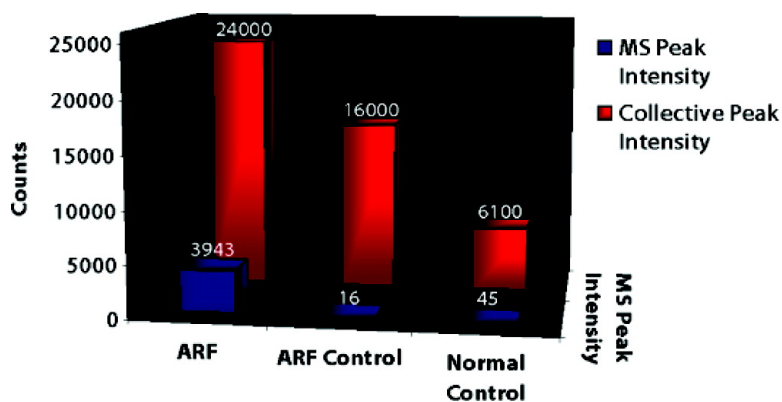


## Body Fluid Proteomics for Biomarker Discovery: Lessons from the Past Hold the Key to Success in the Future

David M. Good, Visith Thongboonkerd, Jan Novak, Jean-Loup Bascands, Joost P. Schanstra, Joshua J. Coon, Anna Dominiczak, and Harald Mischak

*J. Proteome Res.*, **2007**, 6 (12), 4549-4555 • DOI: 10.1021/pr070529w • Publication Date (Web): 31 October 2007

Downloaded from <http://pubs.acs.org> on March 16, 2009



### More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 9 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)



## Body Fluid Proteomics for Biomarker Discovery: Lessons from the Past Hold the Key to Success in the Future

David M. Good,<sup>†</sup> Visith Thongboonkerd,<sup>‡</sup> Jan Novak,<sup>§</sup> Jean-Loup Bascands,<sup>||,⊥</sup>  
Joost P. Schanstra,<sup>||,⊥</sup> Joshua J. Coon,<sup>†,∇</sup> Anna Dominiczak,<sup>○</sup> and Harald Mischak<sup>\*,¶,♯</sup>

*Department of Chemistry, University of Wisconsin—Madison, Madison, Wisconsin 53706, Medical Molecular Biology Unit, Office for Research and Development, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand, Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294, Inserm, U858/I2MR, Department of Renal and Cardiac Remodelling, Team #5, 1 Avenue Jean Poulhès, BP 84225, 31432 Toulouse Cedex 4, France, Université Toulouse III Paul Sabatier, Institut de Médecine Moléculaire de Rangueil, Toulouse, F-31000 France, Department of Biomolecular Chemistry, University of Wisconsin—Madison, Madison, Wisconsin 53706, BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, U.K., and Mosaiques Diagnostics, GmbH, Hannover, Germany*

Received August 14, 2007

Sparked by the article from Lescuyer and colleagues in a recent issue, we aim here to further encourage interest in and discussion of clinically relevant biomarker research. We express our view on proteomics for biomarker discovery by addressing multiple relevant issues, including the inherent differences between biological fluids (and how these differences affect current analytical approaches) and experimental design to maximize the efficiency of moving from the bench to the bedside. Herein, we also include suggestions for definition of the term “biomarker”, based on the use of a set of universal characterization/validation requirements, and illustrate several recent examples of successful transitions of benchtop proteomic studies work to clinical practice.

**Keywords:** Biomarker • Mass spectrometry • Blood • Urine • Validation • Diagnosis • Prognosis

### Introduction

As recently alluded to by Lescuyer et al.,<sup>1</sup> the field of proteomics and, more specifically, the employment of proteomics-based approaches for identification, characterization, and/or validation of biomarkers indicative of disease state has seen rapid expansion in recent years. As we all know, this has led to high hopes, premature claims of enormous success, and consequently, to a harsh landing. However, we should put these errors from the past to good use for present and future studies. It is neither advisable nor wise to now simply dismiss everything that was supposedly correct only a few years ago, and start completely from scratch. In other words, while proteomics of body fluids has not achieved the end goal of improving on the clinical situation through proteome analysis with the previously anticipated ease, it has clearly not been without numerous advancements. Especially when aiming toward realistic goals, several good, scientifically sound, and overall

positive studies have been reported,<sup>2–4</sup> and we will give several further examples below. In addition, we expect that numerous additional studies have been completed, but are still pending publication.

On the basis of these facts, we do not share the opinion of Lescuyer and colleagues<sup>1</sup> that the current proteomic technologies generally do not result in the discovery of useful biomarkers. Recently, we and several other groups demonstrated that certain biofluids, especially urine and cerebrospinal fluid (CSF), are excellent sources of biomarkers for their corresponding organs, inspection of these organs' physiology and pathophysiology, and consequently, also for clinical use.<sup>5–8</sup>

As outlined previously,<sup>9</sup> biomarker discovery for clinical purposes requires separate discovery and validation phases: an early discovery phase (generation of potential biomarkers), testing and characterization in a larger independent cohort, and subsequent validation in a blinded manner on samples from subjects with a particular disease(s), healthy controls, and disease controls, preferably in a multicenter study. However, the validation and subsequent clinical application of these biomarkers does not necessarily exclude mass spectrometry-based techniques and also does not necessitate ELISA techniques for validation. These and additional views are outlined below.

### Prerequisites

As outlined in more detail recently<sup>9</sup> and also below, several prerequisites appear mandatory for a meaningful application

\* Corresponding author: Harald Mischak. Telephone: +49-511-554744-15. Fax: +49-511-554744-31. E-mail: [mischak@mosaiques-diagnostics.com](mailto:mischak@mosaiques-diagnostics.com).

