

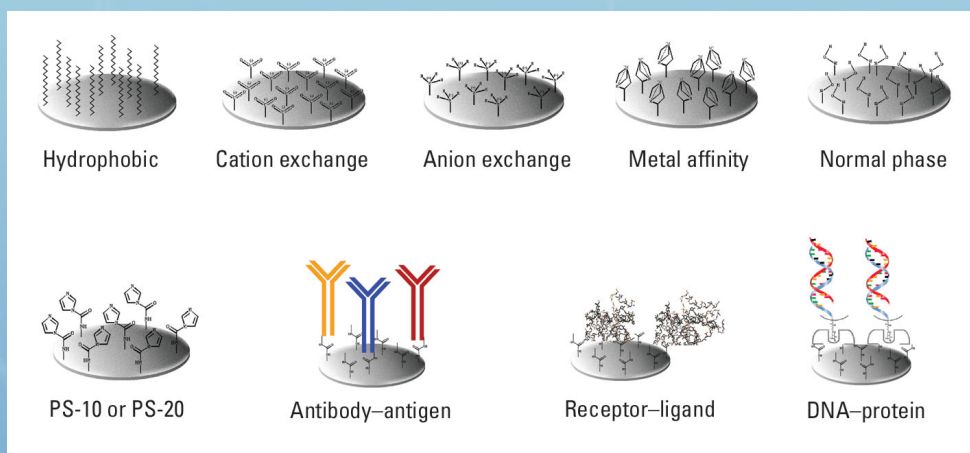
# SELDI-TOF MS *for* Diagnostic Proteomics

By combining chromatographic retention with MS, SELDI-TOF MS can generate protein profiles from as little as 1  $\mu$ L of serum or as few as 25–50 cells.

The urgent need to identify disease biomarkers is underscored by the improved survival rates of patients diagnosed in the early stages of cancer. Discovery, identification, and validation of proteins associated with a particular disease state present a difficult and laborious task, often requiring the analysis of hundreds, if not thousands, of samples. The predominant proteomic method for discovering disease biomarkers is two-dimensional polyacrylamide gel electrophoresis (2-D-PAGE), in which proteins from two distinct samples are analyzed and their protein expression patterns compared. Protein spots of interest are excised from the gel and proteolytically or chemically digested, and the resultant peptides are analyzed by MS to identify the protein.

As a separation technique, 2-D-PAGE provides excellent resolution of complex protein mixtures. However, the method is laborious, has low sensitivity to conventional stains, and cannot resolve proteins with extremes in molecular weight, hydrophobicity, and isoelectric points. Two-dimensional liquid separations such as 2-D-HPLC have

Haleem J. Issaq • Thomas P. Conrads • DaRue A. Prieto  
Radhakrishna Tirumalai • Timothy D. Veenstra  
National Cancer Institute—Frederick



**FIGURE 1.** The arrays consist of either 8 or 16 spots composed of a specific chromatographic surface. The upper arrays represent chemically modified surfaces, and the bottom arrays are biochemically modified surfaces.

not yet been successfully used for disease biomarker detection because of the complexity of proteins in samples, such as serum, plasma, and tissue extracts.

Surface-enhanced laser desorption/ionization time-of-flight MS (SELDI-TOF) is a novel approach introduced by Hutchens and Yip (1). Unlike HPLC/MS, which is based on elution, SELDI-TOF MS combines retention with MS. The principle of this approach is very simple. Proteins are captured by adsorption, partition, electrostatic interaction, or affinity chromatography on a solid-phase protein chip surface. Although SELDI provides a unique sample preparation platform, it is similar to MALDI MS in that a laser ionizes samples that have been co-crystallized with a matrix on a target surface. Unlike MALDI target surfaces, the protein chip chromatographic surfaces in SELDI are uniquely designed to retain proteins from complex mixtures according to their specific properties. After adding a matrix solution, proteins can be ionized with a nitrogen laser and their molecular masses measured by TOF MS.

