SPECIAL ISSUE PAPER

Detlev Suckau · Anja Resemann · Martin Schuerenberg
Peter Hufnagel · Jochen Franzen · Armin Holle

A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics

Abstract A new matrix-assisted laser-desorption/ionization time-of-flight/time-of-flight mass spectrometer with the novel "LIFT" technique (MALDI LIFT-TOF/TOF MS) is described. This instrument provides high sensitivity (attomole range) for peptide mass fingerprints (PMF). It is also possible to analyze fragment ions generated by any one of three different modes of dissociation: laser-induced dissociation (LID) and high-energy collision-induced dissociation (CID) as real MS/MS techniques and in-source decay in the reflector mode of the mass analyzer (reISD) as a pseudo-MS/MS technique. Fully automated operation including spot picking from 2D gels, in-gel digestion, sample preparation on MALDI plates with hydrophilic/hydrophobic spot profiles and spectrum acquisition/processing lead to an identification rate of 66% after the PMF was obtained. The workflow control software subsequently triggered automated acquisition of multiple MS/MS spectra. This information, combined with the PMF increased the identification rate to 77%, thus providing data that allowed protein modifications and sequence errors in the protein sequence database to be detected. The quality of the MS/MS data allowed for automated de novo sequencing and protein identification based on homology searching.

Keywords TOF/TOF · MS/MS · Tryptophan oxidation · de novo Sequencing · Automation · MALDI

Introduction

The new era of post-genomics calls for techniques that further the understanding of gene function by extracting genomic information and even revealing information that is not directly coded in the DNA. The main routes to analyze whole living systems typically rely on the detection of changes in mRNA and in protein expression levels [1]. Some of the most striking conclusions made recently indicate the number of expressed proteins (>100,000) by far exceeds the number of genes (ca. 30,000) [2, 3]. Even the expression levels of proteins only poorly correlate with the transcription of the respective mRNA [4]. To further complicate matters, 2D gels reveal that many proteins are present in the cell in different forms, i.e. they can be identified from more than 20 spots in a gel. This protein diversity results from the presence of protein variants including splice-variants and proteolytic processing such as protein degradation plus heterogeneous post-translational modifications, such as phosphorylation and glycosylation. It has become evident that understanding function in living systems requires the analysis of protein expression levels, their distribution in tissues and cellular compartments, their structural details and molecular interactions that eventually turn the genetic information into function. These requirements indicate the need for fast methods to characterize low levels of a wide dynamic range of proteins in detail, since the proteins with the highest relevance in understanding biological processes are often expressed at lowest levels.

The most widely adopted strategy for characterizing expression levels and identifying the corresponding proteins uses two-dimensional gel electrophoresis (2DGE) and mass spectrometry (MS). The importance of mass spectrometry in this context was even recognized in the 2002 Nobel prize for chemistry [5]. The protein is most commonly digested in-gel by trypsin. Proteolytic peptides are eluted and prepared on a MALDI sample plate, since matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) provides a good high throughput compatibility [1, 6] and a good sensitivity in the low femtomole range and even below [7].

The peptide mass fingerprint (PMF) is used to screen a protein sequence database. A match between the PMF and the patterns computed from the sequence database, employing rules for digest chemistry, are scored and used for

D. Suckau (✉) · A. Resemann · M. Schuerenberg · P. Hufnagel
J. Franzen · A. Holle
Bruker Daltonik GmbH, Bremen, Germany
e-mail: dsu@bdal.de

protein identification. The success of such an identification depends on:

1. the size of the database and the error rate within,
2. the number of the matching peptides and their molecular weight (MW),
3. the mass accuracy, and
4. the control over the digest chemistry.