<sup>†</sup> Department of Chemistry, University of Wisconsin—Madison.

<sup>‡</sup> Mahidol University.

<sup>§</sup> University of Alabama at Birmingham.

<sup>||</sup> Inserm, U858/I2MR, Department of Renal and Cardiac Remodelling.

<sup>⊥</sup> Université Toulouse III Paul Sabatier.

<sup>∇</sup> Department of Biomolecular Chemistry, University of Wisconsin—Madison.

<sup>○</sup> University of Glasgow.

<sup>¶</sup> Mosaiques Diagnostics.

<sup>♯</sup> On behalf of the Urosysteomics consortium.

of what we call clinical proteomics or biomarker discovery. These aspects need to be outlined at the beginning of a study, including:

1. Definition of a clear clinical need and the resulting aim of the study
2. Definition of the sample source, both in the initial phase and also in the application
3. Assessment of reproducibility and comparability of the analysis platform (including assessment of variability of the sample material and a proper sampling protocol)
4. Protocols for quality control, data handling, and statistical evaluation

If these simple steps had been recognized and implemented a few years ago, then most of the studies that subsequently resulted in little success or lack of reproducibility would probably never have been performed, as these were generally deficient in several of the key elements (e.g., very high variability of the sample and low sample number will never result in a statistically sound result). This is especially true for blood-derived samples: the assessment of the variability of serum or plasma samples drawn from the same person at different time-points indicated extremely large variability of such samples.<sup>10</sup>

Consequently, we share Lescuyer and colleagues' view<sup>1</sup> that it is difficult to obtain unbiased detection of useful biomarkers in plasma or serum in general (given the current technological state), although these body fluids may be quite useful in specific situations. These difficulties may be due to large dynamic range of analytes (i.e., serum proteins),<sup>11</sup> sample variability, and the lack of a publicly accessible database that would allow comparison of proteomic data from independent studies.

These thoughts also pinpoint one of the largest problems in clinical proteomics: the requirement of large sample size to reach sufficient power of the analysis. This requirement is mostly due to individual and biological variability and becomes evident when applying proper statistical algorithms by expert biostatisticians as the situation warrants. Biological (or physiological) variability offers only limited options for control (e.g., collection of samples at the same time of day), and consequently, it is imperative to perform large studies; only this approach can help to distinguish between molecules that merely reflect changes due to biological (or physiological) variability and "true biomarkers". Additionally, dynamic changes of the proteomes due to disease staging and progression and the effects of treatment should be also considered. The application of a simple *t* test should be viewed as erroneous and thus avoided, with adjustment due to the Bonferroni method,<sup>12</sup> at the very least, being imperative to identify relevant biomarkers. An even more advantageous way to evaluate the validity of potential biomarkers is the application of a distribution-free resampling test, such as the maxT test.<sup>13</sup>

With the proper application of statistical approaches and tests, it becomes evident that the number of independent samples required for the identification of significant biomarkers frequently exceeds 100. As it appears inefficient to repeatedly analyze quite similar samples in different studies, a database that can provide data on such samples (mostly controls) would be highly beneficial and reduce the costs and time required for a study. In addition, inclusion of a large number of comparable samples would also add to the validity of any study. Such a database requires deposition of metadata that would allow the assessment of reproducibility, accuracy of measurement, and biological variability, as outlined in a recent paper by Taylor et al.<sup>14</sup> in more detail. Only a strict attention to such

metadata would allow meaningful comparison of the data sets. Because of the high biological variability, meaningful assessment of these metadata may be difficult to accomplish when examining blood-derived biological fluids and even more complicated, at the current stage of technology, when using tissue for analyses. We hope that future advancements of mass spectrometry-based technologies, such as tissue imaging, may open new horizons for these types of analyses using tissue.<sup>15–19</sup> On the other hand and as we outline in more detail below, some body fluids like urine and CSF appear to not be riddled to this large extent with high variability, and thus provide a much more promising target.

## Body Fluids

Much focus has recently been placed on biological fluids as sources of biomarkers diagnostic of certain disease states,<sup>20</sup> with other studies placing emphasis on prognosis of disease and/or patient response to treatment regimens.<sup>21–24</sup> Biological fluids have garnered much attention as possible sources for biomarkers when compared to the use of tissue owing to several factors, including ease of accessibility, avoiding risks of invasive tissue sampling through biopsies, relative low cost of obtainment, availability of monitoring based on multiple sampling, and the potential for development of large-scale, valuable prognostic/diagnostic tests.

It is important to point out the differences between using biological fluids and tissues for biomarker analysis. A number of potential problems are associated with tissue analysis; these include difficulty with obtaining samples, standardization in the light of different cell types, or presence of various proteolytic enzymes. Some hurdles may yet be unknown due to the current limited availability of data on these factors. Also, certain biomarkers may be significant in tissue, but not very useful in body fluids, for example, specific proteolytic fragments can display activity of matrix metallo-proteases (MMPs). These fragments may be present in fluid, but not in tissue, while MMPs are present in tissue, but not in fluid.

Several recent papers have illustrated the power of mass spectrometry-based detection methods in the proteome analysis of human biological fluids.<sup>20,25,26</sup> It was concluded that body fluid proteomes are most likely to contain the largest number of clinically useful disease biomarkers. Because of these findings, as well as the current lack of data on using tissues for biomarker extraction, we will focus herein on the major biological fluids currently employed for these analyses and will summarize the individual advantages and disadvantages in the context of biomarker discovery. Each biological fluid is unique and therefore brings with it its own set of challenges.