Proteins from complex biological specimens such as serum, plasma, intestinal fluid, urine, cell lysates, and cellular secretion products have been profiled using SELDI-TOF MS (2–25). SELDI-TOF MS can also be used for more targeted studies, such as characterizing protein–protein (12, 13) and protein–DNA interactions, and is most effective at profiling low-molecular-weight (<20-kDa) proteins and peptides, providing a complementary visualization technique to 2-D-PAGE. SELDI-TOF MS is more sensitive and requires smaller amounts of sample than conventional 2-D-PAGE. Protein profiles can be readily generated from as little as a single microliter of serum and have been generated from as few as 25–50 cells (5, 6).

analysis software. The protein chip arrays are the heart of the SELDI-TOF MS technology and distinguish it from other MS-based systems. The arrays are composed of different chromatographic surfaces that, unlike conventional HPLC or GC, are designed to retain, not elute, proteins of interest. The protein chip arrays are made of a 10-mm-wide  $\times$  80-mm-long aluminum strip with 8 or 16 2-mm spots composed of a chemically (anionic, cationic, hydrophobic, hydrophilic, or metal ion) or biochemically (immobilized antibody, receptor, DNA, enzyme, etc.) active surface (Figure 1). Each surface is designed to retain proteins according to a general or specific physicochemical property of the proteins. Typically, chemically active surfaces retain whole classes of proteins, whereas surfaces to which a biochemical agent, such as an antibody or other type of affinity reagent, is coupled are designed to interact specifically with a single, target protein.

The advantage of biochemically active surfaces is that they can be used to exploit specific molecular recognition mechanisms such as antibody–antigen, enzyme–substrate, receptor–ligand, and protein–DNA interactions to selectively capture target proteins from complex biological samples. The selectivity of chemically active surfaces is illustrated in Figure 2, which shows the protein profiles of low-molecular-weight (<20-kDa) proteins of a human serum filtrate using several surfaces. Note that the figure illustrates that different surfaces retain different proteins, and, when anion and cation exchange surfaces are used, these variations depend on the pH of the sample. Although protein chips containing chemically treated surfaces are commercially available, biochemical surfaces are custom-made by using an open, preactivated platform on which a bait molecule is immobilized. Any crude extract or sample can be applied to the sur-

SELDI-TOF MS has helped researchers discover potential diagnostic markers for prostate (3–7, 23), bladder (8), breast (9), and ovarian (5, 6, 21) cancers; Alzheimer’s disease (10); and interstitial cystitis (23). The technique has also been used to characterize phosphorylated and glycosylated proteins (14, 15), transcription factors (16), and peptides and proteins shed or secreted by various cancer cell lines (17, 20).

### Components

Three major components constitute the instrumentation: protein chip arrays, a mass analyzer, and data

face, thus retaining those target proteins that interact with the bait molecule.

The mass analyzer is a relatively simple TOF mass spectrometer equipped with a pulsed, UV, nitrogen laser (Figure 3). The laser irradiates the sample, which is then desorbed/ionized. The ionized molecules are accelerated in an electric field into a field-free region under vacuum (the so-called TOF tube) toward an ion detector. The  $m/z$  of each species is recorded on the basis of the time required for the ions to pass through the TOF tube. The mass analyzer, though relatively sensitive, provides only low resolution, and hence low mass accuracy.

The output generated from the TOF MS analysis is a trace showing the relative abundance versus the molecular weights of the detected proteins (Figure 4). To simplify the identification of differences in protein abundances between two samples, the software can convert the MS peak trace into a simulated one-dimensional gel electrophoresis display, also known as the “gel view”. In addition, the software can compare these displays to identify unique peaks or those that show a significant abundance difference in one of the samples. The software also provides tools to conduct cluster analyses to identify significant differences in relative protein abundances across many samples.

