



FOCUS: MALDI-TOF AND BIOLOGICAL MS: CRITICAL INSIGHT

The Future of Biological Mass Spectrometry

Marvin L. Vestal

CEO Virgin Instruments Corp, 60 Union Ave. Suite 1R, Sudbury, MA 01776, USA

Abstract

Biological applications of mass spectrometry have grown exponentially since the discovery of MALDI and electrospray ionization techniques. This growth has been further fueled by the massive volume of DNA sequence information that is now available. An ambitious goal of some of this research is to monitor the level and modification of all proteins and metabolites in a biological sample such as plasma. A major research effort in mass spectrometry and related disciplines has been expended over the past several years toward reaching this and other less ambitious goals, and considerable progress has been made; but the presently available tools are clearly not sufficient for these very difficult tasks. In this “critical insight” discussion we suggest that recent advances in time-of-flight (TOF) technology with MALDI ionization may provide some important new tools for achieving the goals of biological research.

Key words: MALDI, TOF, Biological mass spectrometry

Introduction

This paper presents a personal view of the future of mass spectrometry in biology. Whether it includes any critical insight is left for the reader to decide. Mass spectrometry has now entered its second century. I am privileged to have been actively involved in this discipline for the latter half of this first century, and amazed that I may still be able to make some contribution in the second. The discussion includes a very truncated history, and a summary of the present status as necessary background for discussing the future.

The Past

The first half-century was dominated by physicists and applications to basic physical research, but chemical and biological applications began to emerge prior to WWII and now many chemists, biochemists, and biologists perceive mass spectrometry as an essential tool for their research. The earlier work was limited to atoms and small molecules since techniques were not available for generating ions from relatively nonvolatile samples without (often uncontrolled) modification of the chemical structure. The nearly simultaneous discoveries of electrospray (ESI) by Fenn and coworkers [1] and MALDI by

Karas and Hillenkamp [2] twenty-two years ago removed the volatility barrier for mass spectrometry. ESI MS developed very rapidly, at least partly due to ease of interfacing with commercially available quadrupole and trap instruments widely employed for analytical applications. Applications of MALDI have developed more slowly, but the potential of MALDI has stimulated development of improved TOF instrumentation particularly tailored for this ionization technique.

The initial excitement following introduction of these revolutionary ionization techniques focused on production of intact molecular ions from very large molecules including proteins and DNA oligomers. Direct measurement of the mass of these large molecules is often useful for certain applications, but it is now widely recognized that molecular weight alone, even with high accuracy, is not generally sufficient for identification.

Prior to the revolution in ionization techniques for large, nonvolatile molecules, biological mass spectrometry was dominated by large, expensive and very powerful magnetic deflection instruments. One of the earlier techniques for ionizing nonvolatile molecules, known as “Fast Atom Bombardment” [3], was compatible with these analyzers, but the new techniques, MALDI and Electrospray, were difficult to incorporate. Quadrupole mass filters and time-of-flight analyzers were available in many laboratories, and the major applications involved combinations of gas chromatography (GC) with these analyzers. Extension to non-

Correspondence to: Marvin L. Vestal; e-mail: marvin.vestal@virgininstruments.com

volatile molecules required interfacing with liquid chromatography (LC) and the quadrupoles were readily adaptable to LC interfacing with electrospray ionization. Also, the multiple charging inherent in electrospray ameliorated the limited m/z range of the quadrupoles by producing ions even from very large molecules within the range accessible with existing instruments. On the other hand, time-of-flight analyzers provide essentially unlimited mass range, but the available instruments were not initially well suited for either electrospray or MALDI, and interfacing with LC had not been developed.

In 1953, Francis Crick declared to patrons of a Cambridge pub: "We have discovered the secret of life".[4] The discovery of the DNA double helix certainly deserves all of the acclaim that it has received, but most biologists would now agree that the initial exuberance was somewhat exaggerated. In the seventies many geneticists were confident that deciphering full genome sequences would explain all diseases. Some may still cling to the narrow view of "genetic determinism" even though by the nineties it was clear to many that knowing genome sequences is essential, but not sufficient for understanding complex biological events and diseases. [5] Lost in the hype accompanying completion of the human genome is the fact that very few new drugs or therapies have yet been developed using data derived from the human genome project. Sydney Brenner may have summarized the situation best, "When all of this [genome project] mania dies down, we'll get back to hypothesis-driven normal science. What people are forgetting is that there are several tens of thousands of biochemists who can now get stuck in and study protein function. What you have done is provided them with a tool, and that's fine." [6]

Following the announced completion of the human genome it was widely believed that "proteomics was the next big thing." Unfortunately, the proteomics factories spawned from this euphoria failed to deliver on their promises to big Pharma, and these factories have gone out of business. However, it is still generally recognized that metabolomics, functional proteomics and genomics will eventually provide new technologies that revolutionize diagnosis and treatment of disease [7, 8].