Perhaps the most common biological fluids in use today for biomarker analysis are blood (serum and plasma), CSF, and urine. Of these, blood certainly has the longest track record for use in modern diagnostic procedures.<sup>27,28</sup> Blood is a potential source of a plethora of biomarkers, as it comes into contact with all organs of the body and, thus, gives the theoretical possibility to see the whole spectrum of different biomarkers. This, in turn, means that biomarkers found in the blood are often less specific and can generally not be confined to emanating from a single source/organ. In addition to the above-mentioned problems associated with blood, both serum and plasma also suffer from broad dynamic range issues,<sup>11</sup> forcing the complete removal of many predominant proteins to detect minor proteins and reliably and reproducibly separate/fractionate the proteome.<sup>29,30</sup> Such steps introduce additional

variability through the necessary sample preparation steps. Furthermore, the depletion of abundant species from plasma may lead to a loss of information. Depletion of albumin was associated with the reduction of 815 other proteins,<sup>31</sup> suggesting a significant loss of information by using these procedures. The complexity of issues in blood for clinical investigations was recently demonstrated in an article which investigated the anti-inflammatory properties of high density lipoprotein (HDL).<sup>32</sup> The authors noted the presence of unwanted proteins even after extensive sample cleanup, including ultracentrifugation and affinity chromatography. This work illustrated the difficulties of selecting for a specific group of proteins, as multiple analytical techniques led to identification of different associated protein classes. This exemplifies the intricate network of protein-protein interactions within the blood and shows that isolation of a specific, yet complete, subset of these related protein networks cannot be easily achieved using a single preparative or analytical method.

Urine and CSF are fluids very amenable to biomarker analysis.<sup>5,6</sup> CSF is quite specific to the central nervous system (CNS), contains less total protein than serum or plasma, and provides a low fluid-volume-to-organ ratio, thus, increasing the probability of high concentrations of released proteins or peptides within the fluid. Furthermore, transportation of fluids from the spinal cord and brain to the CSF is achieved in a bulk flow manner.<sup>33</sup> This means that molecules of varying sizes have an equal opportunity to be found within the CSF. Unfortunately, obtaining CSF involves a lumbar puncture (LP), which may not be possible for all patients with a CNS disease (e.g., due to contraindications), and is generally not applicable to apparently healthy individuals.

Urine has no such constraints, as it can usually be obtained by strictly noninvasive procedures, and its excretion is a normal and necessary biological function. Like CSF, urine is also more organ-specific than blood, generally containing proteins from only a few organs that are located directly along the path of production and excretion (i.e., the kidney and urinary tract, including the bladder). However, normal urinary proteins can also originate from glomerular filtration of plasma proteins (the glomerular barrier only restricts passage of larger proteins, small to middle-molecular-weight proteins can still pass through), secretion of proteins from renal tubular epithelial cells, shedding of whole cells along urinary passage and of apical membranes of renal tubular epithelial cells, and also from exosome secretion.<sup>7,34,35</sup> Approximately 70% of all urinary proteins originate from the kidney, whereas only 30% are derived from the plasma.<sup>7</sup> Disruption of the glomerular barrier and/or tubular injury can result in an increased proportion of the plasma proteins in the urine. In stark contrast to CSF, urine can be obtained in very large volumes (dL to L amounts) and can be collected over a period of time (even offering the ability for collection of multiple samples within 24 h), allowing for greater robustness and the possibility of using this fluid to monitor patient therapeutic response over very short windows.<sup>22,32,34-36</sup>

In addition to the ease of obtainment and subsequent analysis outlined above, there is a need for the biological fluid to be sufficiently stable for its future use in a clinical context. The sample must be sufficiently stable during the several steps of the procedure including sample collection, sample transportation from site of collection to site of analysis, sample storage, and preparation for analysis. Furthermore, the reproducible sampling of the body fluid at the bedside is also an

issue. Of the biological fluids we have addressed, blood is typically the least reproducible. This appears to be mostly due to blood-containing active proteases, whole cells, and lipids, all of which have to be removed or inactivated prior analysis. Furthermore, different sample types, serum and plasma, can be generated, utilizing different stabilizing agents.<sup>37</sup> Another problem affecting reproducibility when using blood is the presence of coagulation factors (in serum) which render the samples highly variable when using current analytical techniques. In contrast, the proteomes of both CSF and urine have been shown to be quite stable and thus give rise to much more reproducible data sets.<sup>4,38-40</sup>

These considerations underline that the choice of biological fluid for analysis is often one of the first, and also among the most important, decisions that must be made when designing a proteomics-based experiment for mining of clinically relevant biomarkers. Particular fluids cater to specific investigative needs, while also carrying with them their own set of analytical hurdles.

### **Biomarker Discovery: "From Bedside to Bench... And Back to the Bed"**

The move from discovery phase to the subsequent validation and characterization phase(s) is without doubt a difficult, but mandatory transition, as the sole purpose of the biomarker lies in its application. Here, we outline a suggested experimental design for clinically relevant biomarker investigations to better manage this transition.

The foremost suggestion for proteomic studies which aim to draw clinical conclusions is to follow the general guidelines of clinical trials. Furthermore, it is indispensable for a wide variety of researchers (clinicians, statisticians, bioinformaticians, biologists, and clinical biochemists) to be involved with the investigation from its earliest stages.<sup>9,26</sup> This will limit possible misunderstandings between groups with differing specialties and lead to a more clearly defined clinical question from the start. The aim of the study may be for diagnostics, prognostics, and/or prediction of treatment response. All of these questions require profound clinical knowledge and, in most cases, collection and assessment of different samples. A further benefit of a close cooperation of specialists from the different areas is that the result of the study will be based on statistically solid data obtained from sufficient numbers of patients and proper controls and, thus, results in more easily interpretable and clinically applicable data.

