, should have two distinctive steps: discovery and validation²³. Given the current status of measurement reproducibility and lack of standardization of calibration, many researchers use proteomics in the discovery mode. They view proteomics as a rapid screening tool for generating new hypotheses. Candidate proteins are then selected for further evaluation using more traditional techniques, such as ELISA assays upon availability of specific antibodies²⁴.

Targeted or directed proteomics, however, may also have significant potential as a validation method. Targeted proteomics involves quantitative MS to measure a preset generated list of candidate target proteins²⁵. The targeted proteomics strategy often provides greater sensitivity and allows for detection of low-abundant candidate proteins. Anderson and Hunter recently demonstrated the use of Multiple Reaction Monitoring (MRM) for quantitative assaying peptides derived from major plasma proteins²⁵. Targeted MS analysis using MRM enhances the lower detection limit for peptides by up to 100-fold by allowing rapid and continuous monitoring exclusively for the specific ions of interest. Such MRM

assays provide great specificity for peptide/protein identifications and relatively good precision for quantitation. Additionally, MRM can provide a rapid and specific platform for biomarker validation, particularly when coupled with specific enrichment techniques.

Proteomics and Urology

The application of proteomics toward urologic questions has great potential. Urologic disease is ideally situated to benefit from proteomic applications as indicated by the initiation of Human Kidney and Urine Proteome Project (HKUPP) by Human Proteome Organization (HUPO)²⁶. One particular interesting application of proteomics is interrogation of the urinary proteome for urologic disease specific biomarkers. Urine may be an ideal sample for proteomics because it is readily and noninvasively available and is stable²⁷. In addition, urine has the *specific advantage* of being the proximal fluid (primary effluent) to the majority of organs and diseases of interest in the urogenital tract, which is critical to improve biomarker discovery efforts²⁷.

However, urine contains a variable amount of proteins and excreted compounds such as salts and urea, making sample handling and analysis extremely challenging. Because of the variability of the urinary proteome, there are many different methods for urinary protein isolation including: protein precipitation, lyophilization, ultracentrifugation and centrifugal filtration^{27, 28}. Other sample preparation issues such as the use of pooled vs. individual samples, effect of high-abundant protein depletion, removal of interfering compounds and cellular debris, effect of prolonged exposure to room temperature, effect of repetitive freeze-thaw events have been thoroughly studied^{27, 28}. Overall, these technical improvements will help guide standardization of sample handling in the near future²⁶.

Recent urine studies have demonstrated that proteins originate from both plasma via glomerular filtration or tubular secretion, and also from the various organs in the urinary tract^{29–30}. Various techniques have been used to interrogate the urinary proteome and to date the largest number of proteins (or fragments) identified in the adult urinary proteome is 1500³¹. Despite these advances, the consistency and reproducibility of protein identification in the urine remains poor³². This may be secondary to changes in protein content from metabolic or catabolic processes, hydration, diurnal variation, individual discrepancies, specimen acquisition, sample handling, and/or depth and reproducibility of proteomic profiling²⁷.

Benign Urologic Disease

To date there have only been a few studies of the urinary proteome with a Urologic focus. Lee *et al.* used SDS-PAGE and nano-ESI-LC-MS/MS techniques to demonstrate the temporal variations in the urinary proteome during postnatal maturation in the rat³³. They noted dramatic changes in the normal postnatal urinary proteome secondary to physiologic changes related to nephrogenesis and developmental maturation, and documented a set of proteins as potential markers of normal development³³. This preliminary work may provide critical insight into renal development and allow for future comparative studies in rat animal models of urologic disease.

With regard to current clinical urologic applications, proteomics have been used to identify urinary markers for non-cancerous renal disease and various genitourinary cancers. One of the larger areas of urinary proteomic biomarker discovery has been in renal transplant and allograft nephropathy. The majority of the initial discovery work has been conducted using SELDI or CE-MS platforms. The different identified urinary markers may be useful in identifying acute and subacute transplant rejection^{34, 35}; however, the variability between and within each study may reflect the multiple confounding variables present in each patient cohort. Furthermore, no validation studies have been reported to date. Similarly, proteomics has been used to potentially differentiate between various chronic kidney diseases such as IgA nephropathy, diabetic nephropathy, focal-segmental glomerulosclerosis, and others^{36–38}.