A review of functional proteomics by Godovac-Zimmerman and Brown [9] provides an elegant summary of the requirements for truly large-scale global analyses.

"One of the most intellectually challenging, scientifically productive, and socially applicable endeavors of modern biological science is the current effort to understand the function of living cells at the molecular level. New methods are required to monitor and analyze the spatial and temporal properties of the molecular networks and fluxes involved in living cells and to identify the molecular species that participate in such networks via functional stimulation, perturbation, or isolation of these networks. It appears that an ideal starting point for understanding the functioning of a biological system would be a complete quantitative inven-

tory of all of its molecular components as functions of both time and position."

It appears that a new set of tools are needed to allow large numbers of biochemists to efficiently conduct large-scale biological experiments. A key component of these tools is the genomic database; but equally important are new methods for separating, identifying, quantifying, and characterizing proteins and metabolites. Even before the human genome project was completed Eric Lander published a seminal paper in *Science*, entitled "The New Genomics: Global Views of Biology" [10]. One of his recommendations involved monitoring the level and modification of all proteins. He suggested that "A future version of the 2D gel might involve an automated system to take total cellular protein, partially separate it chromatographically, proteolytically cleave each fraction, analyze it by mass spectrometry, and recognize the resulting peptides by comparing their signatures to the complete protein database." A major research effort in mass spectrometry and related disciplines has been expended over the past several years toward reaching this goal, and considerable progress has been made. But the presently available methods are still too slow and cumbersome for many applications. Presently available tools are clearly not sufficient for this very difficult task. [11]

The Present

Advances in chemistry, separations technology, and mass spectrometry during the last 20 years have begun to make global identification and quantification of biological molecules possible, but not yet practical. Mass spectrometry employing either electrospray [1] or MALDI [2] is used to determine molecular weights of intact proteins, and following digestion with enzymes such as trypsin, tandem MS-MS is the method of choice for identifying and characterizing proteins. No single MS-MS instrument or technique has established dominance, although electrospray is presently more widely used than MALDI. Most current MS-MS applications employ triple quadrupoles, hybrid quadrupole-TOF systems, or ion traps, either quadrupole, orbitrap, or magnetic (as in FTICR).

Recently, combinations of linear ion traps with FTICR and orbitraps have significantly improved the performance of electrospray MS and MS-MS for analyses of complex biological mixtures. The high resolving power and excellent mass accuracy of the FTICR and orbitrap for determining peptide molecular weights together with the speed and sensitivity of the linear ion trap for acquiring low resolution fragment spectra make a powerful combination. Nevertheless, these still fall far short of the requirements for an analyzer suitable for use in large scale biological studies in several respects.

Sensitivity and dynamic range are inadequate.

Mass analyzers are too slow to keep pace with fractionations producing large numbers of fractions.

Sample utilization is poor.

Automated data interpretation is unreliable.

Instruments are too complex and expensive for many laboratories.

Highly trained operators are required.

Further improvements in protein chemistry, separations science, mass spectrometry, and bio-informatics are required to complete this revolution. Despite a massive effort, progress has been agonizingly slow and we believe that major bottlenecks involve the combinations of separations and mass spectrometry that are presently available.

The Future

Completion of the human genome project and improved methods for analyzing proteins and small molecules in complex biological samples have been widely predicted to dramatically change drug discovery and clinical practice of medicine in the near future. It is now sixteen years after the term "Proteomics" was coined and ten years after the first announcement of completion of the human genome, and this dramatic improvement has not yet been realized. The Human Proteome Organization has launched major initiatives focused on plasma, liver, and brain proteomics [12]. The results of these initiatives demonstrate that current technology is not yet sufficient for the stated task: advances in high performance ion trap mass spectrometers employing electrospray ionization and nanoflow liquid chromatography have significantly improved our ability to analyze biological samples, but this technology has serious inherent limitations on dynamic range and speed.