Given a database, the success rate of protein identification increases with the number and MW of the peptides and decreases with increasing mass tolerance and chemical heterogeneity. The types of chemical problem typically rendering identification more difficult include incomplete proteolytic degradation and unspecific cleavages, preparation artifacts such as methionine oxidation, deamidation, and cystein acrylamidation. The digestion chemistry and MALDI preparation, and instrument sensitivity and mass accuracy are therefore crucial for high levels of success in PMF-based protein identification.

In certain cases PMFs provide insufficient information for protein identification. MS/MS fragment ion analysis of selected peptides can be used for improved protein identification. Problems affecting protein identification from PMFs include:

1. Low number of proteolytic peptides in the standard mass range of ca. 800–3500 Da due to low protein MW (ca. <15 kDa), unfortunate cleavage site distribution or peptide suppression. Low amounts of protein or high levels of contaminant (detergents, etc.) also produce weak PMF spectra, which will lead to low identification efficiency.
2. Expressed sequence tag (EST) or even genomic databases have error rates of 0.3% and higher[8, 9], i.e. up to 1 in 30 amino acids in the translated protein sequence are wrong or sequence errors that will cause frame shifts. Matching a peptide that spans the splice site between two exons in genomic data is not possible either. Such effects frequently prevent sufficient numbers of peptides in PMFs from being matched[10]. The "normal" error levels in the common sequence databases such as NCBI (http://www.ncbi.nlm.nih.gov) can also cause incorrect PMF identifications.
3. All protein identifications based on mass spectra that could represent a set of proteolytic peptides from several proteins cannot be based solely on PMFs. Such experiments include one-dimensional chromatography or multidimensional chromatography of complex protein mixtures such as proteolytic digests of full proteomes. Due to the extreme complexity of such samples, which is almost incompatible with simultaneous protein quantitation and identification, methods are currently being developed to simplify the nature of the experiments [11]. Isotope-coded affinity tags[12, 13] and diagonal chromatography [14] are the most prominent examples of such approaches. Here, electrospray ionization mass spectrometry (ESI MS) is typically employed on either hybrid quadrupole TOF (Q-TOF) or ion-trap mass spectrometers in conjunction with LC separation. The use of MALDI offers a significant advantage over LC–ESI

MS/MS in that the chromatogram can be "recorded" onto the MALDI target. The MS/MS analysis is therefore decoupled from the temporal constraints of chromatographic separation and various regimes of data analysis-dependent acquisition control become possible [11].

The first mass spectrometric step in an MS/MS experiment is to select a precursor ion and to fragment it into pieces. The second step is to mass analyze the resulting fragments. Fragment ions formed from cleavage of the peptide backbone are most beneficial for protein identification, since they allow for specific identification and fast computer searches of the protein sequence databases [1, 6, 13].

Unfortunately all MALDI-MS/MS methods that generate fragment ions independently of the desorption/ionization process, such as Q-TOFs, suffer lower fragmentation efficiency compared to ESI. Typically, in MALDI singly protonated ions are observed whereas in ESI doubly protonated ions are observed. The formation of sequence-specific fragments is significantly more pronounced for doubly charged ions compared to singly charged ions, regardless of the ionization technique employed [15]. One MALDI-MS/MS technique, however, that has been widely used for several years is post source decay (PSD) introduced by Kaufmann et al. [16, 17]. PSD utilizes the high efficacy of "hot" MALDI matrices, in particular of α-cyano-4-hydroxycinnamic acid, to enhance metastable fragmentation of MALDI generated ions. PSD has been routinely applied to protein identification [18] with unsurpassed sensitivity even down to the mid-attomole range [6] and to the characterization of protein modifications such as phosphorylation [19, 20, 21, 22]. The unimolecular decay of metastable peptide ions (LID) primarily results in single or double backbone cleavages. Successful PSD on low femtomole to attomole levels of sample was predominantly observed on instruments that used a gridless ion reflector and also allowed ions to be space-focused on to small detectors that provided low noise levels. On such instruments the sequence coverage of the N-terminal b- and the C-terminal y-ions [23, 24] significantly exceeds the peptide sequence coverage provided by fragmentation of doubly protonated ions in ESI-MS/MS [25].