Obstructive nephropathy (UPJ) has also been interrogated by CE-MS based urinary proteome analysis. Using a cohort of 36 patients, discovery based methods identified a potential polypeptide pattern that may help determine which children require surgery³⁹. Further prospective validation work is needed to determine the potential of these markers to determine which children will improve without surgery.

Malignant Urologic Disease

Various attempts have been made to identify urinary biomarkers for genitourinary carcinoma. Analysis of patients with renal cell carcinoma and normal controls using SELDI-TOF initially had promising results (sensitivity and specificity 81.8–83.3%). However, further validation studies demonstrated a significant decrease (41.0–76.6%). Similarly discovery studies of the urinary proteome in prostate cancer using CE-MS initially yielded potentially promising results; however they could not be validated in larger studies⁴⁰.

Numerous urinary biomarkers of bladder cancer have been studied by multiple groups using various proteomic methods; however a large number of them have not been tested in validation studies^{41–44}. For instance, urinary microparticles (exosomes), which are small membrane bound proteins that are released from various cells in the urinary tract such as bladder carcinoma cells, were isolated from healthy controls and bladder cancer patients and compared at the proteome level using LC-MS/MS. Eight candidate proteins were identified as potential markers, but require further clinical validation to determine the clinical significance⁴¹.

Of those that have undergone some method of validation, the specificity and sensitivity have ranged and may show some potential in clinical practice. As an example, 2D-DIGE was used as a discovery method to identify regenerating protein-1 (Reg-1), a regulator of apoptosis, and was then validated using traditional biochemical methods (sensitivity 81.3%, and specificity 81.3%). SELDI-TOF profiling patterns were also used to detect bladder cancer but the sensitivity and specificity was poor and may have been a function of the variability of the methods. Finally, using capillary Electrophoresis (CE)-MS, have also shown very promising results in this field³¹. Wittke *et al.* used CE-MS to analyze urinary samples in 19 patients with different grades of subclinical or clinical acute rejection after renal transplantation, 10 patients with urinary tract infection and 29 patients without evidence of rejection or infection. A distinct urinary polypeptide pattern identified 16 out of 17 cases of acute tubulointerstitial rejection, but this pattern was absent in two cases of vascular

rejection⁴⁵. Similarly, CE-MS a polypeptide pattern specific for a subset of urinary peptides was identified and then validated in a larger mixed cohort for a sensitivity of 86–100%. In summary, although many potentially promising results have been identified in both benign and malignant genitourinary disease, larger validation studies that are multiplexed with other potential markers are needed to determine their relevance to clinical practice.

Other potential applications

Outside of the urinary proteome, whole tissue sections have also been successfully interrogated using Imaging MS (IMS), which allows direct profiling of the spatial distribution of proteins/peptides without target specific labels or reagents⁴⁶. Caprioli's group used a combination of Laser Capture Microdissection (LCM) with MALDI-TOF to study the kidney in the context of a glomerulosclerotic rat model⁴⁷. They reported that thymosin β 4 was identified as one of several key differentiators between nonsclerotic and sclerotic glomeruli, suggesting its role as an early activator and indicator of a pro-sclerotic mechanism. Other applications of IMS in urology include the study of gentamicin toxicity biomarkers in rat kidney⁴⁸, comparison of normal prostate development and maturation protein patterns with prostate tumors derived from genetically engineered mice⁴⁹, and others.