Along these lines, we disagree with the suggestion alluded to by Lescuyer and colleagues<sup>1</sup> that a "biomarker" has different meanings for specialists in differing fields. We propose that the term "biomarker" should have a universal meaning, with the only differences coming in terms of the extensiveness of the study performed. As recently outlined at the Grand Rounds in Proteomics at the FDA,<sup>41</sup> a polypeptide biomarker is a specific polypeptide used to measure or indicate the effects or progression of a disease, condition, or treatment. Consequently, the biomarker must also be defined by its intended use. It is important to be aware that a biomarker is not suitable beyond the intended use, unless proven otherwise in a study that clearly validates such a broader application. For example, a biomarker may be well-suited for advanced screening of disease at an early stage, while the same biomarker may not be appropriate for monitoring therapeutic effects, and vice versa.<sup>41</sup> A well-defined clinical proteomics study, framed within the context of searching for an answer to a specific clinical question, should limit

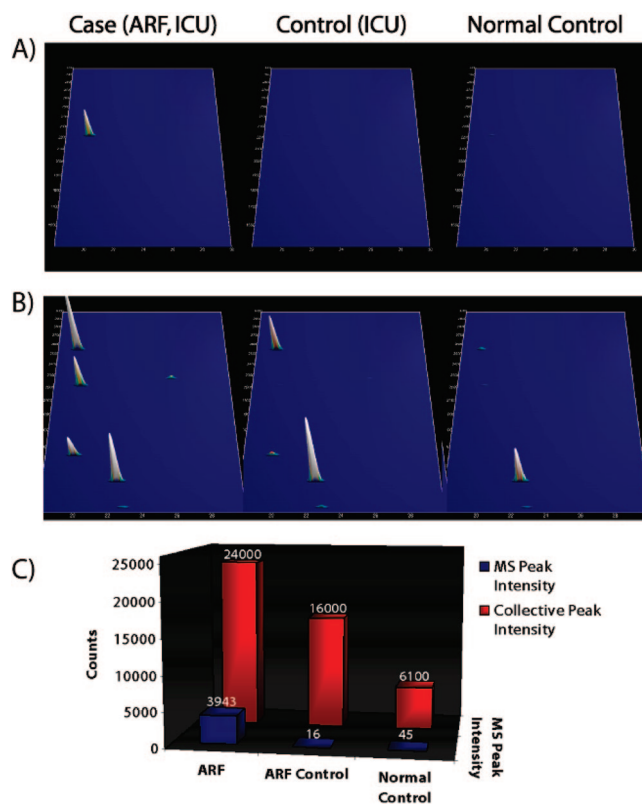
confusion between those performing the initial proteomic studies, clinical chemists, and/or clinicians and allow a clear definition of a biomarker.

We agree that a separate validation phase, aside from the initial discovery phase of the biomarkers, is required.<sup>1,9,26</sup> However, just because the two phases with different emphasis are necessary, this does not obligatorily lead to changes in the analytical techniques employed (i.e., mass spectrometry-based studies can be used for both phases; *vide infra*). The currently held common belief is that mass spectrometry-identified biomarkers can be easily transformed into a successful immuno-based assay, and that such an assay is required for clinical utilization. Below, we describe several reasons why this notion may not be feasible or even desirable.

The strict requirement for an enzyme-linked immunosorbent assay (ELISA) analysis for biomarker assessment in a clinical setting represents a myth or misconception, perhaps driven by the current clinical laboratory practice. In reality, antibodies recognize epitopes, and these can be shared across multiple peptides/proteins. One example of a possible drawback to converting mass spectrometry-derived diagnostics into an immunological test is illustrated in Figure 1. While a particular peptide fragment serves as an excellent biomarker for the detection of kidney failure in preliminary studies, an antibody that would recognize this sequence would also recognize several other peptides within the sample, due to the existence of a similar stretch of sequence. Consequently, the application of such an antibody in an immunological assay would no longer reveal a statistically significant difference between case and control, as the antibody could only detect the combined signal of all these (and possibly more) polypeptides. Thus, this type of assay would not provide useful data, unlike the MS analysis. Furthermore, several recent papers outline the successful use of mass spectrometry-based methods for quantifying specific proteins within complex mixtures.<sup>42,43</sup>

In addition to the difficulties associated with transferring MS-derived data to a working and validated immunological assay, it has been suggested that a single biomarker is of only limited value. The reason lies in the requirement for the existence of a marker peptide within ~100% samples from patients with the disease state of interest. Consequently, diagnostic patterns consisting of multiple peptides may be more advantageous. The diagnostic power of a test can be increased by incorporating multiple biomarkers into a panel so as to enhance the robustness of the diagnostic classification. Several recent studies detailed the use of capillary electrophoresis coupled with mass spectrometry (CE-MS) for both detection and validation of panels of disease-specific biomarkers.<sup>3,4</sup> Employment of these panels of biomarkers generally resulted in the specificity and sensitivity of >90%. The development of an immunological test which integrates multiple diagnostic features (i.e., a diagnostic polypeptide "pattern") would offset one of its main benefits: low cost. While mass producing ELISA-type immunological tests are certainly very cost-effective and such tests allow for high-throughput, this benefit would be reduced by targeting a large panel of peptides. Perhaps, new technologies, such as protein arrays, may change this situation in the future.

