Journal of Chromatography B, 877 (2009) 1963–1981

Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Review

The role of capillary electrophoresis–mass spectrometry to proteome analysis and biomarker discovery

Farid E. Ahmed*

Department of Radiation Oncology, Leo W. Jenkins Cancer Center, The Brody School of Medicine at East Carolina University, Greenville, NC 27858, USA

ARTICLE INFO

Article history:
Received 5 October 2008
Accepted 10 May 2009
Available online 21 May 2009

Keywords:
CSF
2DE
Diagnoses
ESI
LC
MALDI
Plasma
SELDI
Urine
Validation

ABSTRACT

CE offers a low running cost, short separation time, and a high-resolution technique that requires only a small amount of analyte. It has a wide variety of operation modes (CZE, CEC, CIEF, CITP, CAE, CGE and MEKC) that can be interfaced with MS for tissue and body fluid analysis, particularly urine and cerebrospinal fluid, to identify potential proteomic markers for the clinical diagnosis of many diseases (renal, genitourinary, vascular, diabetes mellitus, cancer, arthritis and neurological diseases) and for the monitoring of their therapeutic intervention. It has become evident that no one marker would be sufficient, but a combination of well-selected markers would be needed for that purpose. The potential of CE coupled to MS for studying the pathophysiology of these diseases and the development of biomarkers has been demonstrated. These biomarkers, when validated, will allow greater use of noninvasive methods for diagnosis of diseases, assessment of their progression, and monitoring individuals' response to therapy.

Contents

Abbreviations: Ab, antibody; ACE, affinity capillary electrophoresis; ANN, artificial neural network; APCI, atmospheric-pressure chemical ionization; BGE, background electrolyte; CART, classification and regression tree; CE, capillary electrophoresis; CEC, capillary electrochromatography; CGE, capillary gel electrophoresis; CIEF, capillary isoelectric focusing; CID, collision-induced dissociation; CITP, capillary isotachophoresis; CSF, cerebrospinal fluid; CZE, capillary zone electrophoresis; ECD, electron capture dissociation; EOF, electroosmotic flow; ESI, electrospray ionization; FFPE, formalin-fixed and paraffin embedded; FT-ICR, Fourier-transform ion cyclotron resonance mass spectrometry; HLPC, high-performance liquid chromatography; HT, high throughput; HUPO, The Human Proteome Project; IEF, isoelectric focusing; IRMPD, infrared multiphoton laser desorption; kDa, kilo Dalton; LC, liquid chromatography; IA, immunoaffinity; IT, ion trap; ITP, isotachiphoresis; LIF, laser-induced fluorescence; LOD, limit of detection; LTQ, linear quadrupole ion trap; MALDI, matrix-assisted laser desorption ionization; *m*/*z*, mass-to-charge ratio; MEKC, micellar electrokinetic chromatography; MS, mass spectrometry; MS/MS, tandem mass spectrometry; PSA, prostate specific antigen; PCa, prostate cancer; p*I*, isoelectric pH; PTM, post translational modification; Q, quadrupole; QIT, quadrupole ion trap; RPLC, reversed phase liquid chromatography; SDS, sodium dodecyl sulfate; SELDI, surface-enhanced laser desorption ionization; SPE, solid phase extraction; SVM, support vector machine; TOF, time-of-flight.

* Current address: GEM Tox Consultants & Labs, Inc., 2607 Calvin Way, Greenville, NC 27834, USA. Tel.: +1 252 321 7261; fax: +1 252 258 1008.
E-mail address: GemToxCocsultants@yahoo.com.

1570-0232/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2009.05.023

1. Introduction

A basic principle common to all electrophoretic methods, including CE, is that a charged molecule is moved through a liquid or a gel by an electric force. In capillary electrophoresis, the conduit for the analyte's electrophoretic mobility is a capillary in which the sample is separated into its different ionic parts, which are also detected within the same capillary, resulting in a self-contained system.

CE is a technique that initially has been used for high-resolution separation of peptides in an open quartz U-tube since the 1920s [1,2], and was refined thereafter in 1967 to alleviate the thermal effects produced due to the high current employed [3]. This was followed by the use of polyacrylamide-filled glass tubes to detect small amounts of proteins by staining and autoradiography during the 1970s [4,5]. In the late 1970s and early 1980s, free-zone electrophoresis was rediscovered when capillaries with an i.d. of <100 μm were used [6,7]. However, uncharged particles were not separated unless surfactants were added to the electrophoretic buffer leading to the development of MEKC, which allowed separation of both neutral and charged analytes [8]. Improved separation by various modes during the 1980s such as CIEF [9], gel-filled capillaries or capillary gel electrophoresis (CGE) [10], and gel-coated capillaries [11] contributed towards better separation. The use of various detection modes (e.g., UV, fluorescence, LIF, conductivity, amperometry and MS) further enhanced the capabilities of CE [12].

The introduction of two commercial CZE instruments in the late 1988 [the Paragon CZE 2000®, P/ACE MDQ, by Beckman-Coulter, Fullerton, CA, coupled on-line to an ESI-TOF MS (micro-TOF™, Bruker-Daltonik Inc., Bremen, Germany), and the Capillarys® by Sebia, Paris, France, which utilizes liquid flow electrophoresis], has led to more applications of CE in the clinical setting [13]. More recently, Beckman-Coulter has introduced the ProteomeLab™ PA-800 Protein Characterization System for high-resolution CE separation of proteins by automated SDS, using a formulation of polymers optimized for the resolution of proteins with a wide molecular weight range (10–225 kDa).

2. CE modes, coupling to MS, preconcentration and capillary coating

2.1. Operation modes

CZE, the most popular mode applicable to clinical proteome analysis, offers a high separation efficiency, small sample volume (a few hundred nanoliters compared to milliliters in HPLC), short analysis time, high resolution, inexpensive capillaries, compatibility with many volatile buffers as generally required for detecting ESI, and a setup with a low-risk of analyte loss. The flow rate in CZE is much lower than HPLC, and the analyte migration is determined by the strength of the electric field. This setup leads to a stable, constant flow, without the need for buffer gradients, or the need for continuous adjustment of the ionization voltage for optimum ionization, in addition to compatibility with many buffers. The flow rate in CZE is much lower than HPLC, and the analyte migration is determined by the strength of the electric field. This setup leads to a stable, constant flow, without the need for buffer gradients, or the need for continuous adjustment of the ionization voltage for optics and analytes [14]. CZE separation, based on different migration velocities (i.e., the ratio of charge to size) and a running buffer that determines the effective analyte charge, is the most commonly used electrophoresis separation mode [15] because volatile electrolytes required for sensitive detection are available for a wide pH range (<2 up to >12) [15], with separation spanning <1 min to microsecond, and could be combined with HT analysis leading to multiplex CZE systems [12].

A persistent problem with CE is that many of the buffers normally used are not sufficiently volatile to carry proper separation. Therefore, CE–MS is less broadly used today, compared for example to HPLC coupled to MS, probably due to reasons such as lack of sensitive detectors at the time when CZE was first used, the limited sample volume employed, lack of buffers with sufficient volatility, the coupling of CE to MS represented a greater challenge than the HPLC–MS interfacing, the reduced robustness of CE–MS compared to LC–MS, and the poor reproducibility of the method, all represent factors that resulted in CE being less routinely used by the mass spectrometry community compared to the more widely used LC. Major characteristics of the various CE modes are summarized in Table 1 [16–30].

From a separation perspective, proteins pose unique challenges as their net charge can be positive, negative or neutral depending on whether they are below, above or at their p*I*, respectively. Thus, their electrophoretic mobility in CE depends primarily on the pH of the separation medium, the size (frictional coefficient), and under native condition on their shape (tertiary structure). By manipulating their solvent pH, sieving power and EOF, hydrophobic proteins can be separated due to differences in the charge-to-effective size ratio [31].

2.2. Coupling of CE to MS

While the flow off an HPLC is controlled by the pump and the effluent can directly be ionized using high voltage, this is not the

Table 1
Characteristics of various CE modes, and their advantages/disadvantages when combined with a mass spectrometer.

Mode	Separation feature	Principle	Characteristics	Strengths/weaknesses
CZE	Charge-to-size ratio	pH of the running buffer determines effective analyte charge.	Flow rate in CZE is low, and the analyte migration is determined by the strength of the electric field leading to a stable, constant flow, without the need for buffer gradients, or continuous adjustment of the ionization voltage for optimum Ionization. It is often used as a fast final dimension before ionization for MS analysis [16].	Volatile electrolytes and organic solvents are available in a broad pH range (<2 to >12), such as alcohol, acetonitrile and acetone, which is convenient for MS applications, although biological matrices are known to be water-based; thus, drastically limiting method's separation utility when compared to LC technique for example.
CEC	Hybrid of chromatograph with EOF	Chromatographic separation combined with an electrostatically driven separation.	Combines the advantage of the flat flow separation profile of EOF with the versatility of the loading capacity of chromatography. Packed columns are mostly employed for separation, although open tubular formats gained wider use because of their simplicity and the possibility of commercialization [17].	Because it can be run at a high speed without compromising resolution, it is a suitable technique for a second dimension in 2D separations, and is well suited for separation on microfluidic devices [18]. However, inconsistent repeatability of migration times and peak areas represent problems for robust laboratory work or MS interface.
CIEF	Isoelectric point	Running buffer containing ampholytes generating an electric-stable pH gradient.	A variant of CZE that combines the high resolution power of conventional gel IEF with CE instrumentation. Because of its focusing effect, it is often used as the first step in multidimensional separation of complex mixtures of peptides and proteins [19].	It provides highest efficiency for protein separation and higher sample capacity compared to CZE [20,21]. Major limitations are the incompatibility of the nonvolatile ampholytes with ESI-MS, which can be solved by using a dialysis interface prior to ESI coupling [22].
CITP	Mobility	Discontinuous buffer system between leading and terminating constituents creates different separation zones.	The leading electrolyte is composed of an ion possessing high mobility while the tailing electrolyte contains an ion of relatively low mobility. Upon the application of voltage, the sample components migrate into separate but adjoining zones according to their electrophoretic mobilities.	It provides higher loading capacity compared to CZE. However, the low separation efficiency and difficulties in finding appropriate spacer limits its acceptability It is often used as a preconcentration step before the separation step [23].
CGE	Size	Gels and sieving polymers form a net that permits size-exclusion separation.	The combination of CGE with MS is desirable for size-sieving purposes such as DNA sequences or analysis of intact proteins.	Gels or monomeric impurities strongly decrease the ionization efficiency of ESI and cause contamination of the mass spectrometer [10], although the BGE additive poly(N-vinylpyrrolidine) [24] or a size-sieving gel [25] can permit the analysis.
CAE or ACE	Migration velocity	BGE contains substances that form complexes between analytes and specific substrates.	Useful for analysis of molecular interactions and for the determination of association or dissociation constants of the formed complexes [26].	ACE may generate asymmetric peaks, which could result in erroneous assessment of migration times. Fitting of peaks with mathematical functions may, however, lead to better estimates of migration time [27].
MEKC	Electrophoretic and electroosmotic mobility of ion pairs with opposite charges in EOF-driven system	Surfactants (anionic SDS, cationic, zwitterionic compounds or bile acids) added from charged micelles, allow for separation of neutral analytes.	Differences in migration are mainly due to variation in the compartmentalization characteristics of the uncharged analytes across the micelles [28].	Because SDS – when employed – strongly inhibits ESI efficiency, there is a need to prevent these micelles from entry into the MS by methods such as APCI that involves selectivity different from ESI, counter-migrating micelles or by partial-filling techniques [29], although a MEKC based on perfluorooctanoic acid and ammonia gave good results [30].

Modified from reference [12].

case for CZE where the flow rates are much lower and the analytes do not merely follow the liquid flow, but their migration is mostly determined by the electric field strength. This feature creates a very simple separation principle based on charge that facilitates a homogenous separation and contributes to the power of CZE. However, when required to be interfaced with a MS, it creates a physical problem because an electric field requires both a cathode and an anode in order to be stable. While one of the electrodes can easily be interfaced with the capillary using the buffer at the inlet, the other electrode at the outlet, which has to serve as the interface to the MS, was more difficult to achieve [14].

Coupling of CE to MS could lead to a high throughput (HT) manipulation, exceptional ability to resolve complicated spectra, and quantitative information for proteomic analysis. The development of soft ionization techniques such as MALDI and ESI in the late 1980s coupled to CE has contributed to successful analysis of proteins [32]. Although MALDI-MS has a higher tolerance towards impurities at moderate levels than ESI-MS, main constraints in hyphenation with CE generally arises from restricted compatibility with nonvolatile BGEs, gels or ampholyte compounds. Therefore, sample desalting before MS detection was attempted by protein precipitation with ethanol/chloroform, followed by sample reconstitution [33], or through SPE by means of ZipTip® pipette tips racked with C_{18} material [34]. Off-line hyphenation of CZE and MALDI-MS requires either direct sample desorption into a MALDI target, or CE fraction collection with subsequent MS analysis [35–37].

While direct sample deposition onto the MALDI target can be carried out continuously, fraction collection demands a transient interruption of the electrophoretic separation. At hydrodynamic fractionation, the induced laminar flow causes band-broadening and deteriorates resolution [38]. CIEF has been employed as an off-line micropreparative tool for analyzing model proteins by MALDI-TOF MS. Proteins separated by CIEF were mobilized by pressure and simultaneous application of voltage, with fractionation in 20 μl collection capillaries carried out via a sheath flow Tee connection providing an ammonium hydroxide buffer solution [39].

Combining CIEF with nano-RPLC separation was shown to reduce the complexity of eluted peptides prior to MS detection using an IT MS (e.g., LTQ, ThermoFinnigan, San Jose, CA) and increase in yeast proteome coverage. Future enhancement in the overall peak capacity was realized by simply increasing the number of CIEF fractions collected to 30, which yielded ~8.8% more protein identifications than using four 15-CIEF fraction runs in half the time with a quarter of the total sample usage. This platform was shown to be capable of measuring changes in protein expression as low as 1.5-fold following multiple testing runs [40].

ESI is by far the most frequently applied coupling technique to CE because it allows molecules in the liquid phase to be transferred directly into ions in the gas phase, permits the determination of high molecular weight molecules by bringing about their multiple charging, and separates mostly charged analytes, at least in the frequently applied CZE mode, which makes ESI the ionization method of choice for polar and ionic compounds. Compared to LC–ESI-MS, two additional requirements apply: the electric CE circuit needs to be closed at or close to the sprayer, and very low and BGE-dependent flows need to be handled [31,41].

A general problem with respect to the robustness in the CE–ESI-MS coupling is to prevent the CE current from having an influence on the electrospray current. Currents in CE (>50 μA when using a 90 cm long capillary, i.d. ≥50 μm) are typically three orders of magnitude higher than those found in the electrospray (typically in the 100 nA range). Therefore, a separation of the electrical circuit of CE from the circuit of the EST source is required considering that electrochemical reactions are involved in both processes, possibly influencing the ionization efficiency of the electrospray. This is important, particularly in the case of negative separation potential combined with positive ESI selection. Moreover, the application of high CE current is desired – as a high electric field strength and a concentrated BGE enable fast and efficient separation – as long as heating does not influence the separation. Grounding the ESI needle and applying the ESI voltage on the MS inlet directs all electrical energy from CE to the ground and yields an undisturbed field for ionization. However, only a few manufacturers of ESI-MS instruments employ this hardware configuration [42]. Other challenging instrumental aspects address the absence of an outlet buffer reservoir, and the necessity to complete the electric circuit [38]. Several types of interfaces are currently in use tackling these problems separately as illustrated below and in Fig. 1 [41]. They fall into two classes: (a) liquid supported (or microspray) systems, in which the voltage is applied to the CE buffer *via* a support as exemplified by the popular sheath liquid, or the liquid junction interface, and (b) nonliquid supported systems, in which the voltage is applied directly into the CE buffer with electric contacts either at the border between the separator capillary and sprayer tip, or with electric contact by in-column electrode [43].