Quantitative proteomic techniques have also been extensively utilized in urological studies. SILAC was applied to compare the protein expression levels in microsomal fractions from the highly metastatic PC3M-LN4 and its relatively poorly metastatic PC3M cell line⁵⁰. Around 1000 proteins were identified, out of which 444 proteins were quantified. Of these 22, were found to be down-regulated in the highly metastatic cells, whereas 60 were found to show elevated expression. ICAT was employed to evaluate Platelet-Derived Growth Factor (PDGF)-induced protein changes in lipid rafts of primary culture human bladder smooth muscle cells⁵¹. iTRAQ was used in a mouse model of Autosomal-Dominant Polycystic Kidney Disease (ADPKD)⁵². By characterizing kidney cell plasma membrane from wild-type versus PKD1 knockout mice, 791 proteins were quantified and 67 and 37 proteins showed more than 2-fold up-regulation and down-regulation, respectively. Label free quantitation was used to compare human metastatic prostate cancer cell line, in which Matrix Metalloproteinase-9 (MMP-9) had been down-regulated by RNA interference, with that from the parental cells. From more than 200 proteins identified, 69 showed significant alteration in levels after depletion of the protease. A strong correlation between results by immunoblotting and by LC-MS gave credence to the label free approach⁵³.

Proteomics perspectives

Although proteomics has demonstrated its promise for biomarker discovery, further work is still required to enhance the performance and reproducibility of established proteomics tools before they can be routinely used in clinical medicine. Issues regarding pre-analytical variables, analytical inconsistency and biological variation must be tackled. The standardization of methodologies and dissemination of proteomics data into publicly available databases is becoming commonplace and will result in higher quality results. Improvements in data and bioinformatics tools, such as enhanced mass accuracy and resolution, faster protein identification algorithms, reduced false positive rate and automated data validation strategies, will allow the assimilation of results across multiple studies and

platforms, thus improving the depth of knowledge on a particular question. With these improvements, ambitious goals, such as the creation of a full proteome map of tissues or cell types, will be achieved more readily. Apart from analytical improvements, proteomics should be integrated with genomics, transcriptomics and metabolomics to reach the ultimate aim of systems biology, which is to further the understanding of complex biological systems. Proteomics should not, therefore, be regarded separately, but rather as an important element of systems biology.

Acknowledgments

This study was supported by NIH P50DK65298 (MRF), NIH K08DK077836 (RSL), and a Children's Hospital Boston Career Development Fellowship (RSL). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003; 422:198. [PubMed: 12634793]
2. Pennisi E. Genetics. Working the (gene count) numbers: finally, a firm answer? *Science*. 2007; 316:1113. [PubMed: 17525311]
3. Kelleher NL. Top-down proteomics. *Anal Chem*. 2004; 76:197A. [PubMed: 14697051]
4. Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, et al. Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol*. 1999; 17:676. [PubMed: 10404161]
5. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. Electrospray ionization for mass spectrometry of large biomolecules. *Science*. 1989; 246:64. [PubMed: 2675315]
6. Karas M, Hillenkamp F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry*. 1988; 60:2299. [PubMed: 3239801]
7. Bodnar WM, Blackburn RK, Krise JM, Moseley MA. Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J Am Soc Mass Spectrom*. 2003; 14:971. [PubMed: 12954165]
8. Glish GL, Burinsky DJ. Hybrid mass spectrometers for tandem mass spectrometry. *J Am Soc Mass Spectrom*. 2008; 19:161. [PubMed: 18187337]
9. Poon TC. Opportunities and limitations of SELDI-TOF-MS in biomedical research: practical advices. *Expert Rev Proteomics*. 2007; 4:51. [PubMed: 17288515]
10. Issaq H, Veenstra T. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): advances and perspectives. *Biotechniques*. 2008; 44:697. [PubMed: 18474047]
11. Pappin D, Hojrup P, Bleasby A. Rapid identification of proteins by peptide-mass fingerprinting. *Current Biology*. 1993; 3:327. [PubMed: 15335725]
12. Hernandez P, Muller M, Appel RD. Automated protein identification by tandem mass spectrometry: issues and strategies. *Mass Spectrom Rev*. 2006; 25:235. [PubMed: 16284939]
13. Nesvizhskii AI, Aebersold R. Interpretation of shotgun proteomic data: the protein inference problem. *Mol Cell Proteomics*. 2005; 4:1419. [PubMed: 16009968]
14. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics*. 2002; 1:376. [PubMed: 12118079]
15. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol*. 1999; 17:994. [PubMed: 10504701]
16. Ross PL, Huang YN, Marchese J, Williamson B, Parker K, Hattan S, et al. Multiplexed protein quantitation in *saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics*. 2004; 3:1154. [PubMed: 15385600]