### Sample prep

One of the true powers of the SELDI-TOF MS method is its ability to analyze very crude and complex samples in an array

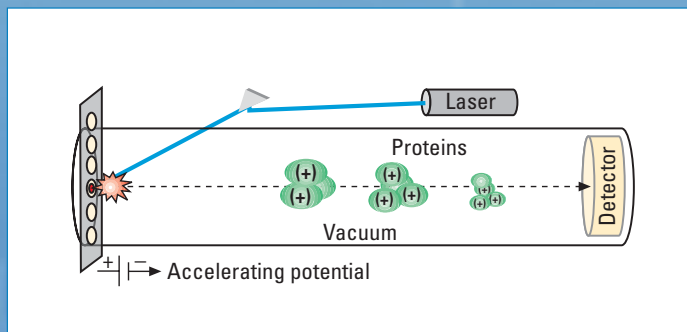


FIGURE 3. Schematic of SELDI-TOF MS.

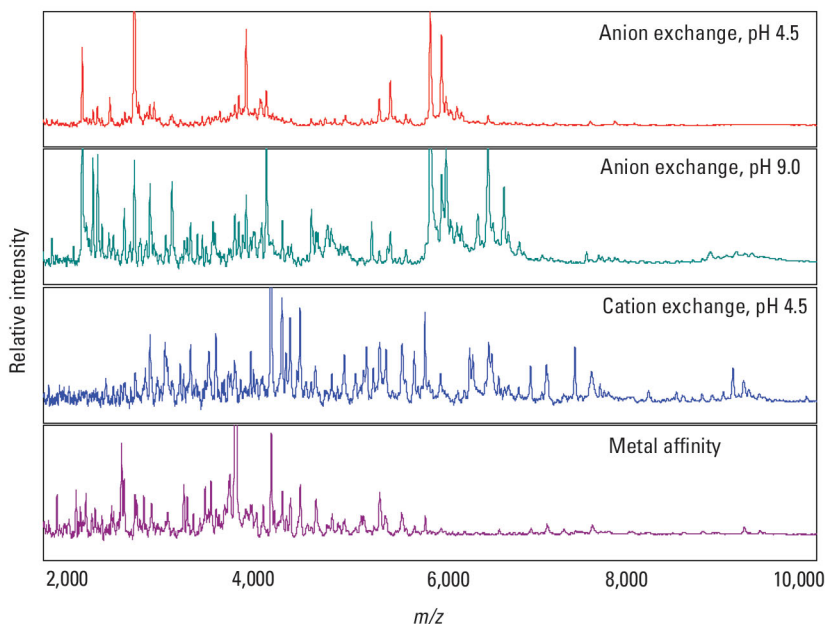


FIGURE 2. Profile of low-molecular-weight proteins in a human serum filtrate on different protein chip array surfaces.

format, which allows for high-throughput measurements. Biofluids such as serum, urine, and plasma can be spotted directly on the protein chip surfaces with little or no sample cleanup or preparation. Samples can be spotted directly onto the protein chip surface or indirectly using a bioprocessor.

In direct spotting, an aliquot as small as 1  $\mu\text{L}$  is dispensed onto the protein chip surface under specific binding conditions that determine which proteins will be retained. For indirect spotting, the sample is applied to the protein chip with the aid of a bioprocessor that is designed to deliver larger sample volumes and

wash solution to individual spots. The bioprocessor captures proteins present at low concentrations from samples. It is available in 8- and 96-well formats, each well having a maximum volume of 500  $\mu\text{L}$ . The 96-well bioprocessor has a top, which contains the wells, gasket, and bracket that holds up to 12 protein chip arrays. After the samples are dispensed into each well, the top is sealed to prevent evaporation. The bioprocessor is then shaken vigorously for 30 min to 2 h, giving the proteins of interest time to adsorb to the chip surface.