2.2.1. Sheath-flow (SF) or sheath-liquid (SL) interface

Developed by Smith et al. [44], in this most commercialized design the separation capillary is surrounded by a second tube of large diameter in a coaxial arrangement. A supportive sheath liquid circumflows the end of the capillary guided through the outer tube either by an external pump or hydrodynamically and mixes with the CE buffer directly in the Taylor cone, and closes the electric circuit. This arrangement, illustrated in Fig. 1a, offers more flexibility for BGE selection (e.g., adding liquids of different surface tension or conductivity) and, therefore, it has been used in a wide range of separations. The CE column is inserted into the atmospheric region of the ESI source through a narrow metal tube that delivers a sheath liquid to the CE column exit. As the liquid flows from the tube it mixes with the column effluent and forms a stable electrospray. Additionally, a third concentric tube may deliver a nebulization gas flow "sheath gas" that can assist in spray formation *via* nebulization and/or scavenging of free electrons to prevent corona discharge. The supply of a coaxial liquid enhances the robustness and delivers a stable spray, which is useful for separations with a low EOF. However, by diluting the analytes, SF reduces sensitivity when compared to sheathless interface. Apart from the electrostatic parameters that should be optimized in order to achieve a suitable separation of the analytes, there are several parameters involving the ESI itself that should also be considered and optimized (i.e., SL composition, SL flow, position of the capillary tip, dry gas temperature, and nebulizing gas flow and pressure) in order to yield a high MS signal and, therefore, low limits of detection in the femtomole range for peptides [29]. Achieving a stable electrospray operation with SL is often a balance of multiple parameters such as capillary position, liquid sheath flow rate and ESI conditions [43]. Use of SL enhances the possibility of optimizing the electrospray process, making it possible to reduce the influence of nonvolatile components, or to improve selectivity [45]. Because of the high coupling stability, the benefits of this form outweigh the lower sensitivity. Moreover, efficient ionization and detection limits approaching the high attomole range can be achieved when the flow rate is reduced to between 200 and 500 nl/min [46].

2.2.2. Sheathless (nanospray) interface

This nanospray interface was proposed by Olivares et al. as the optimum coupling design due to its compatibility with liquid flow and high sensitivity that can be achieved [47]. Since its first introduction, the sheathless interface has experienced several improvements. The main difficulty in this construction, however, is to close the electrical circuit required for CE separation by means

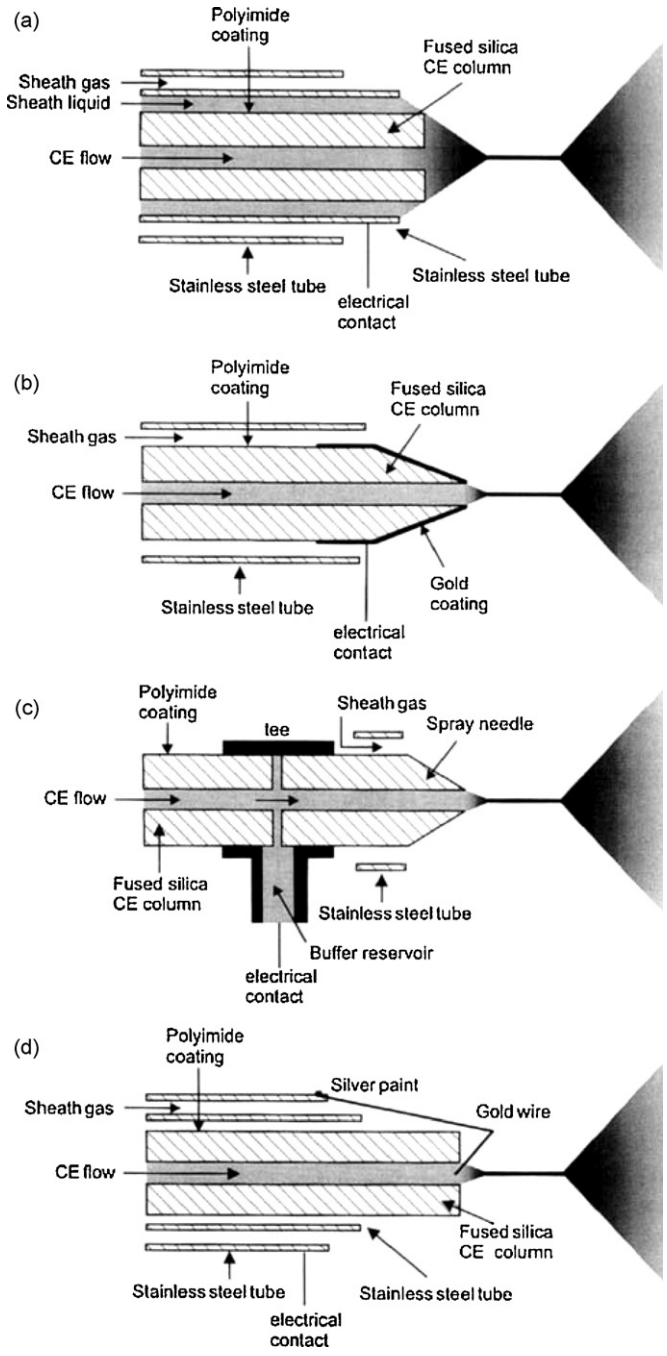

Fig. 1. Different types of interfaces employed for coupling CE to MS: (a) sheath liquid (SL); (b) sheathless (nanospray) with metal coated emitter tip; (c) liquid junction (LJ); and (d) nanospray interface with inserted metal wire (direct electrode). From reference [41]; with permission.

such as coating the capillary outlet with a conductive metal such as gold, polymer, or inserting a conductive wire into the outlet of the capillary, as illustrated in Fig. 1b and d. Moreover, the fabrication of this interface is time-consuming, and its robustness and stability need improvement. Different methods were used to produce durable gold coatings, and different coatings of the inner capillary surface to control the EOF were tested [48].

The sheathless interface was compared to the coaxial sheath flow coupling with respect to performance and sensitivity. While the sheathless interface provides limits of detection (LOD) for a set of standard peptides between 0.1 and 5 fmol, the coaxial sheath flow interface showed reduced sensitivity (3.1–21 fmol). However, in the sheathless interface an EOF close to zero, as produced from neutral coatings or plastic capillaries, cannot be used. To avoid potential corona discharge, SF_6 can be applied as a sheath gas [49].

The sheathless coupling in which the electric field is established using the outer coating of the capillary as one of the two electric poles was reported not to give satisfactory result. This is because the metal coating of the capillaries experienced a high rate of attrition due to the high energy of the electrical current, and to salts present in the buffers that readily react with coating leading to quick erosion. In addition, deposits formed at the tip of the capillary within a short period of time (5–50 min) result in an unstable spray [14]. A graphite coating was reported to be more stable than the metal coating, and therefore might be suitable for CE–MS coupling even in HT analyses [50].

A comparison of ESI sheath flow with sheathless interface that used a mixture of peptide hormones showed both interfaces to give similar LODs in the range of 1–3 μg/μl, although reproducibility was lower with the sheathless compared to the sheath-flow interface (30–80% RSD compared to 10–15% RSD) [51].

2.2.3. Liquid-junction (LJ) interface

First developed by Henion and co-workers [52], the LJ interface is by far the least frequently used interface in coupling CE to MS. The CE column is inserted into a low dead volume tee where the electrical connection for closing the electric circuit with the ES spray needle is provided *via* a buffer reservoir as shown in Fig. 1c. An ESI emitter capillary is positioned opposite the end of the separation capillary with a gap of ca. 10–25 μm between both in such a way that the liquid and analytes from the separation capillary pass to the emitter tube and are sprayed afterwards. The main advantage of this setup is that CE and ESI can be operated independently due to the partial electrical and physical disconnection of the CE separation from the ESI emitter. The main disadvantages have been the difficulty of positioning the transfer capillary in a reproducible way, and the potential loss of separation efficiency in passage through the ESI emitter [53]. Moreover, the ESI emitter can originate a counter hydrodynamic flow in the CE capillary, and, in some cases, a pressure (typically 20–40 mbar) has to be applied to the inlet capillary to counterbalance this effect. Although new improvements on the original design have been reported, some of the original difficulties (i.e., peak broadening) frequently remain, and the construction of the junction between the CE capillary and the ESI emitter continues being a difficult task [29,38]. The LJ approaches based on the connection of CZE to emitter tips did not result in a homogenous contiguous field and gave occasional satisfactory ionization, but poor and nonreproducible resolution in the CZE made these approaches obsolete [14].

2.3. On-line preconcentration

Sensitivity is an important issue in CE separation. The absolute sensitivity of CE–MS is almost independent of the separation method, but depends on the ESI process, the MS used and the mode employed. The concentration sensitivity is proportional to the amount of solution loaded on the CE, which is much less than HPLC. Moreover, in complex biological samples, it is often necessary to perform a purification step prior to separation. Both considerations have necessitated the need for on-line and off-line preconcentration procedures. Some of the on-line approaches to increase CZE–MS sensitivity are outlined below.

2.3.1. Stacking

The simplest and most popular on-line preconcentration procedure is sample stacking. In this method, a sample plug is introduced in with a lower concentration or a buffer with a higher pH than

the separation buffer. Because the mobility of the analytes in the sample buffer differs from that of the analytes in the separation buffer, the sample will focus at the interface between the two buffers. Recently, pH-mediated stacking, using a plug of 1–2 M NH_3 before the sample and 4 M formic acid after the sample, was used with CE–ESI-MS. Using this technique, a 10-fold improvement in sensitivity can be achieved. The so-called large volume sample stacking using electroosmotic flow pump (LVSEP) technique can be employed as an alternative; it operates on the same principle, but requires a capillary coating, which compresses EOF. This technique allows a concentration factor of about 100-fold [54].

Transient CITP in combination with CZE was shown to provide selective analyte enrichment through electrokinetic stacking in which major components may be diluted while trace ones are concentrated. When coupled on line with a nano-ESI-MS, this combination results in an ultrasensitive detection of trace peptides at a subfemtomole mass in complex peptide mixtures with minimal band broadening, and the selective enrichment of trace peptides enables the identification of low abundance peptides comigrating with highly abundant species at a concentration ratio of 1:500,000, which is one to two orders of magnitude more sensitive than conventional electrophoretic or chromatographic-based proteome technologies. Moreover, the speed of CITP/CZE separation and the lack of column equilibration improve the throughput and facilitate integration of this technology with other separation technologies in a multidimensional protein identification platform [55].

For samples in high conductivity media, a sample stacking method that did not require desalting has been developed. For this method, a polyethylene oxide solution was prepared in a 400 mM Tris-borate buffer and followed the injection with a short plug of low pH buffer. Concentration factors of more than 100 were reported [56]. Another technique used a porous joint to connect the concentration region of the capillary to the separation region. This technique has been suitable for the concentration of samples in low ionic strength solutions, acidic solutions, and dissolved in running buffer [57].

2.3.2. Transient ITP

ITP separates the sample into a series of zones between a leading and a tailing ion. This method increases the injected sample volume without significant loss of separation efficiency and resolution. The analytes are first concentrated in discrete sample zones between leading and terminating electrolytes. When the zones disintegrate, CZE separation takes place. The technique was applied to the CZE–ESI-MS qualitative analysis of a tryptic digest of cytochrome *c* and to the quantitative analysis of melagatran and two endogenous peptides. LODs were in the range of 6 nM corresponding to an absolute detection limit of 3 fmol [58]. Amphoteric analytes, like peptides, can be preconcentrated by pH-mediated stacking, which also comprises transient ITP. This method enhanced the loadability by at least 10-fold [42].

2.3.3. Solid phase extraction (SPE) on-line preconcentration

SPE uses the partitioning of molecules into a solid phase (typically C_{18}-coated particles) to extract hydrophobic analytes from dilute aqueous solutions. A method to directly identify proteins in complex peptide mixture by micro-SPE prior to CE that used a sheathless liquid metal junction to interface CE to ESI-MS a multistep elution procedure was employed in the mid 1990s [59]. The micro-SPE served as the first separation dimension and the CE as the second dimension. The multistep elution procedure together with reduced voltage CE extended the resolving power of CE separation to increase the number of acquired MS/MS spectra. Moreover, reduced elution CE extended the analysis time window by reducing the CE running voltage during the period when a large number of peptides were migrating out of the capillary. Eighty to ninety percent of the proteins in the yeast ribosomal complex were identified with that multidimensional CE–MS/MS approach [59].

On-line SPE–CZE–MS have been developed with membranes or stationary phases packing near the capillary injection end. Samples are passed through these devices, washed with BGE, and then an organic solvent was applied to elute the analyte before voltage application [60]. A transient ITP has often been used to further concentrate the zone of eluted analytes before voltage application. In a membrane preconcentration transient ITP–CZE–MS method for sequence determination of tumor peptides using a porous glass injection and a μ-ESI interface, a polystyrene divinyl benzene (PSDVB) membrane preconcentration cartridge was installed at the head of the column facilitating the concentration of the sample. HPLC fractions from a human melanoma tumor cell extract were analyzed by mPC-transient ITP–CZE–MS. Detection limits of less than 50 amol in full-scan and MS/MS modes were observed [61]. An on-line RP-C_{18} particles preconcentrating column, for the concentration and cleaning of proteolytic digests and major histocompatability complex class I bound peptides, resulted in a 10-fold sensitivity improvement and a 100-fold increase in sample volume introduced in the capillary [62].

2.4. Capillary coating

Dynamic coating is less common in CZE–ESI-MS due to their detrimental effect on the ESI process and lack of reproducibility [38]. An undesirable effect in CE separation is the interaction of the analytes with the inner wall of the CE capillary. Under neutral or basic conditions, the silanol groups of the fused silica capillary dissociate and form a negatively charged surface at the inner capillary wall, which interacts with the positive charge on the protein leading to peak broadening, deposit of analyte material in the capillary, poor reproducibility in migration time, and/or low protein recovery rate. To circumvent these problems, several different coatings and coating protocols of covalent nature have been described. In general, the main approach consists of using coatings covalently bonded to the capillary wall that can bear positive, neutral or negative charges. Since most coatings result in a positively charged capillary wall and causes an inverse electroosmotic flow, the electrical field must be reversed. Many of the coatings are tedious, time consuming, with low stability at extreme pH, and of high price for some commercially coated capillaries. On the other hand, the benefits of an ideal coating include very high resolution leading to lower detection limit and reproducibility [14,29].

The cationic polymer polyprene has been a key coating in CE, and in conjunction with acidic buffers, it has been developed early on for CE separation of glycoproteins and glycopeptides in conjunction with ESI-MS detection [63]. Polybrene capillary coating with an in-house on-line CE combined with IT storage/reflectron TOF MS detection was capable of resolving and identifying a large number of digested peptides from hemoglobin variants in a short time (10–15 min) [64]. Use of polybrene in conjunction with ammonium acetate at a pH of ~7.4, narrow capillaries for high separation efficiency and forward polarity CE to avoid acid production and the tip of the capillary were reported as important factors for a successful analysis of peptide complexes [65]. Coating with a bilayer of polybrene and poly(vinyl sulfonic acid) was shown to produce better migration time reproducibility and high separation efficiencies when using UV and MS detection [66]. This double coating yielded consistent and fast CE-TOF MS patterns of amino acids in urine [67], and for profiling of amino acids in CSF and urine with minimal sample pretreatment [68].

Although some coatings such as the aminoalkylsilyl-based [69] one and those with a copolymer (EpyM–DMA) synthesized from 2-ethyl-(2-pyrrolidine) methacrylate (EpyM) and *N*,*N*-dimethylacylamide (DMA) [70] or with a polyamine (PolyE-

323) [71] that provided a fast separation of a few minutes with high efficiency, good deactivation, robust anodal EOF and no bleeding into the mass spectrometer over a pH range of 2–10 and tolerance towards methanol and acetonitrile (two modifiers commonly used in CE–ESI-MS), were both applied successfully in separating peptides and proteins in standard solutions and real samples.