17. Yao X, Afonso C, Fenselau C. Dissection of proteolytic ¹⁸O labeling: endoprotease-catalyzed ¹⁶O-to-¹⁸O exchange of truncated peptide substrates. *J Proteome Res.* 2003; 2:147. [PubMed: 12716128]
18. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinisky JR, et al. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics.* 2005; 4:1487. [PubMed: 15979981]
19. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A.* 2003; 100:6940. [PubMed: 12771378]
20. Collins MO, Yu L, Choudhary JS. Analysis of protein phosphorylation on a proteome-scale. *Proteomics.* 2007; 7:2751. [PubMed: 17703509]
21. Temporini C, Calleri E, Massolini G, Caccialanza G. Integrated analytical strategies for the study of phosphorylation and glycosylation in proteins. *Mass Spectrom Rev.* 2008; 27:207. [PubMed: 18335498]
22. Orchard S, Hermjakob H, Julian RK Jr, Runte K, Sherman D, Wojcik J, et al. Common interchange standards for proteomics data: Public availability of tools and schema. *Proteomics.* 2004; 4:490. [PubMed: 14760721]
23. Lescuyer P, Hochstrasser D, Rabilloud T. How shall we use the proteomics toolbox for biomarker discovery? *J Proteome Res.* 2007; 6:3371. [PubMed: 17655344]
24. Howard BA, Wang MZ, Campa MJ, Corro C, Fitzgerald MC, Patz EF Jr. Identification and validation of a potential lung cancer serum biomarker detected by matrix-assisted laser desorption/ionization-time of flight spectra analysis. *Proteomics.* 2003; 3:1720. [PubMed: 12973732]
25. Anderson L, Hunter CL. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol Cell Proteomics.* 2006; 5:573. [PubMed: 16332733]
26. Yamamoto T, Langham RG, Ronco P, Knepper MA, Thongboonkerd V. 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.* 2008; 8:2156. [PubMed: 18528840]
27. Lee RS, Monigatti F, Briscoe AC, Waldon Z, Freeman MR, Steen H. Optimizing Sample Handling for Urinary Proteomics. *J Proteome Res.* 2008; 7:4022. [PubMed: 18662026]
28. Thongboonkerd V. Practical points in urinary proteomics. *J Proteome Res.* 2007; 6:3881. [PubMed: 17824635]
29. Spahr CS, Davis MT, McGinley MD, Robinson JH, Bures EJ, Beierle J, et al. Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. *Proteomics.* 2001; 1:93. [PubMed: 11680902]
30. Castagna A, Ceconi D, Sennels L, Rappsilber J, Guerrier L, Fortis F, et al. Exploring the hidden human urinary proteome via ligand library beads. *J Proteome Res.* 2005; 4:1917. [PubMed: 16335936]
31. Adachi J, Kumar C, Zhang Y, Olsen JV, Mann M. The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. *Genome Biol.* 2006; 7:R80. [PubMed: 16948836]
32. Sun W, Li F, Wu S, Wang X, Zheng D, Wang J, et al. Human urine proteome analysis by three separation approaches. *Proteomics.* 2005; 5:4994. [PubMed: 16281181]
33. Lee RS, Monigatti F, Lutchman M, Patterson T, Budnik B, Steen JA, et al. Temporal variations of the postnatal rat urinary proteome as a reflection of systemic maturation. *Proteomics.* 2008; 8:1097. [PubMed: 18324733]
34. O'Riordan E, Orlova TN, Podust VN, Chander PN, Yanagi S, Nakazato M, et al. Characterization of urinary peptide biomarkers of acute rejection in renal allografts. *Am J Transplant.* 2007; 7:930. [PubMed: 17331118]
35. Clarke W, Silverman BC, Zhang Z, Chan DW, Klein AS, Molmenti EP. Characterization of renal allograft rejection by urinary proteomic analysis. *Ann Surg.* 2003; 237:660. [PubMed: 12724632]
36. Haubitz M, Wittke S, Weissinger EM, Walden M, Rupperecht HD, Floege J, et al. Urine protein patterns can serve as diagnostic tools in patients with IgA nephropathy. *Kidney Int.* 2005; 67:2313. [PubMed: 15882273]