As outlined above, the ability of a validation method to observe and characterize post-translational modifications (PTM) is of utmost importance, and is another area where a traditional immuno-based assay may fail without an extensive and expensive development of specific and well-characterized monoclonal antibodies. If a PTM is important for the classification



**Figure 1.** Distribution of potential biomarkers for acute renal failure. Urine samples from patients at intensive care units (ICU) were prospectively collected and examined using capillary electrophoresis coupled with mass spectrometry (CE-MS). The data were subsequently grouped according to the occurrence of acute renal failure (ARF) to enable definition of biomarkers for detection of forthcoming ARF. (A) Average distribution of a biomarker indicative for ARF is shown (defined by mass and normalized migration time; peak height reflects average signal intensity), both in patients that subsequently underwent ARF as well as in control ICU patients and apparently healthy normal controls. Sequencing revealed that this potential biomarker shares an epitope with several other peptides. (B) The distribution of all these peptides is shown. As evident, inclusion of all these peptides is of no diagnostic value. (C) Distribution of the signal intensity for the single peak, or for the epitope, as it would be detected by immunological methods.

of a disease as indicated in numerous articles,<sup>42,44-48</sup> its identification and characterization is generally possible through multiple mass spectrometry-based methods. However, the use of an immuno-based assay would typically only allow the detection of either the modified or unmodified form, not the comparison of the abundance of each form when employing only a single antibody.

In addition to the aforementioned CE-MS method that offers the ability to combine the discovery phase and the validation phase using one mass spectrometry-based analytical technique, several other mass spectrometry methods can be employed for the characterization/validation phase of a clinical diagnostic test. Two of these methods include multiple reaction monitoring (MRM)<sup>49</sup> and stable isotope standards and capture by antipeptide antibodies (SISCAPA).<sup>50</sup> The SISCAPA method utilizes antibody-nanoaffinity columns to enrich for specific peptides, as well as stable-isotope-labeled internal standards of the same sequence which have been spiked into the sample. Although this demands prior sequence knowledge of peptides

of interest, the main benefit is the sensitivity of these tests and the ability to quantify results. In MRM, a triple quadrupole MS system can provide excellent sensitivity for quantitation of targeted peptides. The main advantages of this method are the absence of a special preworkup of sample—as long as the sample is compatible with MS, it is practical for this type of analysis—and also the increased ability to characterize species containing PTMs [as compared to using the traditional fragmentation method; collision-activated dissociation (CAD)].

As stated before, we argue that single-entity biomarkers are of limited value, but rather a group of biomarkers/diagnostic features together relate the story of the *in vivo* disease activity with high sensitivity and specificity. Because of its inability to readily deal with modified peptides/proteins, as well as its lack of cost-effectiveness when created for multiple epitopes of interest, ELISA-style immunological tests are not always the most advantageous method for clinical validation, without extensive development of specific and well-characterized antibodies. Therefore, the requirement of establishing an immunological assay such as an ELISA for making a diagnostic method “practical”, in a clinical sense, is strictly based upon the previously accepted clinical development scheme, and no longer applies in light of today’s analytical methods.

We do not imply that MS-based technologies may be well-suited to replace ELISA. On the contrary, a combination of both technologies may be the best advance toward solving yet unmet clinical needs. An ELISA, due to its low cost and ease of handling, may be the best choice to screen a large population for diseases. This should generally be performed aiming at high sensitivity and accepting low specificity (like the PSA assay). A subsequent multidimensional MS-based analysis of samples from potentially positive subjects may be an optimal approach.

### Studies Exemplifying the Potential of Proteomic-Based Body Fluid Biomarker Discovery

By following these guidelines, and investigating biological fluids more conducive to this research, the field of mass spectrometry-based proteomics has proven its worth for biomarker analyses. For example, the above-mentioned CE-MS proteomic technique was recently used in both the successful discovery and validation of biomarker panels in the diagnosis of urothelial cancer<sup>4</sup> and for the prediction of ureteropelvic junction obstruction in newborns.<sup>3</sup> Both studies, using urine samples, had a well-defined clinical question, used separate discovery and validation cohorts, and predicted disease outcome in blinded cohorts with high precision. Since these predictions are based on panels of biomarkers, the transformation, as mentioned above, into an immunoassay in the near future is highly unlikely. This introduces the dilemma of whether one should wait for new technological developments that allow immuno-based assays (if any) or make these kind of MS-based approaches available, after large-scale validation, to the patient.

Analyses of CSF have also proven fruitful for clinically relevant biomarker analyses. Hansson et al. identified a panel of peptides prognostic for Alzheimer’s disease (AD) development, using a combination of T-tau, A $\beta$ 42(A), and P-tau.<sup>51</sup> The use of these markers in combination led to a sensitivity of 95% and a specificity of 83% for detection of incipient AD in patients with mild cognitive impairment. Another recent study investigating diagnostic biomarkers for Alzheimer’s disease focused on synaptic markers and produced a diagnostic panel of polypeptides giving 91% sensitivity and 88% specificity after

utilization of support vector machine (SVM) analysis (Jahn et al., submitted). Upon further definition using a blinded evaluation, the pattern showed a sensitivity of 87% and a specificity of 79%. Though both studies were successful in finding multiple diagnostic markers for early indication of AD, Jahn et al. reported a sensitivity of 88% and specificity of 67% when using CSF measurements of those markers defined by Hansson et al. for their sample analysis. Such examples demonstrate the need to include multiple blinded groups within statistical analyses and also to perform hierarchical re-evaluations of data as the study moves on to increase the specificity of the resulting diagnostic parameters.