For microfabricated devices, static coating necessitates the use of a broader range of material. Therefore, the difficulty lies in the introduction of derivatization reagents. Glass chips can be coated with acrylamide and hexadimethyldisiloxane (HDMS) [72], and polycarbonate microfluidic devices have been employed [73].

Unsatisfactory results were reported for all coatings when clinical samples (e.g., urine) were used. The P/ACE MDQ and PA-800 commercial CE instruments are equipped with bare fused silica capillaries. The optimal approach, therefore, has been to employ uncoated capillaries and use BGE with low pH to reduce the negative charge on capillary surface, which reduces capillary–protein interaction. Moreover, hydrophobic interactions have been reduced by adding an organic solvent (such as acetonitrile) to the BGE [14].

3. MS instrumentation

An expedient tool for the analysis of a large number of complex proteomic samples appears to be CZE using uncoated fused silica capillaries, coupled via a grounded sheath-liquid interface to a mass spectrometer. While the CE–MS analysis in itself does not pose a major problem, several challenges become obvious when HT analysis is contemplated: (a) sample preparation must be highly reproducible irrespective of the protein content, (b) no overloading of any system is permitted, (c) large polypeptides and proteins are frequently denatured at the low pH used, and tend to precipitate in the CE capillary, and (d) the software tools provided by the manufacturers of mass spectrometers are inadequate to analyze the pattern of numerous complex samples in the clinical context, and require several days to complete [14].

CE has been coupled with different types of mass analyzers, i.e., ion traps (ITs), quadrupoles (Qs), Fourier-transform ion cyclotron resonance (FT-ICR), sector field (E, B) and time-of-flight (TOF)-type instruments. Qs have been the most commonly used mass analyzers for CE–MS; however, their slow scanning speed, low resolution and their relatively low sensitivity in full scan modes limit their use [21]. ITs acquire data over a suitable mass range with rates of several spectra per second, making them suitable for MS^n experiments. However, in full scan mode the resolution of an IT is generally too low to resolve the single isotope peaks of greater than three-fold charged molecules, which makes assignment of charges to these peaks difficult [14]. TOF mass analyzers have a high duty cycle, good sensitivity and high resolution. Modern TOF instruments can record >20 spectra/s, summing approximately 1000 spectra per data point, providing a resolution ($m/\Delta m$) of >10,000, and mass accuracy better than 5 ppm. Hadamard transform (HT)-TOF-MS has been coupled with CE, and can sample peaks having width in the millisecond range [74], although it has not shown wide applications. New generation FT-ICR MS with magnetic fields higher than 9 Tesla is capable of sub-attomole sensitivity at nearly 1 s acquisition times [75]. FT-ICR has been generally too slow to be important in CE–MS [76] as the acquisition time is too slow for fast, high resolution CE or for MS/MS.

Other than CID, FT-ICR offers two complementary new ion fragmentation technologies for analysis of PTMs by MS/MS: ECD and infrared multiphoton dissociation (IRMPD). Fig. 2 represents analysis of biomarkers from urine samples (second urine of the morning) separated by the P/ACE MD Q CE system coupled with a hybrid FT-ICR MS [a linear QIT FT-ICR, LTQ-FT mass spectrometer (Thermo Electron)] [77]. Moreover, FT-ICR instrument facilitate the identification of polypeptides >10 kDa [78].

TOF MS is fast and simple [79,80], and is suitable for coupling to CE. Modern TOF instruments can record up to 20 spectra per second and benchtop ESI-TOF MS provide a resolution of more than 10,000 and a mass accuracy >5 ppm [14], making them particularly suitable for interfacing with CE instruments.

The high cost of sophisticated MS equipment (e.g., Q-TOF, FT-ICR) has traditionally prohibited routine coupling of these powerful instruments to CE, although recently their prices have been decreasing due to the introduction of newer technologies and increased use, the wide use of various MS instruments or their combinations, the ease of their coupling and operation, and the entrance of several manufactures (e.g., Thermo-Fisher, Bruker Daltonik, Agilent, Shimadzu, Hitachi, Waters, etc.) into massive MS production for various proteomic users [32].

4. Chip-based CE–MS

Recently, microfabricated devices have been developed to perform CE on chips rather than capillaries. Advantages include: (a) high speed due to short separation channels, (b) HT due to parallelization, (c) different injection schemes and miniaturization steps enable the handling of very small sample volumes, and (d) the two spatial dimensions and variable manufacturing possibilities enable the integration of CE separation on-chip combined with enzyme reactions, sample dilution, derivatization, concentration, and purification, etc. [29,81,82]. The ability of multielement microfluid platform for the analysis of phosphoproteins, which will eventually allow detection of PTMs, has recently been demonstrated. Microchip technology has also been coupled to MEKC to provide HT and high-performance analytic system. Although MEKC is generally incompatible with ESI-MS due to the contamination of the ion source by micelles, the use of the partial filling (PF) technique and pseudostationary phase (PSP) in MCEK–MS is expected to overcome the limitation [83].

An analytical system for the rapid identification of low nanogram amount of yeast proteins separated by a 2DE method consisted of integrating solid phase microextraction/CZE peptide separation device that was connected through a microelectrospray ion source to MS/MS. LOD was 660 amol of sample at a concentration of <33 amol/microlitter for peptide mass measurement and <10 fmol of sample at a concentration limit of <300 amol/μl for peptide analyzed by CID [84].

The ability to handle limited amounts of cellular protein extracts, and to perform detection at the ultra trace level in a HT fashion is a challenge in many areas of biological sciences and in chemical settings. Microdevices will offer significant advantages, particularly in the context of integrating on-chip sample clean-up, preconcentration and microdigestion prior to separation and mass spectrometric detection. The complexity of some of these samples imposes a need for extensive sample prefractionation for achieving chromatographic resolution, and ultimately sensitive and efficient MS detection. Integration and parallelization of analytical processing steps will be essential for promoting HT capabilities, and could potentially result in the fractionation of inexpensive and disposable platforms that prevent sample contamination and carryover, an essential requirement for the reliable analysis of proteomic samples [81,85].

Furthermore, the capability to perform high-speed separation with microchips will be fully matched by HT MS, especially TOF MS. With respect to separation, CE is now widely utilized on chips and is expected to find extensive applications with MS detection [86]. In the long term, the use of parallel processing for separation and subsequent MS analysis would appear to have significant

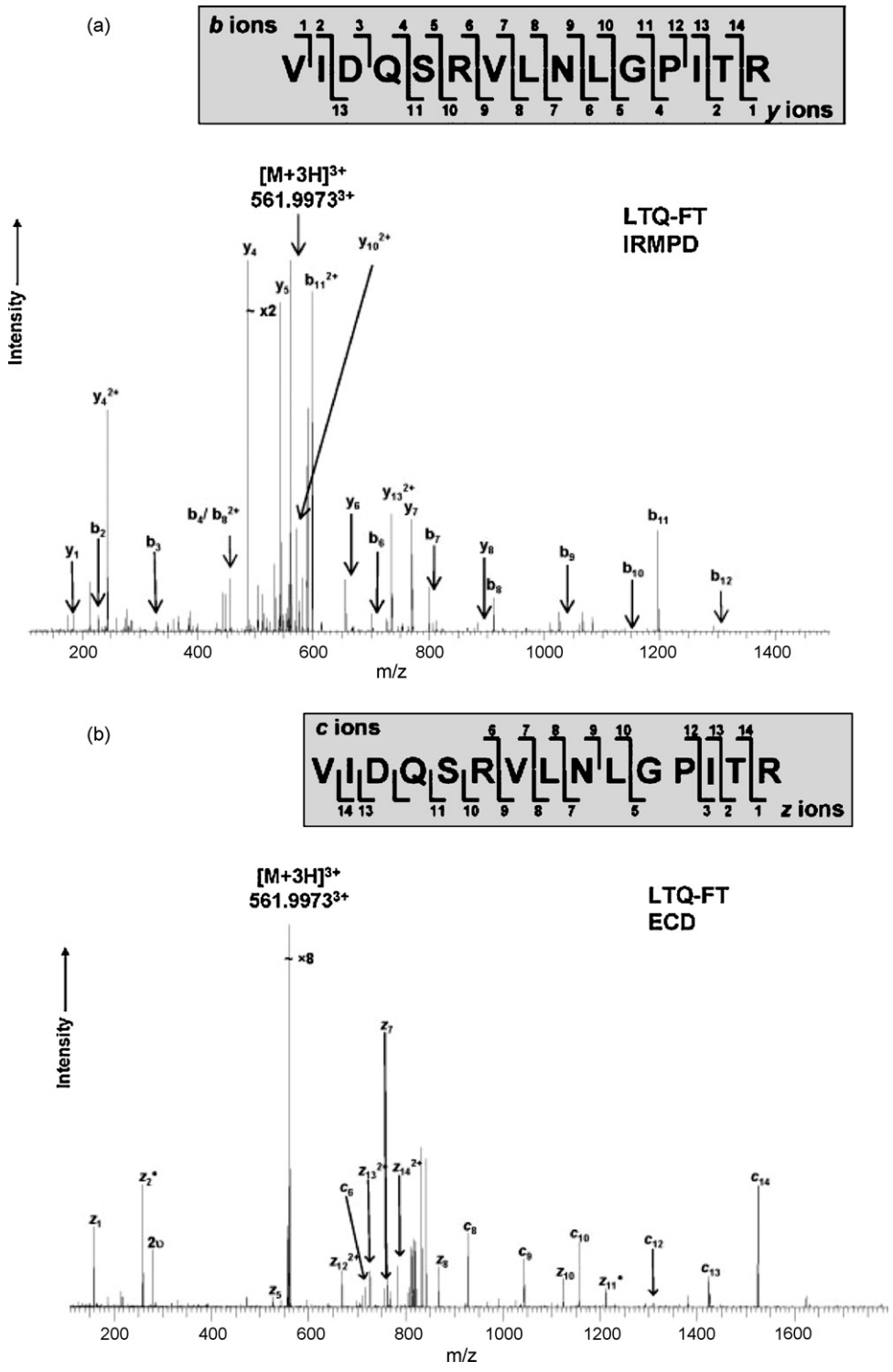

Fig. 2. (a) IRMPD LTQ-FT MS/MS product ion spectra of a triply charged parent ion (561.9973^{3+}) of the urine peptide uromdulin that was separated by the P/ACE MDQ CE instrument. (b) ECD LTQ-FT MS/MS fragment spectrum of the $[M+3H]^{3+}$ precursor ion of the urine peptide uromdulin. Ion fragmentation method IRMPD exhibited several advantages over CID for MS/MS in a QIT (as can be seen by comparing (a) with (b)). From reference [77]; with permission.

potential. This seems to be particularly true in the case of multiplexed, microfabricated separation devices used in conjunction with MALDI-MS detection. With high repetition rate lasers, one to two orders of magnitude above those presently used (i.e., kHz), it is possible to envision parallel processing of separation of 10 to >100 channels simultaneously. The development of miniaturized and multiplexed mass spectrometers will further stimulate the need for major advancement in sample processing instrumentation. Such approaches would ultimately lead to ultrahigh throughput investigations. A number of companies (Advion, Agilent, Gyros, Predicant) have launched efforts to couple novel microfluidic analysis systems with mass spectrometric detection and are developing bioinformatics technology for fast analysis of proteomic and/or clinical samples [81,85].

Earlier work on parallel HT proteomics separation for multiple on-line and off-line infusion MALDI-MS and capillary array elec-

trophoresis (CAE) has been pioneering in the CE–MALDI MS parallel processing arena [87]. An array of eight infusion capillaries was used to vacuum deposit eight individual solutions of a mixture of peptide and matrix simultaneously on an advancing Mylar tape directly in the source chamber of the TOF mass spectrometer. A fast scanning mirror allowed automatic determination of the position of the eight sample traces, after which samples were desorbed with a kilohertz laser and spectra were recorded with a HT data acquisition system. For CAE–MALDI/TOF MS, multiple samples were coupled to eight infusion capillaries using a common liquid junction, deposited on the Mylar tape and analyzed in parallel by the mass spectrometer. Mass resolution above 2000 fwhm (full width half maximum) and low-attomole detection limits were obtained for small peptides [87].

5. Data processing and bioinformatics

Clinical proteome analysis is a multidimensional assay that encompasses comparative analysis of large numbers of variables that exceed the number of samples to be analyzed by orders of magnitude, which necessitates more analysis time and effort. In addition, to obtain statistically significant data, the increasing number of analyzed components requires increasing the number of analyzed samples and consequently the need for greater computing power. Therefore, it is important to find a balance between the desire for maximal data for analysis, and the limitation of needed effort and analysis time [88].

To analyze the vast amount of raw data produced in a single CE–MS, which consists of 500–1000 time sequenced mass spectra, 80,000 data points each that feature Gaussian and non-Gaussian noise, as well as significant baseline offsets produced by unidentifiable analytes across the detection range from 400 to 2500 m/z, several software tools have been developed such as MSight for LC–MS [89], or MosaiquesVisu for CE–MS [90] that can cope with pattern recognition rather than using a single molecule identification to automatically select peaks. MosaiquesVisu accessible at www.proteomiques.com performs a step wise examination of CE (or HPLC) MS spectra. This software was developed for peak identification, deconvolution and the display of refined maps in three-dimensional (3D) formats. The simplest form to depict data from MS is total ion count chromatogram, which is a summary of all spectra collected over the migration time in CE and consists of approximately 1000–1500 single spectra obtained every 3 s, with an error rate below 27, as judged by the 200 most abundant polypeptides, thereby allowing the evaluation of highly complex spectra [90].

In the first step of pattern analysis by MosaiquesVisu, signals are collected fitting the criteria of a certain signal-to-noise level and being present in several consecutive spectra. In the next step, the collected MS peaks (m/z) are charge deconvoluted into one mass based on isotopically resolved peaks, and the data are stored in the database. In the last step, the MS peaks are examined for conjugated peaks up to a charge of 50. This is particularly important for nonresolved MS peaks. The conjugated peaks (and their amplitude) are combined to a single peak, which is characterized by mass and migration time, as well as the combined amplitude of the conjugated peaks. Based on the mass deviation of the conjugated peaks as well as the match of the isotopic distribution, the software also calculates the probability that this peptide, which is calculated based on the calculated charge really exits [14].