After sample application, protein specificity is achieved by applying a series of washes with an appropriate solvent or buffer designed to elute unbound proteins and interfering substances while retaining the proteins of interest. After the chip surfaces dry, a matrix solution such as sinapinic

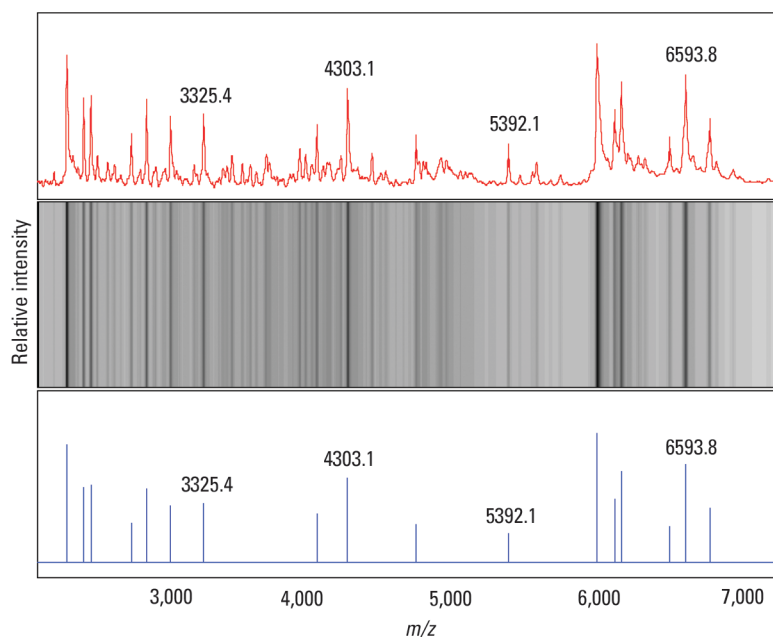


FIGURE 4. Output from TOF MS displayed as (top to bottom) trace, gel, and map views.

acid is added, and the array is inserted into the mass analyzer to measure the  $m/z$  values of the bound species.

### Protein ID

A low-resolution profile of species bound to the protein chip surface is essentially what is observed in a SELDI-TOF MS spectrum. The mass analyzer's resolution, mass accuracy, and lack of tandem MS capabilities make direct protein identification tenuous at best, unless a protein is selectively targeted using an affinity-based surface. So, what is the value of the results?

The value lies in the ability to obtain and compare spectra from a significant number of samples in a relatively short time with very little sample preparation or sophisticated chromatography. For example, a single operator can acquire mass spectra of >150 different samples in a single day. The analysis of a large number of samples will ideally reveal a protein signal that is unique to, or overexpressed in, one sample set when compared with a different sample set. The net result is the molecular mass of a protein(s) that is differentially abundant in a particular sample set.

The next step is the actual identification of the differentially expressed protein(s). Identification most often requires developing a strategy to purify the protein(s) using classical chromatography followed by identification using conventional high-resolution MS. Although SELDI-TOF MS is limited in its ability to identify proteins (indeed, this is not its designed purpose), the process involved in recognizing differentially abundant protein(s) provides the foundation for isolating the protein(s) of interest. The purification strategy can be based on any

type of chromatography; however, the type of protein chip that the protein(s) binds to provides a useful starting point when designing the purification scheme. The purification can be followed by matching the molecular weight(s) of the differentially expressed protein(s) recorded by SELDI-TOF MS with those measured during the purification process. Once the protein is purified, standard high-resolution MS methods such as peptide mapping or tandem MS can be used to identify it.

A simpler approach would be to identify the differentially expressed proteins that are captured directly on the chip, which would be much easier, faster, and less costly. New ion sources are available that allow the protein chips to be analyzed using, for example, a hybrid triple quadrupole TOF mass spectrometer. The aim is to use this high-resolution tandem MS instrument to directly identify peptides desorbed and ionized off the protein chip surface. Although, in theory, this seems to be a reasonable approach, in practice, this method is still unproven with clinical samples.