### Conclusion

We agree with Lescuyer and colleagues that blood-derived samples, with currently available proteomic technologies, are currently not well-suited for most disease-specific biomarker discovery studies as they suffer from the drawbacks outlined above. Other body fluids, such as CSF and, especially, urine, have already yielded robust results and may prove more specific to certain organs, thus, providing suitable media for biomarker discovery, and also for the use of validated markers in clinical application. Other, currently less exploited, body fluids such as bronchoalveolar fluid, synovial fluid, nipple aspirate fluid, saliva, or amniotic fluid have a similar potential.

A requirement for separate discovery and validation phases before clinical application is reasonable, but this does not inherently oblige the use of separate experimental techniques. Here, we have provided examples of how a blind dependence upon validation using the apparent “gold standard” resulted in data not suitable for clinical application, while the use of a carefully developed assay did, in fact, produce statistically significant and easily interpretable results. The frequently mentioned need to display biomarkers using a cheap, immuno-based assay is not based on solid scientific reasoning, but rather on traditional use of this assay in clinical laboratories. Some biomarkers just cannot be displayed using such technology, and the fact that an assay may cost several hundred dollars does not limit its benefit. This is especially true in light of current diagnostic standards, where, in some instances, cost exceeds \$1000 per test, and are in part invasive and risk-associated procedures, which can even result in death of the patient [e.g., see ref 52]. Without doubt, a noninvasive procedure at similar or even lower cost is certainly an advantage, and there should be no impetus in this case for the mandatory development of an ELISA prior to clinical use. One example of an instance in which this may prove to be the case is in determining the necessity for surgical relief in ureteropelvic junction (UPJ) obstruction in newborns. The common clinical practice involves the use of various radiologic investigations such as diuretic ultrasonography and/or excretory urography. These radiologic investigations expose these infants to radiation and may need injection of radiocontrast or radioisotope material and necessitate hospitalization. These cost-related and sociological burden greatly exceed the costs for a MS-based urine analysis.<sup>53</sup> Other examples, where no acceptable diagnostic procedure exists or the current procedures are associated with high costs and risk, and where MS-based biomarker assessment is available, were recently described for graft-versus-host disease<sup>54</sup> and IgA-associated glomerular disease, such as IgA nephropathy and hepatitis C-associated immune complex glomerulonephritis.<sup>55</sup>

In conclusion, mass spectrometry-based proteomics can, and already does, provide clinically relevant results for biomarker discovery, given the selection of an appropriate fluid for analysis, the use of proper technology for performing analyses, and validation of results using appropriate clinical studies.

**Acknowledgment.** D.M.G. was supported through a predoctoral fellowship from the National Institutes of Health; the Biotechnology Training Program, NIH 5T32GM08349. V.T. was supported by Mahidol University, The Thailand Research Fund, and the National Center for Genetic Engineering and Biotechnology. J.N. is supported in part by NIH grants DK78244, DK61525, DK71802, DK47322, and DK64400. J.P.S. was supported by Inserm, the “Direction Régionale de la Recherche” (CHU de Toulouse, France) under the Interface programme and by the Fondation pour la Recherche Médicale. J.J.C. acknowledges the University of Wisconsin—Madison, Thermo Fisher, the Beckman Foundation, and the NIH (1R01GM080148) for financial support. A.D. was supported by the British Foundation Chair (CH98001) and Programme (RG/02/012) and the InGenious HyperCare EU Funding (LSHM-C7-2006-037093). H.M. was supported in part by EUROTRANS-BIO grant ETB-2006-016 and EU Funding through InGenious HyperCare (LSHM-C7-2006-037093) and PREDICTIONS.