Application of MosaiqueVisu for a urine sample of a human volunteer leads to a 3D picture, as shown in Fig. 3a, of the data that contain information on the migration time and signal intensity, ranging from 0 to 25,000 MS counts, which are stored in an access

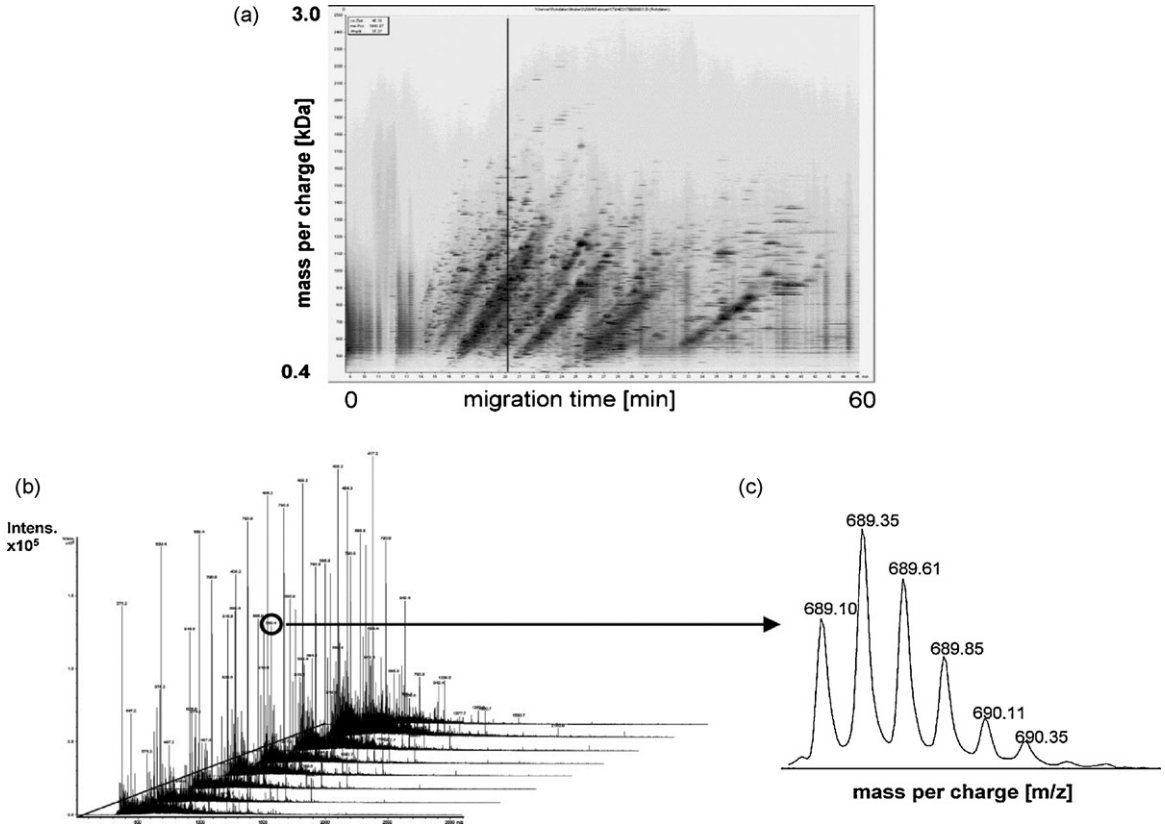

Fig. 3. (a) Contour plot of a typical CE–MS run for a urine sample of a healthy human volunteer; m/z is plotted against the migration time; the signal intensity is reflected by the grey scale. To show the high amount of information present, the eight spectra contained in the area indicated by the grey line are shown in (b). Further magnification of a single peak is shown in (c). From reference [108]; with permission.

database and matched to the peak lists of other samples. Eight consecutive spectra are shown in Fig. 3b. Peaks of different samples are presumed identical if the mass deviation is less than 100 ppm and the migration time deviation is less than 5 min. The magnitude of one peak is shown in Fig. 3c, yielding an isotopic distribution of a particular peptide. Generally, between 900 and 2500 polypeptides with molecular weight from 800 to 3000 kDa can be detected. The raw data are subsequently analyzed for the presence of a real signal and the background is eliminated [7,14,46,91,92].

Generated data are submitted to data mining algorithms in order to minimize the error rate. Data mining strategies fall into two categories: unsupervised (approaches that do not take into account class labels and are analogous to clustering) such as *k*-means clustering, principal component analysis (PCA) and hierarchical clustering; and supervised (approaches that take into account class labels and are analogous to classification) such as classification and regression trees (CART), Bayesian classification, artificial neural networks (ANN), genetic algorithms, and support vector machines (SVM) [93]. Each algorithm has inherent strengths and weaknesses, which must be matched to the specific statistical problems to be addressed [94,95].

To perform feature reduction and classification, the data are generally divided into a training set and a validation set. Preprocessing (e.g., baseline subtraction and peak detection) is performed on the full complement of data, and the data from the validation set are set aside until the time of validation. If the size of the training set is large (i.e., >60 in each class), the training data set can be further subdivided into a training and a test set. Feature selection and model training are performed on the training set and their performance is assessed in the test set; this can be repeatedly performed to assess multiple sets of features and multiple algorithms. If the training data set is too small, as frequently happens, typically cross-validation is performed. Cross-validation, a method by which the sample size may be artificially increased, can still be performed in the training set even if it is a separate set. When cross-validation is performed on algorithms generated on a preselected set of variables, rather than selecting the variables with each iteration, this choice may overestimate the suitability for generalization of an algorithm, and bias may occur [96]. An alternative approach to minimizing selection bias in a proteomic experiment is to perform cross validation during the feature selection process [94].

Regardless of the specific feature selection and classification approaches used to identify the most important classifying peaks in the diagnostic setting, the simplest, most transparent algorithm is desired. Alternatively, it may be advisable to combine the optimal classifiers into a logistic regression model. Additionally, discriminant analysis may generate attractive diagnostic models, which can be linear, quadratic or nonparametric (*k*-NN). For any given model, it is important to access the overall error rate in the cross-validated sample, and to determine how stable the model is and whether it is truly different from a random situation [94].

Since most MS-based proteomic studies suggest that a single biomarker does not provide sufficient specificity, several biomarkers are often combined to form a signature pattern, which results in an increased ability to distinguish between diseased and healthy states, and increases the stability of the pattern because the absence of atypical amplitude of a single peptide does not lead to a significant change in the typical pattern [14,92]. This concept requires a two-step process in which biomarkers are first identified employing statistics for multiple testing, and subsequently they are combined in a predictive model using some of the algorithms depicted above (i.e., CART, SVM, ANN, etc.). CARTs were among the first algorithms to combine multiple markers [96]. However, these approaches are weakened as the number of incorrect predictions made by the classification algorithm increases with the complexity of the decision tree [97]. SVMs provide a tool to overcome some of these limitations due to the theoretical principles upon which they are based [98]. An ANN approach utilizes all available data set to find ideal separating features and can also be used for classification [99] as has been reported for analysis of breast cancer [100] and hepatocellular carcinoma [101].

Unfortunately, SVM or ANN, like black boxes, is unable to provide levels of confidence to any classification. Therefore, the clinician is a left with no information on the statistical significance of such a prediction, which makes them reluctant to use these methods in the clinic because of insurance malpractice considerations. A promising classification method that shares many of the positive characteristics of the SVM, but in addition provides the levels of confidence with each classification prediction, is based on the Gaussian process [102]. No matter which of these mathematical approaches is used, two basic considerations apply: (a) the number of independent variables should be kept to a minimum and should certainly be below the number of samples investigated, and (b) an approach is valid only when it is tested with a blinded validation set in order to determine whether the findings can be generalized across parameters such as patient demographics, clinical subsets and laboratory handling techniques [103], which should be included in any report on potential biomarkers.

The discrimination between different disease groups could be performed with MosaCluster [92]. This software tool generates a model for each of the different diseases investigated based on the polypeptides which are best suited to discriminate disease from control (or other diseases). Each polypeptide used for classification represents one dimension in an *n*-dimensional space. The software allows the classification of samples in the high dimensional parameter space by utilizing SVMs. Such algorithms have performed well for the evaluation of clinical markers [97,104] and for biological evaluations, such as DNA analysis [105].

A key factor in the comparative examination of clinical samples is retrieval of identical polypeptides in consecutive samples. Two parameters are used to assign and tentatively identify a peptide: migration time and molecular weight [106]. Migration time varies depending on the ion content of the sample, and MS signal intensity varies depending on the efficiency of ionization as well as the detector gain. Thus, these parameters have to be normalized by using external standards, or by utilizing a set of polypeptides that are found with high frequency in the sample and therefore serve as internal standards [97,107]. Finally a list of analyzed and clearly identified and standardized sample peptides enables the digital compilation of individual data sets to a specific pattern that can be used for biomarker identification [108].

6. Alternative separation/analysis approaches

From earlier work reported by O'Farrell [109], separation of proteins in 2D configuration, according to the p*I* and molecular mass (Da) in order to produce high-resolution fractionation of complex protein mixtures, is considered the basis of today's proteomics; despite its shortcomings, 2DE still seems the preferred method for the analysis of large proteins [110]. SELDI initially appeared to solve many of the limitations of 2DE–MS. It is based on reducing sample complexity by the selective interaction of peptides and proteins with different reagents in solution on surfaces of chips with different immobilized matrices, which allows for the enrichment/purification of a few fraction of all polypeptides present in the analyte and generation of simple mass spectra from highly complex samples by eliminating interfering compounds, and subsequently characterizing few potential biomarkers from highly complex samples [111]. Shortly after its use, drawbacks of this approach were fiercely voiced in the pre-, analytical and post-analytical steps

[32,112], which limited the mass range attained and diminished separation capabilities for complex mixtures.

LC is a powerful attractive separation method in which large amounts of samples can be loaded onto the LC column [113]. Therefore, if sensitivity is a major consideration, LC is a superior alternative to CE. Multidimensional LC (MudPIT) approaches provide vast amount of analytical information [114], but they tend to be time consuming, and consequently not suited for routine clinical analysis that often require HT methods. LC is less well suited for the separation of larger molecules and analytes covering a broad range of size and hydrophobicity. Moreover, LC is by far more sensitive than CE towards interfering compounds and precipitates [14].

The electrophoretic mobility of a protein or a peptide is proportional to its actual charge and is inversely proportional to its size. The factor that appears to mostly affect electrophoretic mobility is charge or valence. Hence, pH is considered to be the most significant variable for altering CE resolution of peptides and proteins [115].

Electrophoretic migration properties of peptides separated by CE as a function of pH can be predicted from theoretical, semiempirical models, which ultimately depend upon Stoke's law (that can be used to describe the motion of an ion in an electric field) [116]. It is thus possible to predict the electrophoretic mobility of the peptide from a protein digest from its amino acid content, and hence allow for detection of PTMs in CE separation when combined with MS detection [117].

In acidic conditions, CE–MS is not well suited for the analysis of proteins >20–30 kDa, as a fraction of these proteins frequently precipitate in the capillaries, making them unavailable for subsequent MS analysis [82]. For larger proteins, the ProteomeLab® PF2D system (Beckman-Coulter), which is an automated, 2D fractionation system expressly designed for high resolution analysis of complex protein mixtures for down stream proteomic analysis that uses IEF in the first dimension followed by nonporous RP-HPLC selectivity, has been shown to work effectively and reproducibly separating basic proteins [118] as well as highly hydrophobic microsomal proteins [119].

Protein microarrays allow the simultaneous determination of a variety of proteomic parameters in parallel from minute amounts of samples, thus allowing for a HT analysis of translated gene functions and identification of biomarkers in tissue and body fluids. However, there are issues of sensitivity, specificity, difficulty in maintaining the native state of the protein upon surface immobilization, and limitations of current arrays that must be overcome to achieve HT applications and minimize the occurrence of false positive and negative results [120]. The various advantages and disadvantages of application of several discussed protein technologies are summarized in Table 2 [108,119].

The lack of standards and comparability among different proteomic methods could be one of the major limitations of proteome analysis. Recently published suggestions for mandatory standards and guidelines for collection, storage and preparation of samples, as well as requirement for analytical performance and quality control [121] are expected to improve that situation dramatically. Lack of appropriate and user friendly bioinformatics software for data analysis hinders progress towards clinical application. A repository of all data in a common file, together with specific software solution would be an excellent step towards establishing comparable data. In addition, lack of sequence data for many potential biomarkers requires improvement in software solutions for sequence assignment to shrink these shortcomings, as well as solutions for improvements in the detection limits of MS/MS instruments [122].

7. Sample preparations

Sample preparation and fractionation technologies are one of the most crucial processes in proteomic analysis in solubilized samples. Chromatographic or electrophoretic proteomic technologies are also available for separation of cellular protein components. There are, however, considerable limitations in currently available proteomic technologies as none of them allows for the analysis of the entire proteome in a simple step because of the large number of peptides, and because of the wide concentration dynamic range of the proteome in clinical blood samples. The results of any undertaken experiment depend on the condition of the starting material. Therefore, proper experimental design and pertinent sample preparation is essential to obtain meaningful results, particularly in comparative clinical proteomics in which one is looking for minor differences between experimental (diseased) and control (nondiseased) samples [81].

The field of sample preparation for proteomics in general is still considered in its infancy. There is no general standardized strategy for overall sample preparation, separation or purification. The ideal sample preparation protocol should insure minimal loss, or at least reproducible loss of polypeptides and analytes [81]. A crude unprocessed sample should theoretically be analyzed in order to avoid artificial loss or bias arising from the preparation process. How-

Table 2
Advantages and limitations of various proteomic methods.

Method	Advantages	Disadvantages
2DE–MS	Applicable to large molecules, high resolution, allows visualizing changes in molecular mass (M_r), isoelectric point (p*I*) or PTM	Not applicable to peptides <10 kDa, time consuming, hydrophobic or low-abundant proteins or those with extreme p*I*s or M_rs are poorly represented
SELDI	Ease of use, HT, automation, low sample volume, affinity capture, various chip surfaces	Loss of important information, low resolution, limited mass range and limited separation capabilities for complex mixtures, problems at pre-, analytical and post-analytical steps, bias towards high abundant proteins particularly in the low-mass range, performance could change over time
LC–MS	Automation, highly sensitive, multidimensional, MS/MS compatibility, accurate, quantitative	Time consuming, sensitive towards interfering compounds, limited mass range
CE–MS	Automation, relatively sensitive, low sample volume needed, low-cost, MS/MS compatibility	Not well suited for polypeptides >20 kDa, precipitation of polypeptides in capillaries occurs when acidic running buffers are used
IA-CE	HT, ability to analyze low-abundance biomarkers, provides significant enhancement in sensitivity for the purified and enriched affinity tagged analytes, potential for miniaturization and portability	May produce false positive or false negative results characteristic of Ab-based assays due to lack of adequate specificity and accuracy
Arrays	HT, low sample volume, chips have potential for assaying a wide range of biochemical activities, various platforms and detection methods available	Antibodies are not availed for all screened proteins, no standardization is available for biomarker discovery, low sensitivity, qualitative

Modified from references [108,120].

ever, this approach does not appear practical due to the presence of interfering compounds (e.g., aggregates, lipids, carbohydrates, salts, etc.). CE is relatively insensitive towards interfering compounds, but the high salt content of the crude sample could interfere with the CE separation process. Moreover, pre-MS separation is needed in order to cope with samples' complexities and the large dynamic range of proteins in biological material (between 10^7 and 10^{12}) [32]. An efficient analytical method(s) is thus desired to remove unwanted substances without interfering with the composition or representation of proteins in samples. It appears that the use of anion exchange or reverse-phase material fulfill this purpose [14,108].

Major drawbacks to the general acceptance of multidimensional purification strategies are their being technically demanding, time consuming, need to be optimized for recovery and reproducibility, not easy to automate for HT analysis and requirement for extensive MS analysis time that could result in data analysis bottlenecks. Alternative platforms such as magnetic beads, multiwell plates or chips could address many of these issues. Simple batch wise elution techniques can lead to effective fractionation that is amenable to a HT automated application. Use of multidimensional techniques for MS analysis can reduce sample complexity and also increase the number of samples analyzed. Improvement in data processing and analysis by integrating these processes into a linear process will increase overall efficiency [88].

More investigation is needed focusing on reduction of sample complexity, developing promising sample preparation methods such as an improved multiaffinity removal system, multidimension LC, and use of nonporous solid phases. Subcellular fractionation allows access to intracellular organelles and multiprotein complexes; low abundant proteins (LAPs) and signaling complexes can be enriched, and at the same time, the complexity of the sample can be reduced [123].

Fractionation of protein mixtures to isolate species by their common biological activity is an approach that is not yet well established, not because of lack of interest, but rather because of lack of effective methods for immediate implementation. Selection of protease activities would no doubt have a strong impact on the understanding of specific pathway regulations with direct interest in diagnostic. Most proteases are part of very low abundance species that might stay silent for long periods, but can be detected by their specific peptide signature [124]. Presented below are strategies for preparation of various biological samples.

7.1. Frozen tissue samples

The local concentration of the biomarker is expected to be high in the vicinity of the tumor microenvironment. Therefore, fine-needle aspiration biopsies (FNABs) are one way to obtain these samples [32]. Because many different cell types are typically present in tissue biopsies, laser microdissection (LMD) techniques have been developed to provide a rapid method for separating and processing homogenous subpopulation of cells for biochemical analysis [125]. Use of LMD may subject samples to potential artfactual processing, including changes at two different stages: (a) during the stage of tissue sections that enables selection of the relevant cell types, and (b) during the dissection process itself. These changes could impact the level of protein recovery and the quality of subsequent proteomic studies [126]. Isolated cells and captured minute tissue samples can then be directly analyzed using MALDI-MS, or through the use of an automated multidimensional HT separation platform that combines CIEF with another separation technology such as nano-RPLC [127]. The high analyte concentrations in small peak volumes as a result of electrokinetic focusing/stacking and the resolving multidimensional separation results in sensitive proteome analysis by enhancing the dynamic response and detection sensitivity of the coupled MS instruments. Instead of performing multiruns or multidimensional separations, comparable or even better HT proteome results could be achieved by simply increasing the number of CIEF fractions due to the intrinsic high resolvation nature of electrokinetic focusing, a feature that is particularly important for proteome analysis of limited tissue samples [128].