Certain disadvantages to PSD prevented wider acceptance of this technique.

1. One-stage and two-stage reflector instrument designs require the spectrum to be recorded in 10–20 steps, each step measured at a lower reflector voltage than the last. This classical method of acquiring PSD data requires 10–60 min per spectrum.
2. Curved-field reflector instrument designs have a relatively short field free region which means limited time is available for fragmentation, hence the fragmentation is less efficient [26].

In both instrument designs, the kinetic energy of the fragment ions decreases towards lower mass fragment ions thus reducing the efficiency of ion detection using multichannel plate detectors. In addition, fragmentation occur-

Fig. 1 Schematic diagram of the LIFT-TOF/TOF mass spectrometer. TOF1 ranges from the MALDI ion source to the LIFT cell, TOF2 from the second accelerator stage in the LIFT cell to the reflector
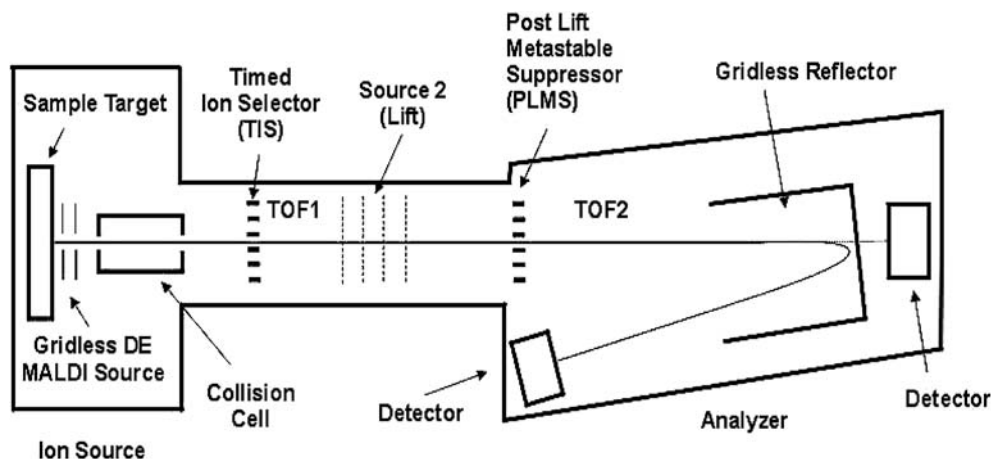
ring in the reflector can give rise to unresolved peaks in the spectrum. This is very disadvantageous, especially for the curved-field reflector as the ions spend most of their flight time inside the reflector, hence increasing the likelihood of unresolved peaks in the spectrum.

Most MS/MS protein identifications are based on nano-electrospray–MS/MS [1, 22] due to the lack of a MALDI-MS/MS technique that is fast and sensitive. The availability of such fast and sensitive MALDI-MS/MS is of high importance in the field of proteomics, since it is currently the only possible way to achieve high throughput protein identification. This technique can also analyze the same sample "stored" on a target for two types of analyses without the need for further sample preparation.

The first approach to overcome the described disadvantages of PSD mainly relied on high-energy collisions to fragment peptides in a TOF/TOF mass spectrometer [27, 28]. This technique can produce fragment ion spectra similar to the type produced by four-sector instruments [29] that show the backbone ions a-, b-, y- and i-, internal ions and the side chain d- and w-ions [30]. In general, MALDI-TOF/TOF instruments are characterized by a co-linear arrangement of two TOF mass analyzers, each equipped with an ion source that allows acceleration and focusing of the ions. In TOF1 the analyte ions are accelerated, precursor ions are selected and fragmented. Fragment ions are allowed to proceed to the "source" of TOF2, where they are accelerated and mass analyzed (Fig. 1).