### A diagnostic tool

An application that has generated much interest in SELDI-TOF MS is its promise as a diagnostic tool in the early detection of diseases, such as ovarian cancer. Serum samples taken from healthy and diseased subjects are spotted onto protein chips and analyzed by TOF MS to generate a protein profile. By surveying different population cohorts (i.e., diseased vs unaffected), unique protein abundance profiles, similar to those in Figure 2, can be produced. A specific pattern can be distinguished using statistical programs (19) and the results used to delineate the stage of the disease and/or the utility of the treatment. Pattern recognition studies do not depend on identifying specific proteins, but rather on fluctuations in protein abundance patterns.

Petricoin and Liotta have successfully used SELDI-TOF MS to generate proteomic profiles of serum samples obtained from patients with breast, prostate, and ovarian cancer (5, 6, 21). In one study, the protein profiles of serum samples from several patients diagnosed at various stages of ovarian cancer were compared with those from unaffected individuals (21). Although no obvious unique distinguishing protein signal was immediately evident, a genetic-based, artificial intelligence algorithm was able to decipher diagnostic patterns within the profiles. This same software program correctly identified serum samples as being taken from patients with ovarian cancer or from healthy

## Spectra from a significant number of samples can be obtained and compared in a relatively short time.

individuals. The provocative finding was that the SELDI-TOF MS analysis was able to correctly distinguish patients with stage I ovarian cancer, which has a 95% survival rate with proper intervention.

Wright et al. have generated data using SELDI-TOF MS to characterize serum taken from individuals with prostate cancer with a specificity and selectivity that is greater than the currently accepted protein-specific antigen test (7, 18, 22). For example, they used a protein chip immunoassay to detect and quantify prostate-specific membrane antigen (PSMA) and used the results to discriminate between benign and malignant prostate disease (Figure 5; 18). PSMA was captured from serum by anti-PSMA antibodies bound to protein chip arrays. The captured PSMA was detected by SELDI MS and quantified by comparing the mass signal integrals to a standard curve established using purified recombinant PSMA. The average serum PSMA value for prostate cancer (623.1 ng/mL) was significantly different from that for benign prostate hyperplasia (BPH) (117.1 ng/mL) and the normal groups (under age 50, 272.9 ng/mL; over age 50, 359.4 ng/mL). The ability to diagnose cancer, especially at early stages, where the prognosis is generally more positive, would have profound effects in the clinical setting.

In our laboratory, SELDI-TOF MS has been used to profile the differences and similarities in proteins secreted from six ovarian cancer human cell lines (20). It is also being evaluated as a diagnostic tool for interstitial cystitis, a bladder infection that predominantly affects women. Preliminary results from our laboratory indicate that this technique, combined with proteomic pattern recognition software, is a potential diagnostic test for this ailment.

### Biological applications

The surfaces of protein chips can be covalently modified with bait molecules, such as antibodies or recombinant proteins, to capture proteins that have an affinity to the bait. If the affinity reagent being used as bait recognizes an epitope in different isoforms of the protein, then these proteins can be separated by their molecular mass, resulting in multiple discrete signals. This direct sampling has a unique advantage over the traditional enzyme-linked immunosorbent assay, in which the result-

ant signal is a weighted average of all of the bound species. Detection limits depend on the affinity reagent used, but they have been demonstrated in the high-attomole range, with standard quantitation curves linear over 2–3 orders of magnitude.

Post-translationally modified proteins can also be detected using SELDI-TOF MS. Glycosylated proteins generally give a broad peak when analyzed by SELDI-TOF MS, resulting from the numerous isoforms of the protein that are present. They can be deglycosylated while still attached to the protein chip and then reanalyzed (15). The presence of a glycosylated protein is confirmed by the collapse of the broad peak into a single, sharp peak.