7.2. Formalin-fixed and paraffin-embedded (FFPE) tissue samples

Besides fresh frozen tissue-based proteomic studies, combined CIEF/nano-RPLC separations equipped with ESI-linear IT MS have been employed to study soluble protein profiles extracted from an archival formalin-fixed and paraffin-embedded (FFPE) tissue of human renal carcinoma, in which tryptic peptides obtained from proteolytic digestion of protein extracts using the heated antigen retrieval (AR) technique at pH 7 are systematically resolved by their difference in p*I* and hydrophobicity. When five FFPE sections were processed with the AR conditions, the amount of tryptic peptides for performing FFPE proteomic analysis was ~5–10 μg, which is comparable to protein extract levels obtained from microdissection-produced tissue specimens [129]. Another method to identify proteins from FFPE cancer tissue is a method termed direct tissue proteomics (DTPs), which uses a shotgun proteomic approach combined with a protein extraction procedure that disrupts cross-linked proteins in order to identify potential diagnostic biomarkers from prostate cancer tissue [130]. Comparison of proteome results of FFPE tissue with fresh frozen tissue, procured from the same case of renal carcinoma, showed that >70% of proteins were identified in both tissue sections in spite of the potentially deleterious and negative impact from the FFPE process on protein extraction [128].

Combined CIEF/nano-RPLC multidimensional separation platform coupled with ESI-linear ion trap (LTQ)-MS were analyzed for HT profiling of membrane proteins within cell pellets of microdissected ovarian carcinoma specimens containing the serous cell type. The SDS detergent-based sample preparation protocol provided effective protein solubilization and complete proteolytic digestion and was compatible with the CIEF/nano-RPLC. A total of 18,861 distinct peptides were detected, leading to the identification of 3303 non-redundant SwissPort protein entries. Among proteins identified, 723 were predicted to contain one or more transmembrane domains corresponding to 22% membrane coverage within the SwissPort [131].

7.3. Body fluids

Because of their easy accessibility compared to tissue biopsies, many of the diagnostic, prognostic or monitoring response to therapy biomarkers used in clinical practice are found in biological fluids (blood, urine, cerebrospinal fluid, etc.). Body fluids are very complex mixtures of molecules with a wide range of polarity, hydrophobicity, and size over a range of several orders of magnitude. When analyzing complex biological samples, major concerns are loss of polypeptides and information as well as reproducibility. Ideally, a crude, unprocessed sample should be analyzed, which would avoid all artificial losses or biases arising from sample preparation. Since all body fluids contain a large amount of different ions, lipids, carbohydrates, etc., these samples cannot be analyzed in the native form in a mass spectrometer. In addition, pre-MS separation is a prerequisite in order to cope with the complexity and dynamic range of these samples [14,110].

CE is quite insensitive towards interfering substances like lipids, carbohydrates, salt, but also towards small amounts of aggregate and large proteins. This could allow the injection of even crude urine samples. However, an increased breakage of glass capillaries was

observed when crude urine samples were used, probably due to the high salt content of these samples. Therefore, it seems preferable to remove salt and other low molecular weight compounds, prior to CE runs, with the use of either anion exchange, or RP material (e.g., RP-C-2 columns that performed better than the RP-C-18 for larger polypeptides) [14,90,91].

7.3.1. Urine

Approximately 70% of all urinary proteins originate from the kidney, whereas only 30% are derived from plasma. Disruption of the glomerular barrier and/or tubular injury can result in increased proportion of plasma protein in urine (proteinurea) when urinary protein excretion exceeds 150 mg/d in adults. Among body fluids, urine is especially attractive for biomarker discovery in urological diseases for many reasons: (a) it is more organ specific than blood, which contain proteins for only a few organs that are located directly along the path of urine production and excretion (i.e., kidney, urinary tract, including bladder). However, normal urinary protein can also originate from glomerular filtration of plasma proteins, secretion of proteins from renal tubular epithelial cells, shedding of whole cells along urinary passage and from exosome secretion, (b) it can be obtained in large quantities using noninvasive procedures, (c) repeated sampling from the same individual is achievable, facilitating longitudinal studies, (d) urine contains proteins and polypeptides of lower molecular mass (<30 kDa) that are highly soluble, which facilitate analysis of such polypeptides in their natural state without the need for additional manipulation (e.g., tryptic digest) [132], and (e) for proteins >30 kDa, urinary polypeptides are stable and do not generally undergo significant proteolysis within several hours of collection, in contrast to blood where activation of proteases and the generation of an array of proteolytic breakdown products is often associated with collection [133].

Two independent studies showed that the urinary proteome did not change significantly when urine was stored up to three days at 4 °C, or up to 6 h at room temperature [134,135]. It is believed that the stability of urine emanates from its storage for several hours in the bladder where proteolytic degradation by endogenous proteases is complete by the time of voiding [132]. On the other hand, urinary exosomes or other fractions of urine may be less stable [136].

Urine has disadvantages as a source for protein markers due to: (a) the wide variation in protein concentration, which is largely due to differences in person's fluid intake. This problem can be mitigated by standardization based on creatinine [137] or peptides present in urine [108], (b) inconsistency of its pH that may alter the activity of proteases in a fraction of the urinary proteome leading to greater variability of the proteome during the day due to factors such as different diets, metabolic or catabolic processes, circadian rhythm, exercise, and circulatory levels of various hormones [138]. However, the basal or housekeeping proteins of urine remains largely unaffected by these changes [108], and (c) clear differences between early stream and mid stream urine samples are found [134]. Therefore, standardization of urine collection protocols is a must [121].

7.3.2. Cerebrospinal fluid (CSF)

CSF is another body fluid amenable to biomarker analysis. CSF is quite specific to the central nervous system (CNS) and the spinal cord, containing less total proteins than the blood (1/200) so that the buffering of protein composition is much weaker in CSF and provides a low fluid-volume-to-organ ratio, thus increasing the probability of high concentrations of proteins or peptides within the fluid. Moreover, transportation of fluids from the spinal cord and brain to the CSF is achieved in a bulk flow manner [139], which implies that molecules of varying sizes have an equal opportunity to be found in CSF. However, 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 welcome by even healthy individuals [140].

Other currently less exploited body fluids such as bronchial fluid, synovial fluid, nipple aspirate fluid, saliva or amniotic fluid have similar potential as urine and CSF [140].

7.3.3. Blood

In contrast to urine and CSF, blood that is easily accessible and can be collected with minimal invasiveness requires meticulous preanalytical handling, and its proteomic analysis is prone to analytical artifacts [141]. A detailed comparison of serum and plasma proteomes revealed that an array of proteases are activated immediately upon clotting, resulting in the generation of many degradation products. Consequently, the Human Proteome Consortium has recommended that blood be examined as plasma rather than as serum and established standardized sample collection protocols [133].

The development of orthogonal high-dimensional proteomic strategies that include two or more protein separations has been shown to overcome to some extent the complexity of the plasma proteome leading to the detection of a large number of LAPs (<100 ng/ml), where cancer biomarkers are expected to be found, and also the identification of their PTMs when using intact protein fractionation schema together with shotgun LC–MS/MS analysis methods that are associated with the disease; this is a feature that is not readily available with regular protein digest-based fractionation approaches [142].

Standardizing sample preparation procedures for plasma/serum profiling, including the type of collection tubes and coagulants, the clotting and incubation time before sample isolation, storage conditions, strategies used for removal of high abundant proteins (HAPs), as well as fractionation techniques employed either to generate several fractions or to selectively obtain a particular subset of peptides/protein, although time-consuming and error prone is critical for obtaining reliable biomarkers and building a biomarker pattern, since slight changes in a given sample preparation could lead to very different protein profiles [143].

The current strategy of comparing plasma protein profiles in normal *versus* diseased subjects has been inefficient during the discovery phase, mostly due to the tremendous dynamic range of protein levels in plasma, ranging from below the nanogram per milliliter (femtomolar concentration of tumor necrosis factor, TNF) to tens of milligrams per milliliter (millimolar concentration of albumin), which is 12 orders of magnitude [110]. Twenty-two plasma proteins make up 99% of the protein mass. Since there is no amplification step for proteins that is analogous to the polymerase chain reaction (PCR) method for amplifying nucleic acids, it is often preferable to deplete most of the proteins in order to reduce the complexity of the sample and to enhance the representation of the minor protein and bring them above the threshold of detection by current proteomic technologies [32]. However, this approach has also failed for a number of reasons: (a) because of the enormous dynamic range, the removal process has to be close to 100% efficient to be effective. Even a perfect fractionation process that removes the 22 most abundant proteins will leave 8–10 orders of magnitude of protein levels in plasma, which exceeds the current power of proteomic technologies that can measure up to 5 orders of magnitude at the most, and (b) the less abundant proteins may also be eliminated during the depletion process resulting in loss of important information and potential artifactual discrepancies between samples [32]. Therefore, it has been suggested to separate the biomarker discovery phase from the validation phase. For example, in case of a tumor, in the discovery phase comparison should first be made between normal tissue and diseased tissue with two or several well separated classification categories to find differ-

ences, and then look at these differences by focused approaches (e.g., MS- or immune-based) in biological fluids such as plasma. That's because the dynamic range of proteins is 6–8 orders of magnitude in cells, 10–12 orders of magnitude in collected plasma, and 8–10 orders of magnitude in perfectly depleted plasma. Moreover, only few intracellular proteins are glycosylated, which enhances the power of protein separation in cells *versus* plasma [144]. Instead, it has been suggested to look for secreted protein markers released in suitable conditioned media by cells or tissue explants "secretome" in animal models [145] as potentially useful source for biomarker search because it was shown that proteins present in the secretion media at concentrations of nanograms/milligram range can be analyzed by standard proteomic tools [146]. Furthermore, secretomes are considered better samples than complete tissue or cell extracts because there is a bias in favor of easily liberated proteins and against tissue proteins present in large complexes (e.g., cytoskeletal proteins) in these secretomes [144].

A CZE approach for tentative identification of proteins in complex samples like plasma, and for monitoring mass changes on the level of intact protein for diagnostic purpose employed anodic separation (AS), which separates plasma components without sample pretreatment in combination with ESI-TOF MS in the positive detection mode. AS involved injecting plasma analytes directly into a fused silica capillary (60 cm in length, 50 μm i.d.) coated with polybrene to avoid hydrophobic interaction with the capillary wall that leads to the appearance of a precipitate, and to obtain a strong reversed EOF. This separation method was presumed to have two advantages: (a) the proteins are completely desalted because the small cations are migrating in the opposite direction and will never reach the detector, and (b) modified proteins like glycosylated forms are separated from nonmodified forms because of the lower positively charge due to noncharge-bearing groups or negatively charged modifications. Glycosylated proteins are easy to allocate, due to the shifted charge envelope, to higher m/z values [14].

8. Biomarker discoveries in clinical samples

8.1. Urinary proteome

As outlined earlier, the urine seems to be an ideal source of biomarkers. Urinary proteins can be analyzed directly or separated by centrifugation into distinct fractions. For example, supernatants from low-speed centrifugation contain proteins that are derived from filtered plasma proteins and secreted by tubular epithelial cells. This supernatant can be further centrifuged at high speed to yield pellet-containing exosomes (small vesicles with diameter <80 nm with cell membrane and cytosolic proteins). These exosomes are derived from epithelial cells that line the urinary tract with a contribution from filtered exosomes from blood cells [136].

A recent study on urine obtained from healthy individuals identified 1543 proteins, a large proportion of which were membrane proteins, probably due to the presence of exosomes [147]. Diseases directly related to kidney damage such as chronic nephropathy [107,148] and uremic toxicity [91,144] were first to be analyzed in urine using CE–MS. When renal failure occurs, the kidney cannot eliminate waste products adequately. Renal replacement therapy (i.e., artificial kidney) is therefore required in some patients with acute-renal failure (ARF) or chronic-renal failure (CRF) to remove waste products and toxins that cause uremic symptoms and complications. Uremic toxins are small proteins with molecular masses up to ~30 kDa [91].

In studies conducted by the European Uremic Toxin Working Group (EUTOX) using CE–MS, the effect of different dialysis membranes (low flux F10 *versus* high flux F70 membranes) on the number of polypeptides from 1 to 10 kDa in the dialysate was investigated. Peptides above 10 kDa were present only in the dialysates obtained from high flux membranes, while most of the polypeptides in dialysates obtained from low flux membranes were less than 10 kDa [91]. In another pilot study for the potential of CE–MS to identify uremic retention molecules in dialysis fluids obtained with low flux *versus* high flux membranes, results also indicated higher efficiency of removal of larger peptides using high flux membranes [149]. These EUTOX studies demonstrated that the polypeptide profile in these fluids yielded a surprisingly low consensus with that of the urine, indicating that the artificial kidney is not comparable with the native kidney, and renal replacement therapy does not substitute all aspects of renal function [149]. In a subsequent study, CE–MS was used to identify polypeptides present in the plasma of dialysis patients that were not present in the plasma of normal control individuals [150]. Combination of data from human plasma and hemodialysate should be able to identify potential uremic toxins, thereby resulting in increased efficiency of dialysis.

IgA nephropathy (IgAN) is the most common glomerular disease in adults. Early stage renal disease (ESRD) develops in about 30% of the patients; therefore early intervention and therapy may prevent or delay the development of ESRD as currently the only diagnosis is invasive renal biopsy [148]. In a recent study that evaluated the utility of SDS-PAGE/Western blot and CE–MS were employed for detection of urinary polypeptide biomarkers of renal disease in patients with IgA-associated glomerulonephritides in a reference cohort of 402 patients with various renal disorders and 207 healthy controls; CE–MS patterns of renal damage and IgAN were defined. In a blinded analysis of a separate cohort of patients with IgAN ($n = 10$), Henoch–Schoenlein purpura (HSP) with nephritis ($n = 10$), and IgA-associated glomerulonephritis due to hepatitis C virus (HCV)-induced cirrhosis ($n = 9$), and healthy controls ($n = 12$), the SDS-PAGE/Western blot and CE–MS against clinical urinalysis for detection of urinary proteins/polypeptides was compared. Urinalysis indicated proteinurea for 50, 90 and 33% of patients, respectively, and for none of the healthy controls. SDS-PAGE/Western blot showed urinary polypeptides abnormality for 90, 80 and 67% of patients, respectively and for none of the healthy controls. CE–MS indicated a renal damage pattern in 80, 80 and 100 of patients, respectively, and in 17% of healthy controls, with the more specific IgAN pattern in 90, 90 and 1%, respectively, and in none of the healthy controls. Based on differences in CE–MS patterns, the disease mechanisms may differ among various IgA-associated glomerulonephritides. These preliminary findings if confirmed in a future prospective study coupled with renal biopsy and urinary testing may allow adapting the CE–MS method to develop new tests to detect renal injury at earlier stages, assess clinical manifestations and monitor response to therapy in patients with IgA-associated renal diseases [151]. Analysis of peptides in urine and serum of humans by CE was shown to be indicative of not only a particular disease, but also the stage of the disease [152].

CE–MS analysis of urine was also compared with an animal model based array for developing biomarkers for neonatal uteropelvic junction (UPJ) obstruction, a frequently encountered pathology in neonates. In a subsequent step, these sets of urinary biomarkers were used for prediction of the clinical evolution of UPJ obstruction patients. A peptide identified as fragment of α 3 chain of type a collagen was present in 34% of the combined UPJ patient group compared to 100% in the healthy newborn group. Another biomarker, a fragment of type V precollagen α 2 chain was present in only 16% of the healthy new born plus non-operated UPJ and in 76% of the operated obstruction patients. The CE–MS-based strategy allowed prediction several months in advance of the clinical evaluation of neonates with UPJ obstruction [153].