The fundamental obstacle to further rapid progress can be simply stated. The range of protein concentrations in biological tissues and fluids is very large (up to 10^{12} in serum) and the dynamic range of the analytical approaches is small (less than 10^3 in most cases). The major limitations are dynamic range, throughput, and reliability of the result. The basic sensitivity of current mass spectrometers is high when separation is carried out with nanoflow systems, but the dynamic range is seriously limited by the inherent low capacity. Many of the protein biomarkers are present at very low concentrations (<1 ng/mL). The obvious solution is to start with enough sample (ca. >1 g protein) and extensively fractionate the sample so that the probability is very small that a low-level sample is found in the same fraction as one at >1000 times higher concentration. But this requires high capacity fractionation and separation systems that generate a very large number of fractions from each sample (ca. >10,000).

MALDI vs. Electrospray

Electrospray has become dominant in recent years with high performance LC-MS and LC-MS-MS using triple quadrupoles, hybrid quadrupole-TOF systems, and ion traps, either quadrupole or magnetic (as in FTICR).

Recent developments in ion trap mass spectrometry, in particular the "Orbitrap", provide resolving power and mass accuracy that competes with FTICR. The conventional wisdom as expressed by commercial development of instrumentation and many of the leaders in biological mass spectrometry is that the future is electrospray, and MALDI is relatively unimportant. We believe, on the contrary, that there are presently no known prospects for further dramatic improvement in electrospray MS technology while systems based on MALDI-TOF can provide powerful new tools for addressing the very difficult problems posed by large-scale applications of biological mass spectrometry. The fundamental problem with electrospray is that it requires on-line direct coupling of the effluent from the chromatograph to the inlet of the mass spectrometer, and high sensitivity "nanospray" systems require very small diameter capillary columns with flow rates in the nanoliter/min range. Thus high sensitivity can be achieved but separation times are long and capacity severely limited.

The main disadvantage of direct coupling between the separation and the mass spectrometer is that all of the measurements on an eluting peak must be made during the time that the peak is present in the effluent. Depending on the speed of the separation technique, this time may be as much as a minute or less than a second. Protein digests derived from complex biological extracts may contain many thousands of peptides in a single sample. Even after LC separation, hundreds of peptides may co-elute. Typically, measurements on these digests involve measurement of the peptide masses in MS mode, deciding which peaks should be measured using MS-MS, and measuring all of the MS-MS spectra of interest. In many cases the separation must be slowed down to accommodate the speed of the mass spectrometer, or some of the potential information about the sample is lost.

MALDI TOF analysis of complex mixtures by LC/MS and MS/MS is currently used in a number of laboratories, and TOF-TOF and Q-TOF instruments are commercially available. To interface MALDI with liquid separation techniques such as HPLC or CE, droplets from the liquid effluent, usually with added matrix solution, are deposited sequentially on a suitable surface and allowed to dry. [13] The surface containing the dried samples is then inserted into the vacuum system of the MALDI mass spectrometer and irradiated by the laser beam. Some systems have been described [14] for directly coupling MALDI with separations, but off-line coupling allows the sample deposition to occur at a speed appropriate to the chromatography, and the mass spectrometer can be operated faster or slower as needed to maximize the information. For example, an entire LC run can be rapidly scanned in MS mode to determine the peptide masses and relative intensities for all peptides in the run. This information can then be used in a true data dependent manner to set up the MS-MS measurement for all of the spots on the plate to obtain the required information

most efficiently. Since it rare for all of the sample to be used in most MALDI measurements, additional measurements can be made on the same plate at any later time as needed.

Early applications of MALDI were limited by relatively low resolving power and mass accuracy of available TOF instruments, and the lack of a commonly available interface with liquid chromatography. Recent work has focused on overcoming these limitations and making high performance MALDI-TOF and TOF-TOF instruments interfaced with high capacity separations available. These developments are being pursued in ongoing research projects, and prototype instruments now provide speed, resolution, dynamic range, and sensitivity substantially superior to the performance available with commercially available instruments.