The desire to overcome the disadvantages of PSD and to utilize LID to its fullest potential triggered the development of LIFT-TOF/TOF. While the basic idea of TOF/TOF, i.e. acceleration of a selected precursor ion together with its fragments is very simple, major refinement of the ion optics was required to successfully gain high mass resolution and sensitivity in MS/MS mode.

The new instrument provides a very high performance MS mode such that resolution, mass accuracy, and sensitivity combine to achieve a high success rate for identification and specificity for PMFs. Two MS/MS modes are available: LID and CID. LID is typically used for protein identification and high energy CID for de novo sequencing, in particular for resolving the Leu/Ile ambiguity. As

an additional "pseudo-MS/MS" method for protein sequence analysis ISD [31, 32, 33], in particular reflector in source decay (reISD) can be used due to the gridless ion optics of the instrument. ReISD of intact proteins as large as 60 kDa produces fragment ion spectra whose peaks are monoisotopically resolved, allowing N-terminal sequences and structures to be characterized [34, 35]. Although not related to the TOF/TOF concept, reISD is a versatile extension of the functionality of the instrument allowing the detailed characterization of recombinant or isolated proteins prior to their use as bait proteins in interaction proteomics experiments [1].

Principle of the new MALDI LIFT-TOF/TOF mass spectrometer

The LIFT TOF/TOF mass spectrometer (Ultraflex TOF/ TOF, Bruker) consists of a gridless MALDI ion source with delayed extraction (DE) electronics [36, 37, 38, 39, 40], a high-resolution timed ion selector (TIS), a "lift" device for raising the potential energy of the ions, a further velocity focusing stage with subsequent post-acceleration (source 2), a post lift metastable suppressor (PLMS), a gridless space-angle and energy focusing reflector, and fast ion detectors for the linear and reflector mode (Fig. 1).

The MALDI ion source accepts target plates of the same size and shape as microtitre plates. AnchorChip plates, equipped with 384 or 1536 hydrophilic anchors on an otherwise hydrophobic surface [41] were used as they enhance the sensitivity for peptide samples at least by a factor of ten [7, 42]. The instrument is equipped with an automated delivery robot for 20 sample plates (Twister, Zymark), which are stored under cover at ambient air pressure to prevent sample degradation by matrix evaporation in vacuum. Either a bar code or a code from an integrated target transponder can be read by the mass spectrometer. The code is used to specify for sample-specific analytical tasks. The targets are automatically introduced into the vacuum system through a vacuum lock. Image recognition methods are applied to find the exact locations of the sample spots. The time to move across the target from

Fig. 2 The full-range focusing method achieves highest mass resolution in an extremely wide range of masses (here from *m/z* 23 to *m/z* 3512) without any acquisition parameter adjustment
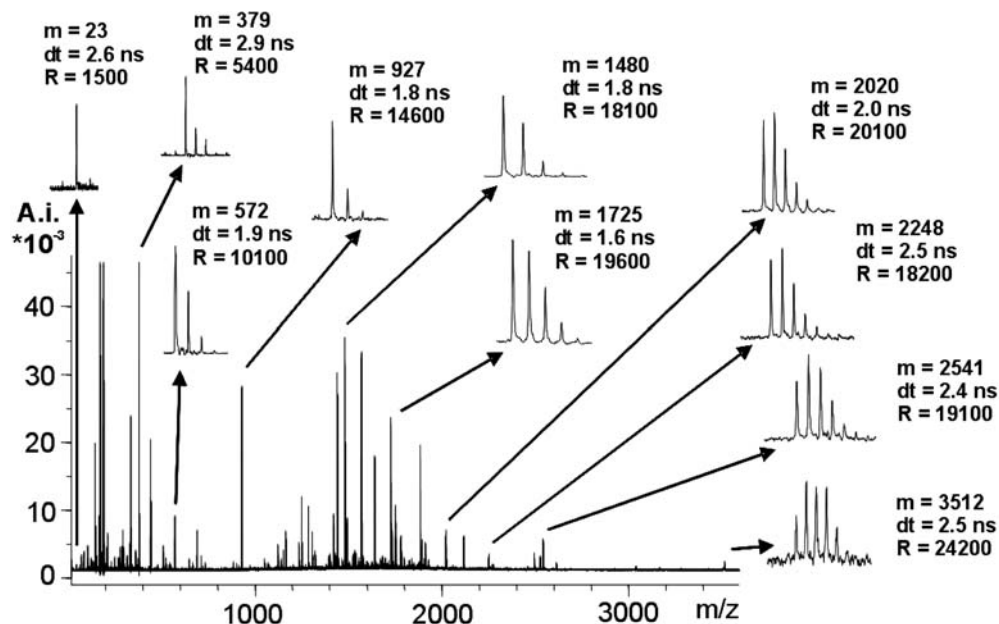
spot to spot amounts to about 0.3 seconds. These aspects combine to allow automated MS and MS/MS analysis in conjunction with robotic sample handling and transfer into the instrument without manual interference.