CE–MS was recently used to find biomarkers for diabetes, diabetic nephropathy and nondiabetic proteinuric renal diseases in 305 individuals. A panel of 40 biomarkers distinguished patients

with diabetes from healthy individuals with 89% sensitivity and 91% specificity. Among patients with diabetes, 102 urinary biomarkers differed significantly between patients with normal albuminuria and nephropathy, and a model that included 65 of these correctly identified diabetic nephropathies with 97% sensitivity and specificity. Furthermore, this panel of biomarkers identified patients who had microalbuminuria and diabetes and progressed toward overt diabetic nephropathy over three years. Differentiation between diabetic nephropathy and other chronic renal diseases reached 81% sensitivity and 91% specificity. Many of the biomarkers were fragments of collagen type I, and quantities were reduced in patients with diabetes or diabetic nephrology. This study showed that the analysis of the urinary proteome by CE–MS may allow early detection of diabetic nephropathy, and may also provide prognostic information [154].

In addition to the definition of disease-specific polypeptide patterns, stage-specific polypeptide markers could be defined by CE–MS as for diabetic nephropathy in patients with diabetes mellitus type I or II [155,156]. In both studies, individual datasets of healthy volunteers (9 and 39, respectively) patients with diabetes type I or II without macroalbuminuria (28 and 46, respectively) and with intermittent 16 or persistent 66 macroalbuminuria were combined to create typical polypeptide patterns. In patients with type I and II diabetes mellitus and a normal albumin excretion rate, the detected polypeptide pattern differed significantly from that in patients with greater albuminuria, suggesting that the urinary proteome contains a much greater variety of polypeptides than previously demonstrated [122].

In addition to the diagnosis and prognosis of disease, urinary proteome has been used to develop markers that are predictive for response to therapy. In a randomized double-blind study, macroalbuminuric patients were evaluated for treatment with daily doses of 8, 16 and 32 mg of candesartan or a placebo for two months. Examination of the urine samples from these patients with CE–MS revealed a significant change in 15 of 113 protein characteristic for diabetic renal damage [157]. Similar data were obtained from patients with vasculitis, for whom the vasculitis-specific protein pattern reverted towards normal after treatment [108].

The urinary proteome was also used to develop markers for cancer (particularly bladder and prostatic cancers). In a blinded prospective study of bladder cancer, CE–MS profiled urine sample from 46 patients with urothelial carcinoma were compared with 33 healthy volunteers and 366 patients with malignant and nonmalignant genitourinary diseases to define cancer-associated signatures of polypeptides. The profiled model was applied to a masked group of 31 patients with urothelial carcinoma, 11 healthy individuals and 138 patients with nonmalignant genitourinary disease. All samples of urothelial carcinoma were correctly identified (sensitivity 100% and specificity ranged from 86% for nephroliths to 100% for healthy controls). The diagnostic pattern for urothelial carcinoma included fibrinopeptide A, a known biomarker for ovarian and gastric cancers [158].

Only 30% of patients with elevated serum prostate specific antigen (PSA) levels who have undergone prostate biopsy are diagnosed with prostate cancer (PCa). Therefore, new methods are needed to reduce the number of unnecessary biopsies for this common cancer in males. Urine has been used to diagnose patients with PCa. The heterogeneity of progressive PCa has hampered development of an effective early detection system. CE was used to define potential urinary markers for PCa in a pilot study. Forty-seven urinary samples from patients who underwent prostate biopsy were analyzed (26 patients had PCa and 21 had benign prostatic disease). Several polypeptides were potential markers for PCa with 92% sensitivity and 96% specificity. Data suggested that early-stream urine was the best sample for the definition of PCa-specific biomarkers, indicating that these biomarkers likely originated from prostatic secretions [107]. In a follow up study by the same investigators, the prostate-specific pattern was refined with 116 urine samples from 54 patients with PCa and 62 patients with benign pathology. A pattern of 26 potential markers was validated in a blinded assessment of urine samples from 36 patients with PCa and 24 patients with benign prostatic condition. The model correctly classified 32 of the 36 patients (89%) with PCa and 16 of the 24 patients (67%) with benign pathology [159].

Using CE–MS, a panel of 12 novel biomarkers for PCa was identified and validated in a blinded multicentric prospective study by comparing first void urine of 51 men with PCa and 35 with negative prostate biopsy. In contrast, mid stream urine samples did not allow the identification of discriminatory markers, suggesting that prostatic fluid may be the source of the defined biomarkers. Hence, first void urine samples were tested using a prostatic fluid informative polypeptide panel (IPP). A combination of IPP and PCaP to predict positive prostatic biopsy was evaluated in a blind prospective study [159]. Two hundred thirteen of 264 samples matched the IPP criteria. PCa was detected with 89% sensitivity and 51 specificity. When age and percent free PSA was added to the proteomic signature, the sensitivity and specificity were 91% and 69%, respectively [159].

The discriminatory potential of polypeptides to distinguish PCa samples from those without any evidence of disease (NED) was analyzed using the receiver operating characteristic (ROC) curves [160]. All polypeptides with an area under ROC curve (AUC) $\geq$0.600 (49 of 1459 polypeptides) were statistically analyzed using the Benjamini and Hochberg procedure [161] for multiple testing correction. Initial statistical analysis indicated that PCa samples of the training set were heterogeneous. To address this heterogeneity, an additional bootstrapping process [162] for the definition of additional PCa specific markers was performed. Biomarker candidates were statistically analyzed using nonparametric methods such as Wilcoxon's test (rank sum test) with $p < 0.05$ as significance level. PCa-specific biomarkers were additionally validated using the Benjamini and Hochberg procedure to control false discovery rate in multiple testing [160].

CE–MS was used to analyze urine from patients with different grades of subclinical or acute clinical rejection, patients with urinary tract infection and patients without evidence of rejection or infection. Substantial differences were found between patients with transplanted kidneys and those with native kidneys, most probably due to treatment with the calcineurin-inhibitor immunosuppressant cyclosporine A. Moreover, a distinct urinary peptide platform identified 16 of the 17 patients with acute tubulointerstitial rejection, which differed from markers of vascular rejection. Potentially confounding variables did not affect the results. An additional polypeptide pattern that allowed differentiating between infection and acute rejection was developed. That pattern was validated blindly on samples from transplanted patients potentially exhibiting real rejection resulting in correct classification of samples [163].

CE–MS analysis of urine was used to develop proteomic patterns that could predict acute graft *versus* host disease (aGvHD) after allogenic haemopoetic stem cell transplantation (allo-HSCT) allowed the correct classification of 13 of 13 aGvHD samples (100% sensitivity) and 49 of 50 (98% specificity) for the training set. A subsequent blinded evaluation of 599 samples enabled diagnosis of aGvHD greater than grade II with a sensitivity of 83% and a specificity of 76% [164]. Other diseases, including systemic lupus erthematosus are currently being investigated [165].

Analysis of urinary polypeptides by CE–MS was reported to improve the assessment of patients with arthritis. A 45 mer collagen type II peptide fragment was reported in the urine, probably due to matrix metalloproteinase (MMP) activity. Therefore, the activity of MMP could be monitored *in vivo* by measuring the urinary excretion of particular collagen fragments [166]. Interestingly, several

of the urinary proteome biomarkers were also collagen fragments [135,167], leading to the speculation that they could indirectly indicate the activity of disease-specific proteases [122]. An effort to optimize proteomic methods for urine analysis and set standards for clinical applications was made by establishing the Human Kidney and Urine Proteome Project (HKUPP) under the auspice of the Disease and Biomarker Institute of HUPO [http://hkupp.kir.jp].

8.2. Blood proteome

Few studies were carried out for analysis of plasma with CE–MS to identify biomarkers. Because a high degree of false appendicitis are diagnosed, a screening study of biomarkers to differentiate two kinds of appendicitis (gangrenous and phlegmonous) was carried out on patient plasma samples with CE–MS. Indicative patterns were found for both pre- and post-surgery of the two types of inflammation, as well as between them. Suggested markers were identified, which need to be validated in future studies [168].

CE has been frequently applied to separation of intact proteins (8–29 kDa) in human blood with attomole sensitivity using very narrow capillaries (5 μm i.d.) with gold-coated ESI tips (2–5 μm i.d.) in a sheathless design that is coupled to an FT–MS as shown by the detection of 7 amol carbonic anhydrase (CA) in a crude blood isolation, which corresponds to the concentration of CA in a single red blood cell (RBC) [169]. CZE–ESI-TOF MS was used for the analysis of α and β chain of ~450 amol of a single intact RBC [170]. LODs of 20 and 44 amol were reported for CAI and CAII, respectively, using a split-flow interface and narrow capillaries (15 μm i.d.) for coupling of CZE with an IT–MS [171], which allowed for the detection of all four major proteins of the intact RBC (α- and β-Hb, CAI and CAII) analysis with minimum sample preparation.

A key application of CE–MS in the area of intact protein is the differentiation of modifications like oxidation or glycosylation as shown by the separation, identification and quantification of the reduced and oxidized forms of cytochrome *c* by CZE–MS [172], characterization of complex apolipoproteins in plasma using CZE–ESI-MS by the addition of 10–20% acetonitrile to the BGN [173], and separation of recombinant human erythropoietin glycoforms by an MS-compatible CZE BGE separation [174]. However, routine analysis of the proteome of a complex biological sample without a preceding separation step is still not a routine method, and it will probably remain so in the foreseeable future.

When analyzing complicated plasma samples, it is often hard to find the information that is relevant to the question at issue, unless it is known in advance what markers to look for, which is not often the case. One way of dealing with this problem is to use multivariate data analysis and PCA. When collecting data from CE–MS, there is a risk of drift in migration time as well as m/z from run to run. Such instrumental variations will usually obstruct the use of PCA as a tool for data exploration. Today, the most common ways to circumvent this situation are: (a) to use an alignment procedure prior to the PCA, or (b) to generate peak lists and perform analysis on them rather than the original dataset. An alternative method, the fuzzy correlation approach, has recently been proposed [175] to blur the dataset with respect to time and m/z, thus reducing the effect of instrument drift when comparing the samples.

8.3. CSF proteome

Due to the blood-brain barrier, the diagnostic potential of blood for the detection of nervous system-related dysfunctions and diseases, such as Alzheimer's disease, is believed to be limited. However, CSF can be of diagnostic value. CSF is a clear and colorless liquid that serves as a buffer for mechanical protection of the brain. It contains sugars, lipids, electrolytes and proteins. Compared to blood, 70% of CSF protein content consists of isoforms of albumin, transferrins and immunoglobulins, which generally provide only limited disease-specific information [108].

A procedure for analyzing CSF involve tryptic digestion of CSF polypeptides, followed by an on-line combination of CE with FT-ICR MS for sample analysis identified 30 proteins with a 95% confidence level, with mass measurement errors of less than 5 ppm [176].

A sample preparation protocol for CE–MS investigation of CSF utilizing depletion of high abundant proteins, instead of tryptic digest, was used and detected about 450 different proteins and polypeptides with molecular masses ranging from 800 to 15,000 Da [67]. These results led to a pilot study, as proof of concept, comparing peptide patterns for healthy volunteers ($n = 4$) and patients with Alzheimer's disease ($n = 8$) and schizophrenia ($n = 7$). Differentiation between the last two groups was possible, which suggested that in principal CSF is suitable for diagnostic analysis using CE–MS. However, because of the limited number of samples analyzed, these preliminary results need to be substantiated in additional clinical studies.

An alternative approach for the analysis of low abundant protein markers in body fluids includes analyte enrichment using immunoaffinity-based CE [177] on lab-chips. Inflammatory cytokines were detected in CSF of patients with head trauma by immunoaffinity electrophoresis on a chip. Following incubation with a fluorescent dye (Alexa-Fluor 633), CE separated cytokines in less than 2 min [178].

CE–MS was applied for analysis of human urine and CSF. Analysis of human urine resulted in biomarkers that allowed classification of a variety of different renal diseases and even similar diseases such as focal-segmental glomerulosclerosis (FSGS) and minimal change disease (MCD), and other markers that enabled evaluation of therapy success. Initial data obtained on human CSF strongly suggest that CE–MS analysis of low molecular weight proteins and peptides reveals several potential biomarkers for schizophrenia as well as Alzheimer's disease [179]. While these data clearly indicated the diagnostic potential of the examination of CSF with CE–MS, the diagnostic potential of the method is questionable as CSF examination received very little patient compliance [179].

9. Concluding remarks

The application of CE–MS technology to proteomic analysis provides unique possibilities to solve many clinical problems such as biomarker discovery and potential therapeutic targets for many diseases because of its potential for automation and high resolution, as well as the availability of a wide variety of CE operation modes (i.e., CZE, CIEF, CITP, CGE, CAE and MEKC). CE has many desirable features, including ease of miniaturization, low reagent consumption, short analysis time, low cost, HT capacity, high separation selectivity and efficiency, excellent mass sensitivity and recovery. Nevertheless, it suffers from poor concentration sensitivity detection due to limited sample-volume capacity of the capillary, and the short capillary optical path, which could limit the sensitivity of the detector. Today, however, it is possible to inject microliters or milliliters of sample volume to enhance sensitivity. A broad spectrum of tissues and body fluids (mostly urine, although CSF and blood have been attempted) have been analyzed by CE–MS providing novel diagnostic tools to monitor changes in the composition of body fluids as well as the PTMs of proteins and peptides in widespread diseases such as renal, genitourinary, dialysis, diabetes mellitus, cancer, arthritis, Alzheimer's diseases and schizophrenia. Therefore, it appears that CE will find wider applications in clinical analysis and biomarker discovery than it has experienced today [180,181]. Examples of applications of CE–MS analyses to biomarker discovery in human diseases are summarized in Table 3 [12,91,92,107,148–159,163–167,179,183–189].

Table 3
Examples of applications of CE–MS analyses to biomarker discovery in human diseases.

Disease	Sample source	CE mode	Reference(s)
Glioblastoma multiforme	Frozen tissue	CIEF and nano- or capillary-RPLC[a]	[183]
Hepatocellular carcinoma	Tissue	CZE and inductively coupled plasma MS[b]	[184]
Alzheimer's disease	CSF	CZE–ESI micro-TOF[c]	[179]
Schizophrenia	CSF	CZE–ESI micro-TOF[c]	[92]
Alzheimer's disease	Blood	ESI-HP^{3D} CE[d]	[185]
Sepsis	Blood	CZE–ESI-TOF[e]	[186]
End stage renal disease	Blood	CZE–ESI-TOF[f]	[91,149]
Renal diseases	Urine	CZE–ESI-TOF	[148,151[g],152[e]]
Renal transplantation	Urine	CZE–ESI-TOF[g]	[163]
Graft *versus* host diseases	Urine	CZE–ESI-TOF	[164[h],165[g]]
Uteropelvic junction obstruction	Urine	CZE–ESI-TOF[f]	[153]
Diabetic nephropathy	Urine	CZE–ESI-TOF	[148[f],155[i],156[f],157[e],167[f],186[e]]
Inborn errors of amino acids metabolism	Urine	CZE–ESI-TOF[j]	[187]
Obstructive pulmonary disease	Urine	CE–LIF-ESI-IT[k]	[188]
Cystinuria	Urine	CE–LIF[l]	[189]
Bladder cancer	Urine	CZE–ESI-MS/MS[m]	[158]
Prostate cancer	Urine	CZE–ESI micro-TOF[c]	[107,159]

Modified from references [12,182].

[a] Two orthogonal multidimensional techniques used: CIEF for prefractionation followed by nano-RPLC or capillary-RPLC for separation using C_{18}-RP-LC column, interfaced with ESI ion source to LCQ IT MS (ThermoFinnigan, San Hose, CA).

[b] CZE used for analysis of metallothionine isoforms and inductively coupled plasma-mass spectrometry for trace metals of copper and zinc in the livers of patients afflicted with hepatocellular carcinoma.