Present Status and Future Prospects for TOF and TOF-TOF

The overall efficiency of a mass spectrometer may be conveniently separated into three components: (1) efficiency of the ionization process (ions/molecule); (2) transmission of the ion optical system (ions out/ions in); (3) detection efficiency (detected pulses out/ions in). In scanning instruments there is an additional efficiency term since only the measured peak is transmitted and all others are discarded. This term is unity for TOF, but not for conventional TOF-TOF since selection of a particular precursor requires rejection of all others in the peptide mass fingerprint spectrum. TOF does not require critical apertures to achieve high resolution and gridless ion optical systems with discrete dynode electron multipliers provide nearly 100% transmission and detection over the range of initial conditions employed in MALDI ionization. In favorable cases for peptides using 4-hydroxy- α -cyanocinnamic acid matrix, the ionization efficiency appears to be very high, but this has not been accurately determined.

Unique properties of TOF-MS

The present generation of MALDI TOF instruments routinely operate at laser rates up to 5 kHz producing full mass spectra over any selected mass range following each laser shot. [15] This speed is orders of magnitude greater than any other mass spectrometer, but there are some practical limitations to the useful speed. It is very rare that a sufficient number of ions are produced in an individual laser shot to provide useful measurements of either peak centroids or intensities, and acquisition of such spectra over any significant time period quickly overwhelms the capabilities of even the most powerful computers. A typical high resolution MALDI TOF spectrum is approximately 1 Mbyte, hence a typical data rate is ca. 5 Terabytes/sec. Thus it is necessary to average a number of spectra using either a time-digital convertor (TDC) or transient digitizer with an on-board averaging. The principal advantage of higher laser rates is the flexibility that it provides to balance higher

sample throughput against increased sensitivity, wider dynamic range, and better sample utilization. In some applications, such as tissue imaging an average of 50 spectra per pixel often produces high quality results at a rate of 100 spectra/s. Producing spectra at these high rates may not be a problem, but storing and interpreting data at these rates is somewhat more demanding. Under these conditions the typical data rate is ca. 100 Mbyte/sec. One day of continuous operation generates 8.64 Terabytes. Storing and managing this data stream is clearly not practical. On the other hand, all of the useful information in the spectrum is contained in the peak centroids, peak intensities, peak widths, and some measure of background noise. A typical averaged spectrum may contain on the order of 100 useful peaks. If the important properties of each peak can be accurately expressed in 16 bytes (4 32-bit words), then the data rate is reduced by a factor of 625, and storage of only 13.8 gigabytes per day is required. This is still substantial, but manageable with available hardware and software.

We have recently developed novel algorithms that allow very rapid and accurate peak detection and mass calibration. Spectra may be saved either as "raw" peak tables or complete spectra. The peak tables contain all the pertinent information about each peak including integrated intensity, centroid, and standard deviation. The new algorithms operate with the necessary speed without sacrificing accuracy as shown by preliminary results. These peak tables constitute the raw data from the measurement and are written to disk. The raw time-of-flight data are discarded, but the peak tables contain sufficient data to allow a statistically equivalent raw TOF spectrum to be regenerated. These high data rates allow very rapid scanning and acquisition of tissue sections.[16] Recent work at Vanderbilt has produced complete images of rat brain sections at a rate of 1 mm²/sec with 100 μ m resolution in a total measurement time of 10 minutes. [17]

In other applications speed may be less important than sensitivity, dynamic range, mass accuracy, and sample utilization. One such application is LC-MS and MS-MS where the speed of the analysis is determined by the chromatograph and in some cases 100,000 laser shots may be used on some fractions. In these applications sensitivity and dynamic range are limited only by the chemical noise.

The performance of the latest generation of MALDI-TOF instruments[18–21] may be summarized as follows: Resolving power > 50,000 for peptides; Mass error < 1 ppm RMS over entire sample plate; Detection limit \sim 1 attomole/ μ L; Dynamic range \sim 10⁵; Performance independent of laser rate (to 5 kHz). A new TOF-TOF instrument provides high-resolution precursor selection with MALDI MS-MS. Single isotopes can be selected and fragmented up to m/z 4000 with no detectable loss in ion transmission and less than 1% contribution from adjacent masses. This instrument also allows 10–50 fold multiplexing in MS-MS. Selected masses must differ by at least 1%, and are preferably within an order of magnitude range in intensity. This allows the generation

of very high quality MS-MS spectra at unprecedented speed. All of the peptides present in a complex peptide mass fingerprint containing a hundred or more peaks can be fragmented and identified without exhausting the sample. Thus speed and sensitivity of the MS-MS measurements can keep pace with the MS results. The combination of high-resolution precursor selection with high laser rate and multiplexing allows high-quality, interpretable MS-MS spectra to be generated on detected peptides at the 10 attomole/ μL level. These specifications represent the current state of research instruments in our laboratory, and it should be noted that not all of these performance limits can be achieved simultaneously. Resolving power is independent of acquisition rate, but mass accuracy, dynamic range, and sensitivity depend on ion statistics that improve in proportion to the square root of the number of laser shots averaged.