MS mode for PMF acquisition

In the MS mode, the instrument operates up to 25 kV as a MALDI-TOF spectrometer of highest quality with all modes of operation: linear, reflector, positive and negative acceleration potential.

In linear mode a gated MCP (micro channel plate) detector is used to detect proteins. Detector gating prevents detector saturation from low-mass ions and neutrals. The detector is placed on the ion optical axis to utilize metastable fragments and neutrals contributing to the protein signal. This increases the sensitivity especially for high-mass proteins.

In reflector mode a fast dual MCP detector is used with micro channels having a bore diameter of only 5 μm and a bias angle of 12°, allowing sub nanosecond pulse widths from single ions. It is optimized for symmetrical peak shapes and minimal ringing, which results in almost Gaussian shaped peaks that are perfect for determining the exact mass of isotopic peaks in the spectrum by software algorithms such as SNAP [43].

The instrument is prepared for a full mass-range high-resolution mode (broadband focusing) by modulation of the delayed extraction pulse shape [44, 45]. As shown in Fig. 2, highest mass resolution is achieved over the mass range from 500 Da to 4000 Da. In the lower mass range, the resolution is limited by the detection system which adds a constant signal width to the ion signal. This broadband focusing mode eliminates the need to adjust the acquisition parameters as a function of analyte mass. However, all work presented in this paper was performed without using this novel focusing mode, except for Fig. 2.

LID-MS/MS mode for protein identification

Acquisition conditions must be modified to generate high fragment ion yields. This is done by increasing laser fluence to provide a larger number of precursor ions per shot. The low initial accelerating voltage of 8 kV provides long flight time (10–20 μs) during which fragmentation occurs. The selected precursor and fragment ions are post-accelerated by an additional 19 kV and mass analyzed in one spectrum.