Another important application is studying phosphorylation states of proteins involved in signal transduction pathways. The 80-Da shift in mass resulting from phosphorylation is readily detectable even with the limited resolution and mass accuracy of

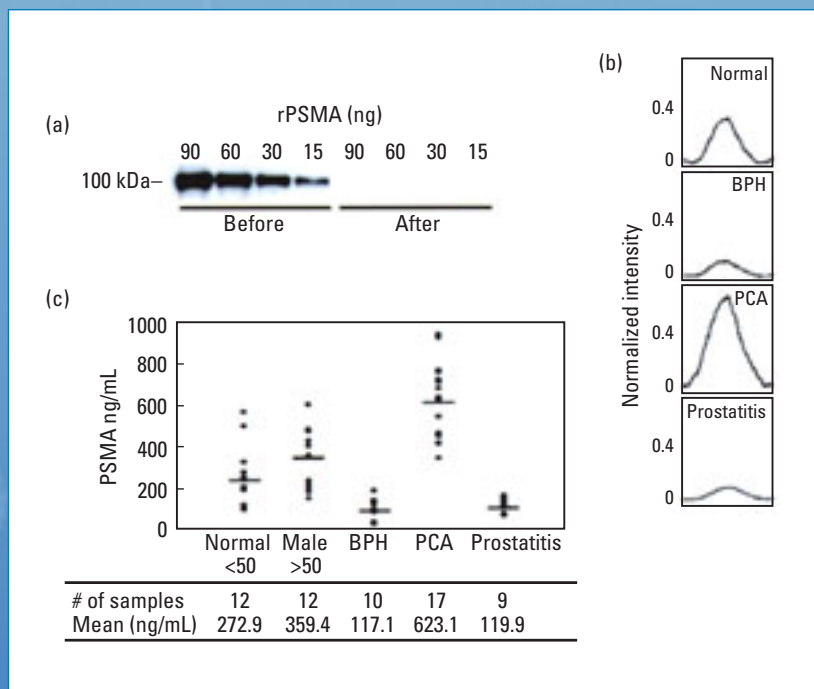


FIGURE 5. Measurement of PSMA levels by SELDI protein chip immunoassay.

(a) Western blot analysis showing the amount of rPSMA added to the chip array before injection and the amount remaining in supernatant after injection. Analysis shows that up to 90 ng of rPSMA could be bound to the array. (b) Representative example of the mass spectrum obtained from a PSMA SELDI immunoassay in the spectra of normal, BPH, prostate cancer, and prostatitis serum samples. (c) Comparison of serum PSMA levels in normal male donors and patients with BPH, PCA, and prostatitis. (Adapted with permission from Ref. 18.)

SELDI-TOF MS. Kinase activity assays can be performed directly on the protein chip surface by adding the proper substrate peptide and ATP to the reaction buffer. After an incubation period, the phosphorylated peptide is detected by an 80-Da increase in the peptide's molecular weight (14). The major advantages of this technique are that it does not require radioactive isotopes, a single assay can be sampled at many different time points, and a discrete signal can be confidently assigned to the phosphopeptide.

Several investigators are using SELDI-TOF MS as an alternative to gel shift assays (16). Streptavidin can be covalently coupled to the protein chip surface, and biotinylated DNA or RNA can be bound to make a nucleic acid affinity surface. Nuclear extracts can then be examined for possible associations with specific consensus sequences of the DNA or RNA. Proteins bound to the nucleic acid are then detected by TOF MS. Furthermore, the protein(s) associating with the nucleic acid affinity surface can be digested on the chip, and the captured peptides can be identified by peptide mapping.