MALDI LC-MS and MS-MS

A typical MALDI LC-MS analysis may involve 100–200 fractions deposited on the sample plate, and with the laser operating at 5 kHz MS spectra generated by averaging ca. 1000 laser shots per fraction require a maximum of 40 seconds per separation. Our current sample plate accommodates 2500 fractions; thus either a series of very fast separations or parallel separations with longer retention time can be accommodated. Recording high quality MS spectra from all 2500 fractions takes less than 10 minutes. In general, acquisition of MS-MS spectra requires somewhat longer acquisition times, but with 10-fold multiplexing and 5000 laser shots/spectrum, MS-MS spectra for all 2500 fractions can be recorded in 45 minutes including plate loading time. The estimated analysis time may be modified as required by the results of the MS measurement, since some fractions may contain no interesting peaks and others may include a large number of relatively weak peaks requiring more laser shots. In our experience about 100,000 laser shots are required to desorb most of the sample in a fraction; thus the maximum time per fraction is 20 seconds. Approximately 60,000 fractions per day can be analyzed with this approach. The bottlenecks are then moved to sample preparation and separation since this corresponds roughly to 600 samples/day for injection into the LC, and to data interpretation, archiving, and informatics required to efficiently deal with ca. 600,000 MS-MS spectra/day. While these instruments and interfaces may provide the performance required for global analysis of biological samples, development of improved database and bioinformatics tools are essential for converting the massive volume of data that can be generated by these systems into useful information for addressing biological problems.

Summary

We believe that scientists can be sorted into three groups: users, innovators, and inventors, although some may span

more than one group. We define users as those primarily interested in research on a particular problem or discipline doing the best work they can with available tools and protocols. Innovators discover new ways to use existing tools to address previously intractable problems, and inventors develop new tools. The users constitute by far the largest group and are directly responsible for most of the measurable progress. The innovators generally receive the most attention and prestigious awards since they serve the vital function of adapting and demonstrating new tools for specific applications. The inventors are by far the smallest group, and they are largely ignored by users, innovators, manufacturers, and funding agencies until it is demonstrated by the innovators that the tools they invent are necessary for further advance of the science. The history of DNA sequencing leading to the genomic revolution offers some insight into the processes and interactions that lead to major advances, and this history may serve as a useful guide for similar revolutions in proteomics and metabolomics.

In 1977, twenty-four years after the importance of DNA sequence was recognized, two general methods for DNA sequence analysis were developed independently by Sanger and Coulson [22] and Maxam and Gilbert [23]. These were adopted by many laboratories, and ten years later more than 15,000,000 bases of DNA had been sequenced. Using these techniques a skilled biologist could produce about 50,000 bases of finished DNA sequence per year under ideal conditions. In 1987 an automated approach to DNA sequencing was described [24], and the prospects for automated DNA sequencing and analysis of the human genome was reviewed by Hood, Hunkapiller, and Smith [25]. In that review they envisioned a process involving three overlapping phases: phase I, technology development and mapping, 5–10 years; phase II, sequencing human and other genomes, 10–20+ years; and phase III, understanding human and other genomes, 100's of years. In retrospect, their time estimate for phase I was quite accurate. Phase II was accomplished in a much shorter time than anyone could imagine in 1987, and the time required for phase III remains to be determined.

Why did this take so long? In 1953, when Crick spoke of having found the secret of life, much of the technology eventually employed in the successful DNA sequencers either did not exist or was so new that it was understood and practiced only by the scientists involved in discovering the basic chemistry and physics. These scientists generally had little knowledge of the needs of biologists, and biologists were largely ignorant of recent advances in analytical chemistry and physics. These early sequencers were developed by scientists in university laboratories interested in the applications. There was little support from the academic establishment for developing tools rather than doing research in a particular area of biology. Commercial enterprises provided little input or support, and there was no mechanism for collaborative efforts between academic science and industry.