All precursor ions have a velocity corresponding to $E=1/2mv^2$ after acceleration to 8 kV. Fragments that are formed after acceleration maintain the velocity of the corresponding precursor. An "ion family" consisting of a precursor ion and its fragments will reach the timed ion selector (TIS) together. Different ion "families" from different precursor ions will reach the TIS at appropriate times. The TIS is capable of deflecting all ion families except the one under investigation by switching the gate voltage off while the selected ions pass through. A small fringe field results from this design that allows for high resolution when selecting ion families [46]. The TIS switches from a positive deflection voltage to ground and after delay of only a few nanoseconds to the opposite polarity. This compensates for the small effects of the fringe field that would otherwise reduce the resolving power of the device [47]. At threshold laser power, a TIS resolution, $R=m/\Delta m$ (FWHM), in excess of $R=1000$ is achieved and single isotopes can be selected (Fig. 3). In practical terms $R=1000$ means that ions at a distance of ±0.5 Da ($\Delta m=1$) from $m=1000$ have peak intensity reduction of up to 50%. Taking the Gaussian shape of the suppression characteristics, at ±1 Da the intensity would be down to 6.3% for $R=1000$, as shown in Fig. 3. However, at an increased laser power setting typical for MS/MS the mass resolution obtained from the first TOF region decreases from 1500 (at threshold) to 200–400. As a result the isotopic resolution cannot
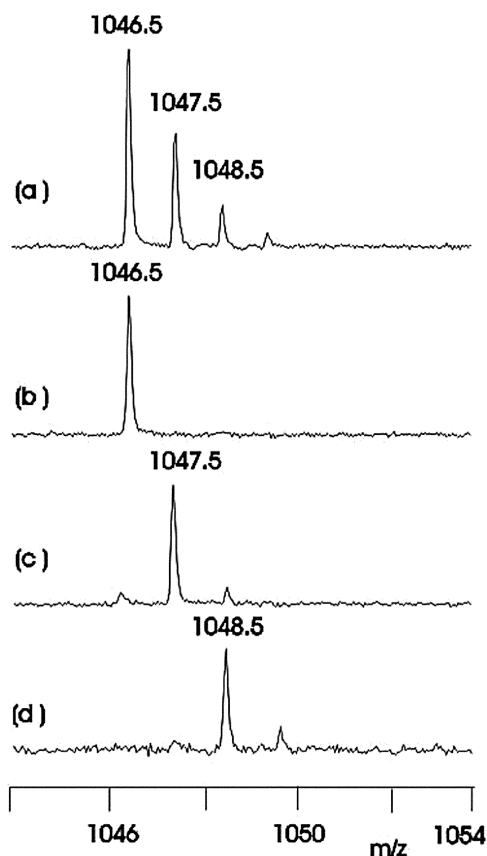
Fig. 3 Resolution achievable by the timed ion selector (i.e. the first TOF) at threshold laser radiation, showing (**a**) the isotope peaks of angiotensin II, and (**b**–**d**) the individually selected isotopic peaks. In practice, however, solely the entire isotopic cluster of a precursor ion is selected as shown in (**a**), even in the presence of other precursor ions nearby

Fig. 4 Calculated ion trajectories of the spatial focussing properties of the gridless reflector (SIMION). (**a**) Focussing of a divergent ion beam onto a small detector area with various beam origins but constant ion energy; (**b**) spatial focussing of ions with various energies (velocities) as they result from the LIFT process

be maintained in the typical peptide m/z range and the selection of the whole isotopic cluster becomes mandatory (data not shown). In fact, the resolution resembles that of a short linear TOF instrument, where isotopic resolution is obtained at laser threshold but is lost at conditions of higher laser fluence. Consequently, the electronics of the commercial instrument are adapted to select the whole isotopic cluster of a precursor ion, rather than single isotopes. The isotopic information can help interpret the MS/MS spectrum and assign fragment masses confidently. The selected "ion family" leaves the TIS and enters the "LIFT" device. Since ions that have an identical mass but a small velocity distribution start to drift away from each other, velocity-focusing is required.