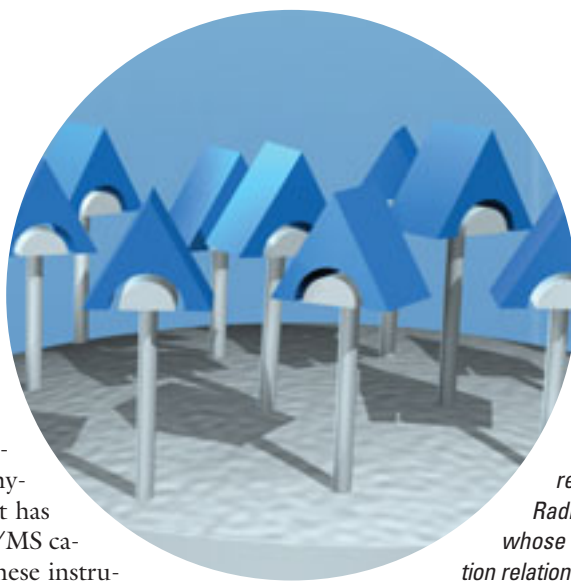
**Proteomic diagnostics may only be effective using methods that group these variances and compare them to indicate a disease state.**

### **Advantages and limitations**

When used with a robotics system, SELDI-TOF MS is a high-throughput technique that allows hundreds of samples to be screened for disease biomarker identification in a relatively short time. Otherwise, the system is manual, time consuming, and prone to human error. We have been fortunate in our ability to obtain good day-to-day, operator-to-operator, and sample-to-sample reproducibility. At first glance, the cost of protein chip arrays seems to be high at \$75/chip (8 spots), but it is not exorbitant when the savings in time and labor and the wealth of information generated is considered.

As with all analytical methods, SELDI-TOF MS has its limitations. Although it can rapidly generate proteome patterns from complex mixtures, with the goal of recognizing unique biomarkers related to a particular disease state, routine identification of these biomarkers cannot be achieved. The low resolution, and hence mass accuracy, coupled with the inability to do TOF MS/MS, prevents reliable identification based on con-

ventional bioinformatic searching. Identification still relies on classical methods to enrich and/or isolate the species recognized as potential biomarkers in the initial proteomic pattern screening process. Although a SELDI interface is available for high-resolution instruments, such as the hybrid quadrupole TOF instrument that has a much higher mass accuracy and MS/MS capabilities, the routine application of these instrument platforms for biomarker identification directly from the applied sample has yet to be demonstrated. Chiefly successful at discovering proteins in the low-molecular-weight range, SELDI-TOF MS has not yet shown itself consistently successful in studying high-molecular-weight proteins.



ration techniques for proteomic research. Thomas P. Conrads is the associate director of the MS center. His current interests include novel applications of high-resolution MS for the early detection of cancer. DaRue A. Prieto is a senior research associate whose primary research interest is diagnostic proteomics.

Radhakrishna Tirumalai is a senior scientist whose research interests include structure–function relationships in proteins and diagnostic and therapeutic proteomics. Timothy D. Veenstra is the director of the MS center, and his research interest is using MS to further understand disease states. Address correspondence to Issaq at Analytical Chemistry Laboratory, SAIC–Frederick, National Cancer Institute–Frederick, P.O. Box B, Frederick, MD 21702 (issaqh@mail.ncifcrf.gov).

### Patterns, not specific proteins

The SELDI-TOF MS approach goes against many of the current trends in proteomics today, including development of instrumentation to identify hundreds or thousands of proteins in a single analysis. SELDI-TOF MS provides a simple, low-resolution pattern generated from proteins retained on a specific chromatographic surface. In most instances, it does not allow the direct identification of proteins that may be potential disease biomarkers.

Why then is the technology generating such interest? SELDI-TOF MS can rapidly screen and generate proteomic patterns for hundreds of crude samples. This simple answer may have a profound impact on how proteomics is viewed in the future. The pharmaceutical industry is expending much effort on identifying disease biomarkers in biofluids. Results from conventional 2-D-PAGE MS as well as SELDI-TOF MS suggest that no “smoking gun” biomarker exists for many disease states. Serum from healthy and diseased subjects exhibits a wide range of variability. Proteomic diagnostics may only be effective using methods that group these variances and compare them to indicate a particular disease state. Indeed, the serum samples from the ovarian cancer patients studied by Petricoin and Liotta did not generate a single unique pattern, nor did all of the control samples. The multiple patterns generated by the ovarian cancer samples, however, did not overlap with those generated by the normal samples, suggesting that patterns may be a more realizable goal than identifying specific proteins unique to a particular disease state.