In the period between 1977 and 1983 automated instruments for both DNA and protein sequencing were developed by Leroy Hood and coworkers at Cal Tech. Plans to produce commercial sequencing machines were presented to 19 different companies, and rejected by all. Applied Biosystems was formed in 1983 by Hood, Hunkapiller, and colleagues, and the Model 377 automated DNA sequencer was introduced in 1987. The Model 470 Protein Sequencer employing automated Edman chemistry was introduced earlier, and was a more immediate commercial success. About 6000 Model 377 DNA Sequencers were sold between 1987 and 1997, and approximately \$100 million was invested in improvements. These led to the introduction of the Model 3700 that provided the performance needed for the human genome project. The rapid completion of the human genome after 1997 can be ascribed to several factors including shotgun sequencing, formation of Celera Genomics, and the infrastructure developed within the Human Genome Project, but without the 3700 or equivalent, completion would still be years in the future.

Functional Proteomics

Phase I of functional proteomics and metabolomics essentially began in 1988 with the discovery of electrospray and MALDI, and the vast collection of DNA sequences now available have substantially accelerated progress, but we contend that substantial further developments in mass spectrometry integrated with separations and bioinformatics are essential for completing Phase I. We believe that new TOF technology may provide an important tool, but other tools are at least equally important. Global proteomics is clearly much more challenging than sequencing the human genome, and much more powerful tools are required for significant progress in Phase II of the proteomics analog to sequencing of the human genome.

Phase III, understanding human and other proteomes (and genomes), may well occupy the biological research community for 100's of years, but inclusion of proteins and metabolites in these studies will expedite rather than impede this research. A beginning for Phase III might be defined as understanding in detail the biochemistry of the living cell, often described as "functional proteomics" and "metabolomics". An excellent review of functional proteomics has been published by Godovac-Zimmerman and Brown. [9] Their definition of functional proteomics is as follows:

"We define functional proteomics as the use of proteomics methods to monitor and analyze the spatial and temporal properties of the molecular networks and fluxes involved in living cells. Conversely, functional proteomics is also the use of proteomics methods to identify the molecular species that participate in such networks via functional stimulation, perturbation, or isolation of these networks."(p15)

Other quotations from this review that amplify the definition are given below.

"An ultimate ideal goal for proteomics is to be able to monitor all cellular proteins. There are several requirements for this goal: all proteins must be quantitatively extracted from the original biological material; the proteins must be resolved and displayed; each protein must be accurately quantified; and the identity of each protein must be determined."(p29) "Identification of proteins only at the gene level will miss much of the important biology."(p43) "Eventually the construction of a functionally useful, complete virtual proteome for an organism will certainly require identification of displayed proteins at the phenotypic level." (p43)

"For many functional proteomics studies, characterization of proteins at the phenotypic level will be essential. These analyses will be substantially more complex than the gene level identifications for several reasons. First it is necessary to characterize the entire peptide sequence of a protein. Second, post-translational modifications must be fully detected. Third, phenotypic characterizations may have to be carried out on mixtures of small numbers of proteins. Finally it will be essential to identify markers of functional activity such as incorporated isotopes or chemical labels." (p46)

This review of functional proteomics provides a dramatic demonstration of the recognized fact that proteomics is orders of magnitude more demanding than DNA sequencing. Mass spectrometry is clearly required for functional proteomics, but the present status of available techniques and instruments falls far short of these requirements. New and more powerful tools are badly needed. Demonstration of these tools may signal the beginning of the end of Phase I of the technology development required for functional proteomics. In Phase II of the human genome project the sequencing effort was concentrated in a small number of large laboratories funded by the Human Genome Initiative or by Celera. We envision a different approach for proteomics in which a large number of laboratories, equipped with the necessary platform technology, will conduct large-scale studies on problems pertinent to their research. Results will be stored in local databases that are linked through a public database that makes all of the information available to the entire biological research community in a timely fashion. A number of difficult issues must be addressed to make this feasible and it will require some modifications in the focus of both the funding agencies such as NIH and the scientific associations such as HUPO.