The potential lift is the heart of the LIFT technology. It consists of three stages between four grids: the first stage is the actual potential lift, the second stage is a focusing cell to modulate the speed of the ions, and the third stage is a post-acceleration cell. Stage 1: As soon as the ions have completely entered the lift cell, the potential on the two adjacent grids forming the cell is rapidly increased from ground to 19 kV. Stage 2: The ions continue to travel at the same speed and enter the focusing cell that is also
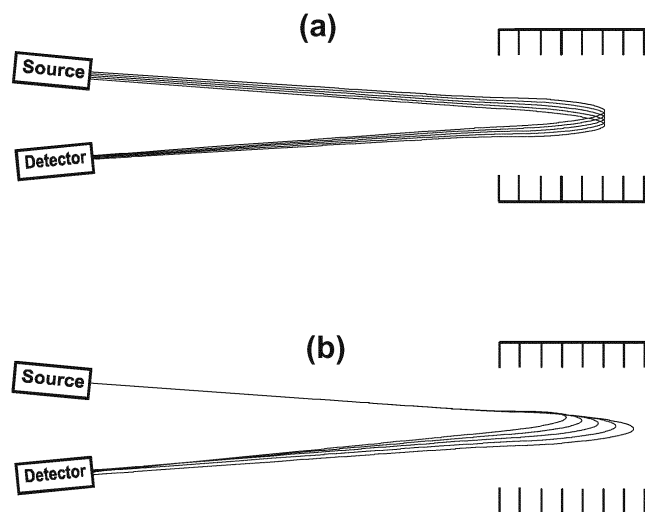
held at 19 kV. The potential on the third grid is now reduced by 2–3 kV and the ions are accelerated towards the third cell. Stage 3: In the third cell the ions are accelerated to full speed (similar to the delayed extraction process in a regular MALDI ion source) and time-focused onto the detector.

Precursor ions that leave the post-acceleration region are fully accelerated to a kinetic energy of 27 keV: 8 keV comes from the ion source, and 19 keV from the lift. While the precursor ions gain 8 keV in the ion source, the fragment ions have only a fraction of the energy, proportional to their mass; e.g., a fragment ion with m/z 100 derived from a precursor of m/z 1000 will have only 0.8 keV kinetic energy. As the potential lift adds 19 keV to this particular fragment, it will have a total energy of 19.8 keV. The kinetic energies of the fragment ions may range from 19 to 27 keV, a range that is most suitable for detection after passing the reflector (Fig. 4b). The gridless two-stage reflector also space-focuses the divergent ion beam onto the small area of the detector (Fig. 4a), a capability that reflectors with grids cannot achieve. Furthermore, most of the ions formed by metastable decay within the reflector field do not reach the detector and the gridless reflector allows a reduction in the number of times the ions pass a grid. Both properties are essential for the high sensitivity and high signal-to-noise ratio achieved by utilizing this design.

The instrument contains an additional device to suppress precursor ions. This "post lift metastable suppressor" (PLMS), situated between the LIFT device and reflector, deflects the remaining intact precursor ions and prevents undesired fragment ion formation after post-acceleration. The PLMS is an ion deflector much like the TIS. It is set to ground potential while the fragments pass; it is then raised to a high potential prior to the passage of

the precursor ions. Due to the elimination of the precursor ion, further fragmentation in the second TOF and the reflector is impossible; this results in low chemical background disrupting the LIFT spectra.

At present the laser operates with a repetition rate of 50 Hertz. For typical MS spectra with good signal-to-noise ratio and good dynamic range about 50–200 shots need to be accumulated. In MS/MS mode typically 300–600 shots need to be added, especially when analyzing low femtomole amounts of real-life samples. On the other hand, using picomole amounts it is possible to obtain a searchable spectrum from a single shot. Resolution is typically high enough to enable monoisotopic masses to be assigned to fragment ions, when the precursor ion mass does not exceed 2000 Da (see below) and a fragment mass accuracy of 0.2–0.4 Da is typical.

CID-MS/MS mode for Leu/Ile differentiation

In addition to the LID-MS/MS mode described above, a high-energy CID mode is available. This is useful for leucine/isoleucine differentiation and for the fragmentation of natural compounds or polymers. The ion source housing contains a collision region for high-energy collisions (Fig. 1). A collision gas (typically argon) is introduced under computer control, to increase the source pressure to 6×10^{-6} mbar. Under such conditions, tryptic peptides tend to create additional high-energy w-type ions [30] via side-chain fragmentation. This allows the isobaric amino acid residues leucine and isoleucine to be distinguished by their different side-chain fragmentation reactions.