[c] CZE by P/ACE MDQ CE apparatus (Beckman-Coulter), ESI sprayer kit (Agilent Technologies, Palo Alto, CA, USA) and micro-TOF (Bruker Daltonik, Bremen, Germany).

[d] ESI sheath-liquid interface with orthogonal coaxial triple tube ESI interface (Agilent Technologies) to a capillary electrophoresis instrument HP^{3D} CE (Agilent Technologies, Waldbronn, Germany).

[e] P/ACE MDQ CE, ESI sheath sprayer (Agilent Technologies, Waldbronn, Germany) and TOF (Mariner Biospectrometry Workstation (Aplied Biosystems, Foster City, CA, USA).

[f] CZE by P/ACE MDQ coupled to ESI-TOF (Perceptive Biosystems, Farmingham, MA, USA).

[g] CZE by P/ACE MDQ coupled to ESI sheath flow sprayer kit (Agilent) to a micro-TOF (Bruker Daltonik).

[h] CZE by P/ACE MDQ coupled to ESI-TOF Mariner Biospectrometry Workstation (Preceptive Biosystems, Farmingham, MA, USA).

[i] CZE by P/ACE MDQ coupled to ESI sheath flow sprayer kit (Applied Biosystems) to a TOF (ABI Mariner).

[j] CZE by Agilent Capillary Electrophoresis HP^{3D} CE (Agilent Technologies, Waldbron, Germany), G1607 Agilent ESI sprayer kit connected to API 300 TQ MS/MS (Applied Biosystems, Foster City, CA, USA).

[k] CE–LIF Beckman P/ACE 2200 equipped with a laser Model 488 (Beckman) interfaced with ESI ion source to LCQ DECA IT MS (ThermoFinnigan).

[l] P/ACE 5510 with argon-ion laser-induced fluorescence (LIF) detector 488/522 nm (Beckman).

[m] CZE by P/ACE MDQ coupled to sheath flow sprayer kit from Agilent, and MS/MS by Ultraflex MALDI TOF–TOF from Bruker Daltonik.

Because biomarkers are excreted at high concentration in the vicinity of their site of production and are considerably diluted after they are excreted into various body fluids (often after they have traveled long distance from the original source of their production) [190], the capabilities of CIEF-based multidimensional separations for performing proteome analysis from minute tissue samples create new opportunities in the pursuit of biomarker discovery using enriched cell populations produced from tissue specimens [128]. Immunoaffinity capillary electrophoresis (IA-CE) is emerging as a promising HT method for the analysis of low-abundance biomarkers because (a) bioselective adsorption, (b) subsequent recovery of markers from an immobilized affinity ligand, and (c) finally separation of enriched compound [191].

Instead of using complicated proteomic plasma samples for initial biomarker discovery, it was suggested that success could be enhanced by first comparing normal tissue from which the biomarker is believed to emanate with diseases ones, or diseased tissue with two or several well separated etiologies to find differences in proteins detected, and then it would be easier to look for these differences by focused approaches such as MS or immunoassays in body fluids [144].

CE–MS is a powerful tool for elucidating the pathophysiological relevance of biomarkers because it combines excellent performance for biomarker discovery with a unique separation platform-independent biomarker sequence, which contributes to a better understanding of the disease under consideration [108].

Conflict of interest

The author declares no financial interests or conflict related to the contents of this article.

Acknowledgements

I express my gratitude to the many colleagues cited in the reference section for kindly providing many materials and for clarification and interesting discussions throughout preparing this manuscript, the anonymous reviewers and the assigned editor of this manuscript for making constructive comments that enhanced the readability of this review.

References

[1] A. Tiselius, Nova Acta Regiae Socieatarium Upsaliensis 3 (1930) 7.
[2] A. Tiselius, Biochem. J. 31 (1937) 313.
[3] S. Hjertén, Chromatogr. Rev. 9 (1967) 122.
[4] R. Virtanen, Acta Polytech. Scand. 51 (1974) 123.
[5] V. Neuhoff, W.B. Schill, H. Sternbach, Biochem. J. 117 (1970) 623.
[6] F.E.P. Mikkers, F.M. Everaerts, T.P.E.M. Verheggen, J. Chromatogr. 169 (1979) 12.
[7] J.W. Jorgenson, K.D. Luckacs, Science 222 (1983) 266.
[8] S. Terabe, K. Otsuka, K. Ichikawa, A. Tsuchuya, T. Ando, Anal. Chem. 56 (1984) 111.
[9] S. Hjertén, K. Elenbring, F. Kilar, J.L. Lao, A.J. Chen, C.J. Siebert, M.-D. Zhu, J. Chromatogr. 403 (1987) 47.
[10] A.S. Cohen, B.L. Karger, J. Chromatogr. 397 (1987) 409.
[11] S. Hjertén, Z. Liu, S.L. Zhao, J. Biochem. Biophys. Methods 7 (1983) 101.
[12] M. Sniehotta, E. Schiffer, P. Zürbig, J. Novak, H. Mischak, Electrophoresis 28 (2007) 1407.
[13] X. Bossuyt, Advan. Clin. Chem. 42 (2006) 43.
[14] W. Kolch, C. Neusüss, M. Pelzing, M. Mischak, Mass Spectrom. Rev. 24 (2005) 959.
[15] M.L. Riekkola, J.A. Joenssorg, Pure Appl. Chem. 76 (2004) 443.
[16] D.C. Simpson, R.D. Smith, Electrophoresis 26 (2005) 1291.
[17] G. Choudhary, A. Apffel, H. Yin, W. Hancock, J. Chromatogr. A 887 (2000) 85.
[18] Y. Huo, W.T. Kok, Electrophoresis 29 (2008) 80.
[19] A.E. Herr, J.I. Molho, K.A. Drouvalakis, J.C. Mikkelsen, P.J. Utz, J.G. Santiago, T.W. Kenny, Anal. Chem. 75 (2003) 1180.
[20] W. Tang, A.K. Harrata, C.S. Lee, Anal. Chem. 69 (1997) 3177.

[21] L. Yang, C.S. Lee, S.A. Hofstadler, L. Pasa-Tolic, R.D. Smith, Anal. Chem. 70 (1998) 3235.
[22] M.H. Lamoree, U.R. Tjaden, J. Van der Greef, J. Chromatogr. A 777 (1997) 31.
[23] T. Hirokawa, H. Okamoto, B. Gaš, Electrophoresis 24 (2003) 496.
[24] A. Harsch, P. Vouros, Anal. Chem. 70 (1998) 3021.
[25] T. Freudemann, A. Von Brocke, E. Bayer, Anal. Chem. 73 (2001) 2587.
[26] C. Galbusera, D.D. Chen, Curr. Opin. Biotechnol. 14 (2003) 126.
[27] T. Le Saux, A. Varenne, P. Gareil, Electrophoresis 26 (2005) 3094.
[28] T.J. Pappas, M. Gayton-Ely, L.A. Holland, Electrophoresis 26 (2005) 719.
[29] J. Hernández-Borges, C. Neusüβ, A. Cifuentes, M. Pelzing, Electrophoresis 25 (2004) 2257.
[30] P. Petersson, M. Jörntén-Karlsson, M. Stålebro, Electrophoresis 24 (2003) 999.
[31] R. Jabeen, D. Payne, J. Wiktorowicz, A. Mohammad, J. Petersen, Electrophoresis 27 (2006) 2413.
[32] F.E. Ahmed, Exp. Rev. Proteomics 4 (2008) 841.
[33] V. Klyishnichenko, A. Rodenbrock, J. Thőmmes, M.-R. Kula, H. Heine, M. Biselli, Biotechnol. Appl. Biochem. 27 (1988) 181.
[34] S.S. Rubakhin, J.S. Page, B.R. Monroe, J.V. Sweedier, Electrophoresis 22 (2001) 3752.
[35] I.J. Preisler, P. Hu, B.L. Karger, Anal. Chem. 72 (2000) 4785.
[36] J.S. Page, S.S. Rubakhin, J.V. Sweedler, Analyst 125 (2000) 555.
[37] S.I. Snovida, V.C. Chen, O. Krokhin, H. Perreault, Anal. Chem. 78 (2006) 655.
[38] H. Stutz, Electrophoresis 26 (2005) 1254.
[39] K.K.C. Yeung, A.G. Kiceniuk, L. Li, J. Chromatogr. A 931 (2001) 153.
[40] B.M. Balgley, W. Wang, T. Song, X. Fang, L. Yang, C.S. Lee, Electrophoresis 29 (2008) 3047.
[41] C. Neusüβ, M. Pelzing, M. Macht, Electrophoresis 23 (2002) 3149.
[42] J.F. Banks, Electrophoresis 18 (1997) 2255.
[43] P. Schmitt-Koplin, M. Frommberger, Electrophoresis 24 (2003) 3837.
[44] R.D. Smith, J.A. Olivares, N.T. Nguyen, H.R. Udseth, Anal. Chem. 60 (1998) 436.
[45] F. Foret, D.P. Kirby, P. Vouros, B.L. Karger, Electrophoresis 17 (1996) 1829.
[46] T. Kaiser, S. Wittke, I. Just, R. Krebs, S. Bartel, et al., Electrophoresis 25 (2004) 2044.
[47] J.A. Olivares, N.T. Nguyen, C.R. Yonker, D. Smith, Anal. Chem. 59 (1987) 1230.
[48] J.F. Kelly, L. Ramaley, P. Thibautt, Anal. Chem. 69 (1997) 51.
[49] S.A. Hofstadtler, J.C. Severs, R.D. Smith, F.D. Swanek, A.G. Ewing, Rapid Commun. Mass Spectrom. 10 (1996) 919.
[50] S. Nilsson, M. Wetterhall, J. Bergquist, L. Nyholm, K.E. Markides, Rapid Commun. Mass Spectrom. 15 (2001) 1997.
[51] V. Sanz-Nebut, E. Balaguer, F. Benavente, J. Barbosa, Electrophoresis 26 (2005) 1457.
[52] E.D. Lee, W. Mück, J.D. Henion, T.R. Covey, Biomed. Environ. Mass Spectrom. 18 (1989) 844.
[53] J.C. Severs, A.C. Harms, R.D. Smith, Rapid Commun. Mass Spectrom. 10 (1996) 1179.
[54] N.A. Guzman, R.J. Stubbs, Electrophoresis 22 (2001) 3602.
[55] A. Yaming, J.W. Cooper, B.M. Balgley, C.S. Lee, Electrophoresis 27 (2006) 3599.
[56] W.L. Tseng, H.T. Chang, J. Chromatogr. A 924 (2001) 93.
[57] C.W. Huck, G.K. Bonn, Methods Mol. Biol. 384 (2008) 507.
[58] M. Larsson, E.S.M. Lutz, Electrophoresis 21 (2002) 2859.
[59] W. Tong, A. Link, J.K. Eng, J.R. Yates III, Anal. Chem. 71 (1999) 2270.
[60] G.M. Janini, M. Zhou, L.R. Yu, J. Blonder, G. Michelle, T.P. Conrads, H.J. Issaq, T.D. Veenstra, Anal. Chem. 75 (2003) 5984.
[61] R.E. Settlage, P.S. Russo, J. Shabanowitz, D.F. Hunt, J. Microbiol. 10 (1998) 281.
[62] M.B. Barroso, A.P. deJong, J. Am. Soc. Mass Spectrom. 10 (1999) 1271.
[63] J.F. Kelly, S.J. Locke, L. Ramaley, P. Thibault, J. Chromatogr. A 720 (1996) 409.
[64] M.X. Li, L. Liu, J.T. Wu, D.M. Lubman, Anal. Chem. 69 (1997) 2451.
[65] A. Nguyen, M. Moini, Anal. Chem. 80 (2008) 7169.
[66] J.R. Catai, J. Sastre-Torano, P.M. Jongen, G.J. de Jong, G.W. Somsen, J. Chromatogr. B 852 (2007) 160.
[67] R. Ramautar, O.A. Mavboroda, R.J. Derks, C. van Nieuwkoop, J.T. Dissel, G.W. Somsen, A.M. Deelder, G.J. de Jong, Electrophoresis 29 (2008) 2714.
[68] R. Ramautar, O.A. Mavboroda, A.M. Deelder, G.W. Somsen, G.J. de Jong, J. Chromatogr. B 871 (2008) 370.
[69] N. Johannesson, M. Wetterhall, K.E. Markides, J. Bergquist, Electrophoresis 25 (2004) 809.
[70] C. Simó, C. Elvira, N. Gonzalez, J. San Román, C. Barbas, A. Cifuentes, Electrophoresis 25 (2004) 2056.
[71] S. Ullsten, A. Zuberovic, M. Wetterhall, B. Hardenborg, K.E. Markides, J. Bergquist, Electrophoresis 25 (2004) 2090.
[72] S.W. Tsai, M. Loughran, A. Hiratsuka, K. Yano, I. Karube, Analyst 128 (2003) 237.
[73] J. Gaudioso, H.G. Craighead, J. Chromatogr. A 971 (2002) 249.
[74] F.M. Fernández, J.M. Vadillo, J.R. Kimmel, M. Wetterhall, K. Markides, N. Rodriguez, R.N. Zare, Anal. Chem. 74 (2002) 1611.
[75] S. Martinovic, T.D. Veenstra, G.A. Anderson, L. Pasa-Tolic, R.D. Smith, J. Mass Spectrom. 37 (2002) 99.
[76] R.D. Smith, J.H. Wahl, K.J. Light-Wahl, B.E. Winger, J. Tox, Environ. Health 40C (1993) 147.
[77] P. Zürbig, M.B. Renfrow, E. Schiffer, J. Novak, M. Walden, S. Wittke, I. Just, M. Pelzing, C. Neusüβ, D. Theodorescu, K.E. Root, M.M. Ross, H. Mischak, Electrophoresis 27 (2006) 2111.
[78] M.J. Chalmers, C.L. Mackay, C.L. Hendrickson, S. Wittke, M. Walden, H. Mischak, D. Fliser, I. Just, A.G. Marshall, Anal. Chem. 77 (2005) 7163.
[79] K.J. Wu, T.A. Shaler, C.H. Becker, Anal. Chem. 66 (1994) 1637.
[80] D.C. Muddiman, A.I. Gusev, A. Proctor, D.M. Hercules, R. Venkataramanan, W. Dinen, Anal. Chem. 66 (1994) 2362.
[81] F.E. Ahmed, J. Sep. Sci. 32 (2009) 771.
[82] A. Dawoud, H.A. Sarvaiya, I.M. Lazar, Electrophoresis 28 (2007) 4645.
[83] F. Kitagawa, K. Otsuka, J. Sep. Sci. 31 (2008) 794.
[84] D. Figeys, A. Ducret, J.R. Yates 5rd, R. Aebersold, Nat. Biotechnol. 14 (1996) 1579.
[85] I.M. Lazar, J. Grym, F. Forest, Mass Spectrom. Rev. 25 (2006) 573.
[86] Y. Zheng, H. Li, Z. Guo, J.M. Lin, Z. Cai, Electrophoresis 28 (2007) 1305.
[87] J. Preisler, P. Hu, T. Rejtar, E. Moskovets, B.L. Karger, Anal. Chem. 74 (2002) 17.
[88] F.E. Ahmed, Exp. Rev. Proteomics 6 (2009) 171.
[89] P.M. Palagi, D. Walther, M. Quadroni, C. Sébastien, J. Burgess, R.D. Appel, Proteomics 5 (2005) 2381.
[90] S. Wittke, D. Fliser, M. Haubitz, S. Bartel, R. Krebs, F. Hausadel, M. Hillmann, I. Golovko, P. Koester, H. Haller, T. Kaiser, H. Mischak, E.M. Weissinger, J. Chromatogr. A 1013 (2003) 173.
[91] T. Kaiser, A. Hermann, J.T. Kielstein, S. Wittke, S. Bartel, R. Krebs, F. Hausadel, M. Hillmann, I. Golovko, P. Koester, H. Haller, E.M. Weissinger, D. Fliser, H. Mischak, J. Chromatogr. A 1013 (2003) 157.
[92] S. Wittke, H. Mischak, M. Walden, W. Kolch, T. Rädler, K. Wiedemann, Electrophoresis 26 (2005) 1476.
[93] F.E. Ahmed, Exp. Rev. Mol. Diag. 6 (2006) 736.
[94] E.T. Fung, S.R. Weinberger, E. Gavin, F. Zhang, Exp. Rev. Proteomics 2 (2005) 847.
[95] A. Haoudi, H. Bensmail, Exp. Rev. Proteomics 3 (2006) 333.
[96] N.J. Camp, M.L. Slattery, Cancer Causes Control 13 (2002) 813.
[97] E.M. Weissinger, S. Wittke, T. Kaiser, H. Haller, S. Bartel, R. Krebs, I. Golovko, H.D. Rupprecht, M. Haubitz, H. Hecker, H. Mischak, D. Fliser, Kidney Int. 65 (2004) 2426.
[98] Z.R. Yang, K.C. Chou, Bioinformatics 20 (2004) 735.
[99] F.E. Ahmed, Mol. Cancer 4 (2005) 29.
[100] Q.C. Ru, L.A. Zhu, J. Silberman, C.D. Shriver, Mol. Cell. Proteomics 5 (2006) 1095.
[101] T. Gobel, S. Vorderwulbecke, K. Hauck, H. Fey, D. Haussinger, A. Erdhardt, World J. Gastroenterol. 12 (2006) 7604.
[102] M. Girolami, S. Rogers, Neural Comput. 18 (2006) 1790.
[103] D.F. Ransohoff, Nat. Rev. Cancer 4 (2004) 309.
[104] F. Dieterle, S. Müller-Hagedorn, H.M. Liebich, G. Gauglitz, Artif. Intell. Med. 28 (2003) 265.
[105] M.P. Brown, W.N. Grundy, D. Lin, N. Cristianini, C.W. Sugnet, T.S. Funey, M. Ames, D. Haussler, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 262.
[106] R.E. Offord, Nature 211 (1966) 591.
[107] D. Theodorescu, D. Fliser, S. Wittke, H. Mischak, R. Krebs, M. Walden, M. Ross, E. Eltze, O. Bettendorf, C. Wulfing, A. Semjonow, Electrophoresis 26 (2005) 2792.
[108] E. Schiffer, H. Mischak, J. Novak, Proteomics 6 (2006) 5615.
[109] P.H. O'Farrell, J. Biol. Chem. 250 (1975) 4007.
[110] F.E. Ahmed, Exp. Rev. Proteomics 5 (2008) 469.
[111] N. Tang, P. Tornatore, S.R. Weinberger, Mass Spectrom. Rev. 23 (2004) 34.
[112] D.E. van der Mere, K. Oikonomopoulou, J. Marshall, E.P. Diamandis, Adv. Cancer Res. 96 (2007) 23.
[113] G. Chen, B.N. Pramanik, Exp. Rev. Proteomics 5 (2008) 435.
[114] E.I. Chen, J. Hewel, B. Flding-Habermann, J.R. Yates III, Mol. Cell. Proteomics 5 (2006) 53.
[115] N.J. Adamson, E.C. Reynolds, J. Chromatogr. B 699 (1997) 133.
[116] V.J. Katz, Math. Mag. 52 (1979) 146.
[117] J. Kim, R. Zand, D.M. Lubman, Electrophoresis 23 (2002) 782.
[118] Y.K. Shin, H.J. Lee, J.S. Lee, Y.K. Paik, Proteomics 6 (2006) 1143.
[119] H.J. Lee, E.Y. Lee, M.S. Kwon, Y.K. Paik, Curr. Opin. Chem. Biol. 10 (2006) 42.
[120] F.E. Ahmed, Curr. Genomics 7 (2006) 399.
[121] H. Mischak, R. Apweiler, R.E. Banks, M. Conaway, J. Coon, A. Dominiczak, J.H.H. Ehrich, D. Fliser, et al., Proteomics Clin. Appl. 1 (2007) 148.
[122] H. Mischak, B.A. Julian, J. Novak, Proteomics Clin. Appl. 1 (2007) 792.
[123] T. Stasyk, L.A. Huber, Proteomics 4 (2004) 3704.
[124] P.G. Righetti, A. Castagna, P. Antonioli, E. Bochetti, Electrophoresis 26 (2005) 297.
[125] F.E. Ahmed, Cancer Genom. Proteom. 3 (2006) 217.
[126] P. Chaurand, M.E. Sanders, R.A. Jensen, R.M. Caprioli, Am. J. Pathol. 165 (2004) 1057.
[127] J. Chen, B.M. Balgley, D.L. DeVoe, C.S. Lee, Anal. Chem. 75 (2003) 3145.
[128] T. Guo, C.S. Lee, W. Wang, D.L. DeVoe, B.M. Balgley, Electrophoresis 27 (2006) 3523.
[129] S.R. Shi, C.S. Liu, B.M. Balgley, C.S. Lee, C.R. Taylor, J. Histochem. Cytochem. 54 (2006) 739.
[130] S.I. Hwang, J. Thumar, D.H. Lundgren, K. Rezaul, et al., Oncogene 26 (2007) 65.
[131] W. Wang, T. Guo, P.A. Rudnick, T. Song, J. Li, Z. Zhuang, W. Zheng, D.L. DeVoe, C.S. Lee, B.M. Balgley, Anal. Chem. 79 (2007) 1002.
[132] D. Theodorescu, H. Mischak, World J. Urol. 25 (2007) 435.
[133] G.S. Omenn, D.J. States, M. Adamski, T.W. Blackwell, R. Menon, H. Hermjakob, R. Apweiler, B.B. Haab, R.J. Simpson, J.S. Eddes, E.A. Kapp, R.L. Moritz, D.W. Chan, A.J. Rai, A. Admon, R. Aebersold, J. Eng, W.S. Hancock, S.A. Hefta, H. Meyer, Y.K. Paik, J.S. Yoo, P. Ping, J. Pounds, J. Adkins, X. Qian, R. Wang, V. Wasinger, C.Y. Wu, X. Zhao, R. Zeng, A. Archakov, A. Tsugita, I. Beer, A. Pandey, M. Pisano, P. Andrews, H. Tammen, D.W. Speicher, S.M. Hanash, Proteomics 5 (2005) 3226.
[134] S. Schaub, J. Wilkins, T. Weiler, K. Sangster, D. Rush, P. Nickers, Kidney Int. 65 (2004) 323.