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract no. N01-CO-12400.

Haleem J. Issaq is principal scientist and head of the separation technology group in the analytical chemistry laboratory at NCI–Frederick. His research interests are developing multidimensional–multimodal separation

### References

- Hutchens, T. W.; Yip, T. T. *Rapid. Commun. Mass Spectrom.* **1993**, *7*, 576–580.
- Merchant, M.; Weinberger, S. *Electrophoresis* **2000**, *21*, 1164–1167.
- Eggeling, V. F.; et al. *Electrophoresis* **2001**, *22*, 2898–2902.
- Srinivas, P. R.; Srivastava, S.; Hanash, S.; Wright, G. L. *Clin. Chem.* **2001**, *47*, 1901–1911.
- Paweletz, C. P.; Liotta, L. A.; Petricoin, E. F. *Urology* **2001**, *57*, 160–163.
- Paweletz, C. P.; et al. *Drug Dev. Res.* **2000**, *49*, 34–42.
- Wright, G. L.; et al. *Prostate Cancer and Prostatic Diseases* **2000**, *2*, 264–276.
- Vlahou, A.; et al. *Am. J. Pathol.* **2001**, *158*, 1491–1501.
- Wulfkuhle, J. D.; et al. *Proteomics* **2001**, *1*, 1205–1215.
- Austen, B. M.; Frears, E. R.; Davies, H. J. *Pept. Sci.* **2000**, *6*, 459–469.
- Hampel, D. J.; Sansome, C.; Sha, M.; Brodsky, S.; Lawson, W. E.; Goligorsky, M. S. *J. Am. Soc. Nephrol.* **2001**, *12*, 1026–1035.
- Stoica, G. E.; et al. *J. Biol. Chem.* **2001**, *276*, 16772–16779.
- Hinshelwood, J.; Spencer, D. I. R.; Edwards, Y. J. K.; Perkins, S. J. *J. Mol. Biol.* **1999**, *294*, 587–599.
- Cardone, M. H.; et al. *Science* **1998**, *282*, 1318–1321.
- Chernyak, A.; Karavanov, A.; Ogawa, Y.; Kovac, P. *Carbohydr. Res.* **2001**, *330*, 479–486.
- Forde, C. E.; et al. *Biochem. Biophys. Res. Commun.* **2002**, *290*, 1328–1335.
- Sato, K.; Sasaki, K.; Akiyama, Y.; Yamaguchi, K. *Cancer Lett.* **2001**, *170*, 153–159.
- Xiao, Z.; et al. *Cancer Res.* **2001**, *61*, 6029–6033.
- Fung, E. T.; Enderwick, C. *BioTechniques*, in press.
- Prieto, D. A.; Conrads, T. P.; Scudiero, D. A.; Veenstra, T. D.; Issaq, H. J. *J. Proteome Res.*, submitted.
- Petricoin, E. F.; et al. *Lancet* **2002**, *359*, 572–577.
- Adam, B.-L.; et al. *Cancer Res.* **2002**, *62*, 3609–3614.
- Prieto, D. A.; et al. Presented at the 13th Annual Frederick CE Conference, Frederick, MD, October 21–22, 2002.
- Li, J.; Zhang, Z.; Rosenzweig, J.; Wang, Y. Y.; Chan, D. W. *Clin. Chem.* **2002**, *48*, 1296–1304.
- Issaq, H. J.; Veenstra, T. D.; Conrads, T. P.; Felschow, D. *Biochem. Biophys. Res. Commun.* **2002**, *292*, 587–592.