## References

- Lescuyer, P.; Hochstrasser, D.; Rabilloud, T. *J. Proteome Res.* **2007**, *6*, 3371–3376.
- Roessler, M.; Rollinger, W.; Mantovani-Endl, L.; Hagmann, M. L.; Palme, S.; Berndt, P.; Engel, A. M.; Pfeffer, M.; Karl, J.; Bodenmueller, H.; Ruschoff, J.; Henkel, T.; Rohr, G.; Rossol, S.; Rosch, W.; Langen, H.; Zolg, W.; Tacke, M. *Mol. Cell. Proteomics* **2006**, *5*, 2092–2101.
- Decramer, S.; Wittke, S.; Mischak, H.; Zurbig, P.; Walden, M.; Bouissou, F.; Bascands, J. L.; Schanstra, J. P. *Nat. Med.* **2006**, *12*, 398–400.
- Theodorescu, D.; Wittke, S.; Ross, M. M.; Walden, M.; Conaway, M.; Just, I.; Mischak, H.; Frierson, H. F. *Lancet Oncol.* **2006**, *7*, 230–240.
- Fliser, D.; Novak, J.; Thongboonkerd, V.; Argiles, A.; Jankowski, V.; Girolami, M.; Jankowski, J.; Mischak, H. *J. Am. Soc. Nephrol.* **2007**, *18*, 1057–1071.
- Zhang, J. *Proteomics Clin. Appl.* **2007**, *1*, 805–819.
- Thongboonkerd, V. *Proteomics Clin. Appl.* **2007**, *1*, 780–791.
- Mischak, H.; Julian, B. A.; Novak, J. *Proteomics Clin. Appl.* **2007**, *1*, 792–804.
- Mischak, H.; Apweiler, R.; Banks, R. E.; Conaway, M.; Coon, J. J.; Dominzak, A.; Ehrlich, J. H.; Fliser, D.; Girolami, M.; Hermjakob, H.; Hochstrasser, D. F.; Jankowski, V.; Julian, B. A.; Kolch, W.; Massy, Z.; Neususs, C.; Novak, J.; Peter, K.; Rossing, K.; Schanstra, J. P.; Semmes, O. J.; Theodorescu, D.; Thongboonkerd, V.; Weissinger, E. M.; Van Eyk, J. E.; Yamamoto, T. *Proteomics Clin. Appl.* **2007**, *1*, 148–156.
- Kolch, W.; Neususs, C.; Pelzing, M.; Mischak, H. *Mass Spectrom Rev.* **2005**, *24*, 959–977.
- Anderson, N. L.; Anderson, N. G. *Mol. Cell. Proteomics* **2002**, *1*, 845–867.
- Abdi, H. *Bonferroni and Sidak Corrections for Multiple Comparisons*; Sage: Thousand Oaks, CA, 2007.
- Westfall, P. H.; Young, S. S. *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment*; Wiley: New York, 1993.
- Taylor, C. F.; Paton, N. W.; Lilley, K. S.; Binz, P. A.; Julian, R. K., Jr.; Jones, A. R.; Zhu, W.; Apweiler, R.; Aebersold, R.; Deutsch, E. W.; Dunn, M. J.; Heck, A. J.; Leitner, A.; Macht, M.; Mann, M.; Martens, L.; Neubert, T. A.; Patterson, S. D.; Ping, P.; Seymour, S. L.; Souda, P.; Tsugita, A.; Vandekerckhove, J.; Vondriska, T. M.; Whitelegge, J. P.; Wilkins, M. R.; Xenarios, I.; Yates, J. R., III; Hermjakob, H. *Nat. Biotechnol.* **2007**, *25*, 887–893.
- Altelaar, A. F.; Luxembourg, S. L.; McDonnell, L. A.; Piersma, S. R.; Heeren, R. M. *Nat. Protoc.* **2007**, *2*, 1185–1196.
- Norris, J. L.; Cornett, D. S.; Mobley, J. A.; Andersson, M.; Seeley, E. H.; Chaurand, P.; Caprioli, R. M. *Int. J. Mass Spectrom.* **2007**, *260*, 212–221.
- Simonson, A. B.; Schnitzer, J. E. *J. Thromb. Haemost.* **2007**, *5* (Suppl. 1), 183–187.
- McLean, J. A.; Ridenour, W. B.; Caprioli, R. M. *J. Mass Spectrom.* **2007**, *42*, 1099–1105.
- Jang, W.; Kim, K. P.; Schwob, J. E. *J. Comp. Neurol.* **2007**, *502*, 367–381.
- Hu, S.; Loo, J. A.; Wong, D. T. *Proteomics* **2006**, *6*, 6326–6353.
- Elkind, M. S.; Tai, W.; Coates, K.; Paik, M. C.; Sacco, R. L. *Arch. Intern. Med.* **2006**, *166*, 2073–2080.
- Rossing, K.; Mischak, H.; Parving, H. H.; Christensen, P. K.; Walden, M.; Hillmann, M.; Kaiser, T. *Kidney Int.* **2005**, *68*, 193–205.
- Cicenas, J.; Urban, P.; Vuaroqueaux, V.; Labuhn, M.; Kung, W.; Wight, E.; Mayhew, M.; Eppenberger, U.; Eppenberger-Castori, S. *Breast Cancer Res.* **2005**, *7*, R394–R401.
- Cicenas, J.; Urban, P.; Kung, W.; Vuaroqueaux, V.; Labuhn, M.; Wight, E.; Eppenberger, U.; Eppenberger-Castori, S. *Eur. J. Cancer* **2006**, *42*, 636–645.
- Schmidt, A.; Aebersold, R. *Genome Biology* **2006**, *7*, 242.
- deVera, I. E.; Katz, J. E.; Agus, D. B. *Clin. Adv. Hematol. Oncol.* **2006**, *4*, 541–549.
- Ludwig, J. A.; Weinstein, J. N. *Nat. Rev. Cancer* **2005**, *5*, 845–856.
- Etzioni, R.; Urban, N.; Ramsey, S.; McIntosh, M.; Schwartz, S.; Reid, B.; Radich, J.; Anderson, G.; Hartwell, L. *Nat. Rev. Cancer* **2003**, *3*, 243–252.
- Whiteaker, J. R.; Zhang, H.; Eng, J. K.; Fang, R.; Piening, B. D.; Feng, L. C.; Lorentzen, T. D.; Schoenherr, R. M.; Keane, J. F.; Holzmann, T.; Fitzgibbon, M.; Lin, C.; Zhang, H.; Cooke, K.; Liu, T.; Camp, D. G.; Anderson, L.; Watts, J.; Smith, R. D.; McIntosh, M. W.; Paulovich, A. G. *J. Proteome Res.* **2007**, *6*, 828–836.
- Hoffman, S. A.; Joo, W. A.; Echan, L. A.; Speicher, D. W. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2007**, *849*, 43–52.
- Shen, Y.; Kim, J.; Strittmatter, E. F.; Jacobs, J. M.; Camp, D. G.; Fang, R.; Tolle, N.; Moore, R. J.; Smith, R. D. *Proteomics* **2005**, *5*, 4034–4045.
- Vaisar, T.; Pennathur, S.; Green, P. S.; Gharib, S. A.; Hoofnagle, A. N.; Cheung, M. C.; Byun, J.; Vuletic, S.; Kassim, S.; Singh, P.; Chea, H.; Knopp, R. H.; Brunzell, J.; Geary, R.; Chait, A.; Zhao, X. Q.; Elkon, K.; Marcovina, S.; Ridker, P.; Oram, J. F.; Heinecke, J. W. *J. Clin. Invest.* **2007**, *117*, 746–756.
- Cserr, H. F.; Cooper, D. N.; Suri, P. K.; Patlak, C. S. *Am. J. Physiol.* **1981**, *240*, F319–F328.
- Pisitkun, T.; Johnstone, R.; Knepper, M. A. *Mol. Cell. Proteomics* **2006**, *5*, 1760–1771.
- Hoorn, E. J.; Pisitkun, T.; Zietse, R.; Gross, P.; Frokiaer, J.; Wang, N. S.; Gonzales, P. A.; Star, R. A.; Knepper, M. A. *Nephrology* **2005**, *10*, 283–290.
- Ngai, H. H.; Sit, W. H.; Jiang, P. P.; Thongboonkerd, V.; Wan, J. M. *J. Proteome Res.* **2007**, *6*, 3313–3320.
- Omenn, G. S.; States, D. J.; Adamski, M.; Blackwell, T. W.; Menon, R.; Hermjakob, H.; Apweiler, R.; Haab, B. B.; Simpson, R. J.; Eddes, J. S.; Kapp, E. A.; Moritz, R. L.; Chan, D. W.; Rai, A. J.; Admon, A.; Aebersold, R.; Eng, J.; Hancock, W. S.; Hefta, S. A.; Meyer, H.; Paik, Y. K.; Yoo, J. S.; Ping, P.; Pounds, J.; Adkins, J.; Qian, X.; Wang, R.; Wasinger, V.; Wu, C. Y.; Zhao, X.; Zeng, R.; Archakov, A.; Tsugita, A.; Beer, I.; Pandey, A.; Pisano, M.; Andrews, P.; Tammen, H.; Speicher, D. W.; Hanash, S. M. *Proteomics* **2005**, *5*, 3226–3245.
- Schaub, S.; Wilkins, J.; Weiler, T.; Sangster, K.; Rush, D.; Nickerson, P. *Kidney Int.* **2004**, *65*, 323–332.
- Fiedler, G. M.; Baumann, S.; Leichle, A.; Oltmann, A.; Kase, J.; Thiery, J.; Ceglarek, U. *Clin. Chem.* **2007**, *53*, 421–428.
- Kaiser, E.; Schonknecht, P.; Thomann, P. A.; Hunt, A.; Schroder, J. *Neurosci. Lett.* **2007**, *417*, 193–195.
- Goodsaid, F.; Bandow, J. E.; Mischak, H. *Proteomics* **2007**, in press.
- Lu, Y.; Bottari, P.; Aebersold, R.; Turecek, F.; Gelb, M. H. *Methods Mol. Biol.* **2007**, *359*, 159–176.
- Kiernan, U. A. *Expert Rev. Proteomics* **2007**, *4*, 421–428.
- Wuhrer, M. *Expert Rev. Proteomics* **2007**, *4*, 135–136.
- Tomana, M.; Novak, J.; Julian, B. A.; Matousovic, K.; Konecny, K.; Mestecky, J. *J. Clin. Invest.* **1999**, *104*, 73–81.
- Renfrow, M. B.; Cooper, H. J.; Tomana, M.; Kulhavy, R.; Hiki, Y.; Toma, K.; Emmett, M. R.; Mestecky, J.; Marshall, A. G.; Novak, J. *J. Biol. Chem.* **2005**, *280*, 19136–19145.
- Moldoveanu, Z.; Wyatt, R. J.; Lee, J. Y.; Tomana, M.; Julian, B. A.; Mestecky, J.; Huang, W. Q.; Anreddy, S. R.; Hall, S.; Hastings, M. C.; Lau, K. K.; Cook, W. J.; Novak, J. *Kidney Int.* **2007**, *71*, 1148–1154.