37. Schiffer E, Mischak H, Novak J. High resolution proteome/peptidome analysis of body fluids by capillary electrophoresis coupled with MS. *Proteomics*. 2006; 6:5615. [PubMed: 16991199]
38. Weissinger EM, Wittke S, Kaiser T, Haller H, Bartel S, Krebs R, et al. Proteomic patterns established with capillary electrophoresis and mass spectrometry for diagnostic purposes. *Kidney Int*. 2004; 65:2426. [PubMed: 15149356]
39. Decramer S, Wittke S, Mischak H, Zurbig P, Walden M, Bouissou F, et al. Predicting the clinical outcome of congenital unilateral ureteropelvic junction obstruction in newborn by urinary proteome analysis. *Nat Med*. 2006; 12:398. [PubMed: 16550189]
40. Theodorescu D. Cryotherapy for prostate cancer: what we know, what we need to know. *J Urol*. 2008; 180:437. [PubMed: 18550127]
41. Smalley DM, Sheman NE, Nelson K, Theodorescu D. Isolation and identification of potential urinary microparticle biomarkers of bladder cancer. *J Proteome Res*. 2008; 7:2088. [PubMed: 18373357]
42. Saito M, Kimoto M, Araki T, Shimada Y, Fujii R, Oofusa K, et al. Proteome analysis of gelatin-bound urinary proteins from patients with bladder cancers. *Eur Urol*. 2005; 48:865. [PubMed: 15964125]
43. Irmak S, Tilki D, Heukeshoven J, Oliveira-Ferrer L, Friedrich M, Huland H, et al. Stage-dependent increase of orosomucoid and zinc-alpha2-glycoprotein in urinary bladder cancer. *Proteomics*. 2005; 5:4296. [PubMed: 16196100]
44. Celis JE, Gromova I, Moreira JM, Cabezon T, Gromov P. Impact of proteomics on bladder cancer research. *Pharmacogenomics*. 2004; 5:381. [PubMed: 15165174]
45. Wittke S, Haubitz M, Walden M, Rohde F, Schwarz A, Mengel M, et al. Detection of acute tubulointerstitial rejection by proteomic analysis of urinary samples in renal transplant recipients. *Am J Transplant*. 2005; 5:2479. [PubMed: 16162198]
46. McDonnell LA, Heeren RM. Imaging mass spectrometry. *Mass Spectrom Rev*. 2007; 26:606. [PubMed: 17471576]
47. Xu BJ, Shyr Y, Liang X, Ma LJ, Donnert EM, Roberts JD, et al. Proteomic patterns and prediction of glomerulosclerosis and its mechanisms. *J Am Soc Nephrol*. 2005; 16:2967. [PubMed: 16079267]
48. Meistermann H, Norris JL, Aerni HR, Cornett DS, Friedlein A, Erskine AR, et al. Biomarker discovery by imaging mass spectrometry: transthyretin is a biomarker for gentamicin-induced nephrotoxicity in rat. *Mol Cell Proteomics*. 2006; 5:1876. [PubMed: 16705188]
49. Chaurand P, Rahman MA, Hunt T, Mobley JA, Gu G, Latham JC, et al. Monitoring mouse prostate development by profiling and imaging mass spectrometry. *Mol Cell Proteomics*. 2008; 7:411. [PubMed: 17991918]
50. Everley PA, Krijgsveld J, Zetter BR, Gygi SP. Quantitative cancer proteomics: stable isotope labeling with amino acids in cell culture (SILAC) as a tool for prostate cancer research. *Mol Cell Proteomics*. 2004; 3:729. [PubMed: 15102926]
51. MacLellan DL, Steen H, Adam RM, Garlick M, Zurakowski D, Gygi SP, et al. A quantitative proteomic analysis of growth factor-induced compositional changes in lipid rafts of human smooth muscle cells. *Proteomics*. 2005; 5:4733. [PubMed: 16267816]
52. Han CL, Chien CW, Chen WC, Chen YR, Wu CP, Li H, et al. A multiplexed quantitative strategy for membrane proteomics: Opportunities for mining therapeutic targets for autosomal-dominant polycystic kidney disease. *Mol Cell Proteomics*. 2008
53. Xu D, Suenaga N, Edelmann MJ, Fridman R, Muschel RJ, Kessler BM. Novel MMP-9 substrates in cancer cells revealed by a label-free quantitative proteomics approach. *Mol Cell Proteomics*. 2008; 7:2215. [PubMed: 18596065]