References

1. Fenn, J., Mann, M., Meng, C.K., Wong, S.F.: Electrospray Ionization of Large Biomolecules. *Science* **246**, 64-71 (1989)
2. Karas, M., Hillenkamp, F.: Laser desorption ionization of proteins with molecular masses exceeding 10,000 Daltons. *Anal. Chem.* **60**, 2299-2301 (1988)
3. Barber, M., Bordoli, R. S., Sedgewick, R. D., Tyler, A. N.: Fast atom bombardment of solids (F.A.B.): a new ion source for mass spectrometry. *J. of the Chemical Society, Chemical Commun.* 325-7 (1981)
4. Watson, J.D., Crick, F.H.C.: Molecular structure of nucleic acid: a structure for DNA. *Nature* **171**, 737-738 (1953)

5. Strohmman, R.C.: The coming Kuhnian revolution in biology. *Nat. Biotechnol.* **15**, 194–200 (1997)
6. S. Brenner interview, June 24, 2000 cited by K. Davies, *Cracking the Human Genome*, The Free Press, New York 2001, p. 248
7. Celis, J.E.: Cancer Proteomics. *Mol. Cell. Proteomics* **3**, 297 (2004)
8. Vitzthum, F., Behrens, F., Anderson, N.L., Shaw, J.H.: Proteomics: from basic research to diagnostic application. A review of applications and needs. *J. Proteomics Res* **4**, 1096–1097 (2005)
9. Godovac-Zimmerman, J., Brown, L.R.: Perspectives for mass spectrometry and functional proteomics. *Mass Spectrom. Rev.* **20**, 1–57 (2001)
10. Lander, E.: The new genomics: global views of biology. *Science* **274**, 536–539 (1996)
11. Lescuyer, P., Hochstrasser, D., Rabilloud, T. How Shall We Use The Proteomics Toolbox for Biomarker Discovery? *J. Proteome Res.* (2007), ASAP Article 10, 1021
12. Omenn, G.: The HUPO plasma proteome project pilot phase. *Proteomics* **4**, 1235–1240 (2004)
13. Gu, X., Deng, C., Yan, G., Zhang, X.: Capillary array reversed-phase liquid chromatography-based multidimensional separation system coupled with MALDI-TOF-TOF-MS detection for high-throughput proteome analysis. *J. Proteome Res.* **5**, 3186–3196 (2006)
14. Preisler, J., Hu, P., Rejtar, T., Muskovets, E., Karger, B.L.: Capillary array electrophoresis-MALDI mass spectrometry using a vacuum deposition interface. *Anal. Chem.* **74**, 17–25 (2002)
15. Vestal, M.L., Hayden, K.: High-performance MALDI-TOF mass spectrometry for proteomics. *Int. J. Mass Spectrom.* **268**, 83–92 (2007)
16. Vestal, C. H., Parker, K., Hayden, K., Mills, G., Vestal, M. L., Cornett, S., Caprioli, R. Tissue Imaging by 5 kHz High-Performance MALDI-TOF, 57th ASMS Conference on Mass Spectrometry, Philadelphia, May 31-June 4, 2009
17. Spraggins, J.; Private communication. Jan 6, 2011
18. Vestal M. L., Hayden K. New MALDI-TOF MS with Very High Resolving Power and Mass Accuracy. Paper TOC pm 02:30 presented at 56th ASMS Conference on Mass Spectrometry
19. Hayden, K., Gabeler, S. C., Dahl, M.; Vestal, M. L. “MALDI-TOF-TOF with High Resolution Precursor Selection and Multiplexed MS-MS”, 57th ASMS Conference on Mass Spectrometry, Philadelphia, May 31-June 4, 2009
20. Vestal, M.L.: Modern MALDI time of flight mass spectrometry. *J. Mass Spectrom.* **44**, 303–317 (2009)
21. Vestal, M. L. “Design, Optimization, and Performance Evaluation of New MALDI-TOF MS and MS-MS Instruments”, 57th ASMS Conference on Mass Spectrometry, Philadelphia, May 31-June 4, 2009
22. Sanger, F., Nicklen, S., Coulson, A.R.: DNA Sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci. U. S. A.* **74**, 5463–5467 (1977)
23. Maxam, A.M., Gilbert, W.: A New Method for DNA Sequencing. *Proc. Nat. Acad. Sci. USA* **74**, 560–564 (1977)
24. Connell, C., Fung, S., Heiner, C., Bridgham, J., Chakerian, V., Heron, E., Jones, B., Menchen, S., Mordan, W., Raff, M., Recknor, M., Smith, L., Springer, J., Woo, S., Hunkapiller, M.: Automated DNA sequence analysis. *Biotechniques* **5**, 342–348 (1987)
25. Hood, L.E., Hunkapiller, M.W., Smith, L.M.: Automated DNA sequencing and analysis of the human genome. *Genomics* **1**, 201–212 (1987)