- (48) Novak, J.; Tomana, M.; Shah, G. R.; Brown, R.; Mestecky, J. *J. Dent. Res.* **2005**, *84*, 897–901.
- (49) Kuhn, E.; Wu, J.; Karl, J.; Liao, H.; Zolg, W.; Guild, B. *Proteomics* **2004**, *4*, 1175–1186.
- (50) Anderson, N. L.; Anderson, N. G.; Haines, L. R.; Hardie, D. B.; Olafson, R. W.; Pearson, T. W. *J. Proteome Res.* **2004**, *3*, 235–244.
- (51) Hansson, O.; Zetterberg, H.; Buchhave, P.; Londos, E.; Blennow, K.; Minthon, L. *Lancet Neurol.* **2006**, *5*, 228–234.
- (52) Preda, A.; Van Dijk, L. C.; Van Oostaijen, J. A.; Pattynama, P. M. *Eur. Radiol.* **2003**, *13*, 527–530.
- (53) Decramer, S.; Bascands, J. L.; Schanstra, J. P. *World J. Urol.* **2007**, *25*, 457–465.
- (54) Weissinger, E. M.; Schiffer, E.; Hertenstein, B.; Ferrara, J. L.; Holler, E.; Stadler, M.; Kolb, H. J.; Zander, A.; Zurbig, P.; Kellmann, M.; Ganser, A. *Blood* **2007**, *109*, 5511–5519.
- (55) Julian, B. A.; Wittke S.; Novak, J.; Coon, J. J.; Zurbig, P.; Schiffer, E.; Haubitz, M.; Moldoveanu, Z.; Calcaterra, S. M.; Wyatt, R. J.; Sykora, J.; Sladkova, E.; Hes, O.; Mischak, H.; McGuire, B. M. *Electrophoresis* **2007**, in press.

PR070529W