Figure 1. Number of research articles published in “PubMed” during 2000–2007 with **A.** Proteomics in their title or as keyword and **B.** Number of research articles on Urinary proteomics published during 2000–2007.

Figure 2.

Current Proteomics workflows. Samples, in form of body fluids/tissues/cells are prepared by extraction, LCM or cell culture. Highly abundant proteins can be depleted or specific PTMs can be enriched. Proteins are then separated according to their different physicochemical properties in one or several dimensions. This step is usually followed by protein identification by mass spectrometry. In the shotgun approach, proteins are first digested and then analyzed by LC-MS/MS and identified using bioinformatics tools. SELDI, protein/

peptide microarrays and imaging mass spectrometry are among alternative approaches to the two previously mentioned mainstream techniques.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Mass Analyzer	Mass Accuracy	Mass Resolution	Sensitivity	Cost
Time-Of-Flight (TOF)	10-100 PPM	20000	++++	\$\$
Quadrupole (Q)	100-1000 PPM	2000	++	\$
Quadrupole Ion Trap (LCQ)	100-1000 PPM	10000	+++	\$
Linear Quadrupole Trap (LTQ)	100-500 PPM	10000	++++	\$\$
Fourier Transform	0.1-2 PPM	1000000	++	\$\$\$\$
Orbitrap	0.1-10 PPM	100000	+++	\$\$\$

Figure 3.
Characteristics and performances of commonly used mass analyzers.

Table 1

Advantages, Disadvantages and Limitations of Urinary Biomarkers in Urologic Disease

Advantages	Disadvantages	Limitations
<ul style="list-style-type: none">• Noninvasive nature of sample collection• Readily available• Proximal fluid (primary effluent) to the diseases of interest in the urogenital tract, which may enhance biomarker discovery• Improved high-throughput methods for sample analysis that potentially allow for rapid screening as compared to conventional biochemical methods	<ul style="list-style-type: none">• Difficulty of sample handling because of urinary salts and buffers (e.g. urea) limits easy application of proven proteomic techniques for high-throughput discovery methods.• Inherent low protein concentration of urine• High-abundant proteins (e.g. albumin) may lead to suppression effects on less abundant proteins• Proteins of interest are most likely low abundant proteins	<ul style="list-style-type: none">• Variability in sample collection and handling• Requires exacting cohort development and study design• Normal variability of age, gender, time of day and disease state may create artificial variations in the urinary proteome• Large size and depth of urinary proteome may still be beyond the current capabilities of proteomic technologies• Currently, proteomics may be more suited for discovery and not high-throughput validation