[135] D. Theodorescu, S. Wittke, M.M. Ross, M. Walden, M. Conway, I. Just, H. Mischak, H.F. Frierson, Lancet Oncol. 7 (2006) 230.
[136] H. Zhou, P.S. Yuen, T. Pisitkun, P.A. Gonzales, H. Yasuda, J.W. Dear, P. Gross, M.A. Knepper, R.A. Star, Kidney Int. 69 (2006) 1471.
[137] P. Vestergaard, R. Leverett, J. Lab. Clin. Med. 51 (1958) 211.
[138] M. Kussmänn, F. Raymond, M. Affolter, J. Biotechnol. 124 (2006) 758.
[139] H.F. Cserr, D.N. Cooper, P.K. Suri, C.S. Patlak, Am. J. Physiol. 240 (1981) F319.
[140] D.M. Good, V. Thongboonkerd, J. Novak, J.L. Bascands, J.P. Schanstra, J.J. Coon, A. Dominiczak, H. Mischak, J. Proteome Res. 6 (2007) 4549.
[141] D. Fliser, J. Novak, V. Thongboonkerd, À. Argilés, V. Jankowski, M.A. Girolami, J. Jankowski, H. Mischak, J. Am. Soc. Nephrol. 18 (2007) 1057.
[142] V. Faca, S.J. Pitteri, L. Newcomb, V. Glukhova, D. Phanstiel, A. Krasnoselsky, Q. Zhang, J. Strthers, H. Wang, J. Eng, M. Fitzgibbon, M. McIntosh, S.J. Hanash, J. Proteome Res. 6 (2007) 3558.
[143] J. Villanueva, J. Philip, D. Entenberg, C.A. Chaparro, M.K. Tanwar, E.C. Holland, P. Tempst, Anal. Chem. 76 (2004) 1560.
[144] D. Lescuyer, T. Hochstrasser, Rabilloud, J. Proteome Res. 6 (2007) 3371.
[145] E. Thouvenot, M. Lafon-Cazal, E. Demettre, P. Jouin, J. Bockaert, P. Marin, Proteomics 6 (2006) 5941.
[146] M. Chevallet, H. Diemer, A. van Dorsselaer, C. Villers, T. Rabilloud, Proteomics 7 (2007) 1757.
[147] J. Adachi, C. Kuman, Y. Zhang, J.V. Olsen, M. Mann, Genome Biol. 7 (2006) R80.
[148] M. Haubitz, S. Wittke, E.M. Weissinger, M. Walden, H.D. Ruppecht, J. Floege, H. Haller, H. Mischak, Kidney Int. 67 (2005) 2313.
[149] E.M. Weissinger, T. Kaiser, N. Meert, R. DeSmet, M. Walden, H. Mischak, R.C. Vanholder, Nephrol. Dial. Transplant. 19 (2004) 3068.
[150] E.M. Weissinger, T. Nguyen-Khoa, C. Fumeron, C. Saltiel, M. Walden, T. Kaiser, H. Mischak, T.B. Drüeke, B. Lacour, Z.Z. Massy, Proteomics 6 (2006) 993.
[151] B.A. Julian, S. Wittke, J. Novak, D.M. Good, J.J. Coon, P. Zürbig, E. Schiffer, M. Haubitz, Z. Moldoveanu, S.M. Calcatera, R.J. Wyatt, J. Sýkora, E. Sládková, O. Hes, H. Mischak, M.B. McGuire, Electrophoresis 27 (2007) 4469.
[152] B.A. Julian, S. Wittke, M. Haubitz, P. Zürbig, E. Schiffer, M.B. McGuire, R.J. Wyatt, J. Novak, World J. Urol. 25 (2007) 467.
[153] S. Decramer, J.L. Bascands, J.P. Schanstra, World J. Urol. 25 (2007) 457.
[154] K. Rossing, H. Mischak, M. Dakna, P. Zürbig, J. Novak, B.A. Julian, D.M. Good, J.J. Coon, L. Tarnow, J. Am. Soc. Nephrol. 19 (2008) 1283.
[155] H. Mischak, T. Kaiser, M. Walden, M. Hillmann, S. Krueppel, H. Haller, D. Fliser, Clin. Sci. 107 (2004) 485.
[156] M. Meier, T. Kaiser, A. Herrmann, S. Knueppel, M. Hillmann, P. Koester, T. Danne, H. Heller, D. Fliser, Diabetes Complications 19 (2005) 223.
[157] K. Rossing, H. Mischak, H.H. Parving, P.K. Christensen, M. Walden, T. Kaiser, Kidney Int. 68 (2005) 193.
[158] D. Theodorescu, S. Wittke, M.M. Ross, M. Walden, M. Conway, I. Just, H. Miscak, H.F. Frrerson, Lancet Oncol. 7 (230) (2006).
[159] D. Theodorescu, E. Schiffer, H.W. Baure, F. Douwes, F. Eichhorn, R. Polley, T. Schmidt, W. Schőfer, P. Zürbig, D.M. Good, J.J. Cohn, H. Miscak, Proteomics Clin. Appl. 2 (2008) 556.
[160] N.A. Obuchoeski, M.L. Lieber, F.H. Wians Jr., Clin. Chem. 50 (2004) 1118.
[161] Y. Benjamini, Y. Hochberg, J. R. Stat. Soc. B 57 (1995) 125.
[162] M.K. Kerr, G.A. Churchill, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 8961.
[163] S. Wittke, M. Haubitz, M. Walden, F. Rhode, A. Schwarz, M. Mengel, H. Michak, H. Haller, W. Gwinner, Am. J. Transplant. 5 (2005) 2479.
[164] E.M. Weissinger, E. Schiffer, B. Hertenstein, J.L. Ferrara, E. Holler, M. Stadler, H.J. Kolb, A. Zander, P. Zürbig, M. Kellmann, A. Ganser, Blood 109 (2007) 5511.
[165] E.M. Weissinger, B. Hertenstein, H. Mischak, A. Ganser, Exp. Rev. Proteomics 2 (2005) 639.
[166] O.V. Nemirovskiy, D.R. Dufield, T. Sunyer, P. Agrawal, D.J. Welsch, W.R. Mathews, Anal. Biochem. 361 (2007) 93.
[167] S. Decramer, S. Wittke, H. Mischak, P. Zürbig, M. Walden, F. Bouissou, J.L. Bascands, J.P. Schanstra, Nat. Med. 12 (2006) 398.
[168] N. Johannesson, L. Olsson, L. Bä kström, M. Wetherhall, R. Danielsson, J. Bergquist, Electrophoresis 28 (2007) 1435.
[169] G.A. Valaskovic, N.L. Kellehr, F.W. McLafferty, Science 273 (1996) 1199.
[170] P. Cao, M. Moini, J. Am. Soc. Mass Spectrom. 10 (1999) 184.
[171] M. Moini, S.M. Demars, H. Huang, Anal. Chem. 74 (2002) 3772.
[172] T. He, N. Chandramouli, E. Fu, A. Wu, Y.K. Wang, Anal. Biochem. 171 (1999) 189.
[173] L.J. Deterding, J.M. Cutalo, M. Khaledi, K.B. Tomer, Electrophoresis 23 (2002) 2296.
[174] V. Sanz-Nebot, F. Benavente, A. Vallverdu, N.A. Guzman, A. Barbosa, Anal. Chem. 75 (2003) 5220.
[175] J. Kraly, M.A. Fazal, R.M. Schoenherr, R. Bonn, M.M. Hardwood, E. Turner, M. Jones, N.J. Dovichi, Anal. Chem. 78 (2006) 5124.
[176] M. Wetterhall, M. Palmblad, P. Hakansson, K.E. Markides, J. Bergquist, J. Proteome Res. 1 (2002) 361.
[177] N.A. Guzman, Anal. Bioanal. Chem. 378 (2004) 37.
[178] T.M. Phillips, Electrophoresis 25 (2004) 1652.
[179] S. Wittke, H. Mischak, M. Walden, W. Kolch, J. Rädler, K. Wiedemann, Electrophoresis 26 (2005) 1576.
[180] D. Fliser, S. Wittke, H. Mischak, Electrophoresis 26 (2005) 2708.
[181] J.R. Petersen, A.O. Okorodudu, A. Mohammad, D.A. Payne, Clin. Chim. Acta 330 (2003) 1.
[182] R. Bakry, C.W. Huck, M. Najm-ul-Haq, M. Rainer, G.K. Bonn, J. Sep. Sci. 30 (2007) 192.
[183] Y. Wang, P.A. Rudnick, E.L. Evans, J. Li, Z. Zhuang, D.L. Devoe, C.S. Lee, B.M. Balgley, Anal. Chem. 77 (2005) 6549.
[184] T. Kawata, S. Nakamura, A. Nakayama, H. Fukuda, M. Ebara, T. Nagamine, T. Minami, H. Sakurai, Biol. Pharm. Bull. 29 (2006) 403.
[185] E. Varesio, S. Rudaz, K.H. Krause, J.L. Veuthey, J. Chromatogr. A 974 (2002) 135.
[186] T. Kaiser, S. Wittke, I. Just, R. Krebs, S. Bartel, D. Fliser, H. Mischak, E.M. Weissinger, Electrophoresis 25 (2004) 2044.
[187] T. Soga, Y. Kakazu, M. Robert, T.M. Martin, T. Nishioka, Electrophoresis 25 (2004) 1964.
[188] L. Annovazzi, S. Viglio, E. Perani, M. Luisetti, et al., Electrophoresis 25 (2004) 683.
[189] P. Lochman, T. Adam, D. Friedecky, E. Hidkova, Z. Skopkova, Electrophoresis 24 (2003) 1200.
[190] F.E. Ahmed, Curr. Proteomics 5 (2008) 224.
[191] N.A. Guzman, Am. Lab. 40 (2008) 18.