Reflector ISD mode for fragmenting intact proteins

The MS/MS methods LID and CID that utilize fragmentation processes that occur after the precursor ions leave the ion source allow for the preselection of precursor ions. In contrast, in-source decay is a very fast fragmentation process that takes place during the MALDI process before the ions are accelerated out of the DE ion source [31, 32]. This prevents the dissipation of the excitation energy across the peptide that would result in the preferred cleavage of the most labile bonds such as the bonds N-terminal to proline or C-terminal of aspartic acid. The dissociation of a backbone bond takes place close to the excitation site causing uniform fragmentation patterns that are well suited for sequence determinations. In reflector mode (reISD), monoisotopic sequence ions can be observed, which are predominantly c-ions but to a lower extent also y- and a-ions [35]. They are well suited for sequencing the N-terminal region of intact proteins even larger than 60 kDa and allow its characterization even in the case of a blocked N-terminus [34]. Mass resolution of 5000–8000 and a mass accuracy of 50 ppm is typically observed in reISD spectra. Internal recalibration can be applied and mass errors of better than 10 ppm can be achieved.

ISD does not accommodate precursor ion selection, hence monodisperse analytes, e.g. chromatographically purified products are required. Typically, 5–10 pmol of pure protein are used for these measurements that require an extremely sensitive mass analyzer to obtain satisfactory signal-to-noise spectra since the efficiency of producing fragment ions by ISD is rather low compared to LID.

Material and methods

Gels

A colloidal Coomassie-stained Klose 2D gel [48] of a lysate of a human endothelial cell line was supplied by Martin Blüggel, Protagen, Dortmund. Protein lysate of the cell line (5 µg µL^{-1}, 50 µL) was deposited on the 2D Gel. One half of the gel was mounted on the A4 format scanner of the spot picker covering the pI range 5–10.

Sample preparation

Ninety-six protein spots with various intensity of stain were picked with a robot (Proteineer SP, Bruker) and deposited into microtitre plates dedicated to the Proteineer robotic platform. The microtitre plates were transferred to the digest/preparation robot (Proteineer DP, Bruker). A tryptic digestion (Trypsin, proteomics grade, Roche Diagnostics) was performed over 4 h at 30 °C in 10 µL digest buffer solution using a standard digestion kit for 96 samples (DP 96 Kit, Bruker) and the standard digest protocol on the robot. Documented methods were used for the thin layer sample preparation on 600 µm AnchorChip targets (Bruker) [42].

MALDI acquisition and data analysis

The AnchorChip target was inserted into an Ultraflex TOF/TOF instrument (Bruker) and submitted to an automated analysis loop using external mass calibration. Following MS acquisition each spectrum was submitted to a peptide mass fingerprint search using both Mascot (Matrix Science) and Profound (Genomic Solutions) search engines in the ProteinScape proteomics database system (Bruker). After automated assessment of the search results only those samples not unambiguously identified by PMF were automatically submitted to LIFT TOF/TOF acquisition by the work-
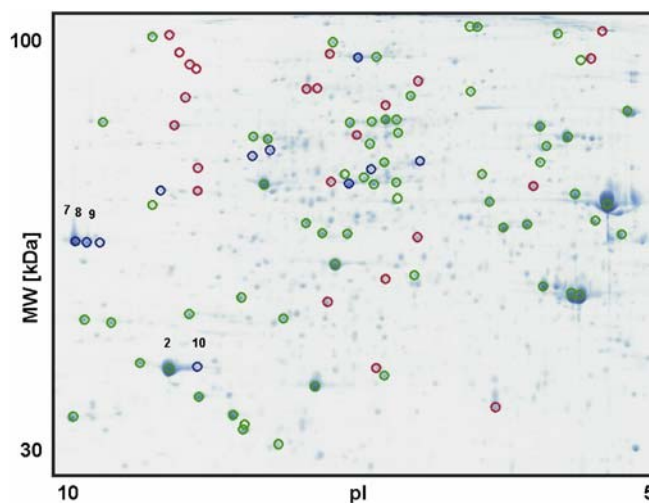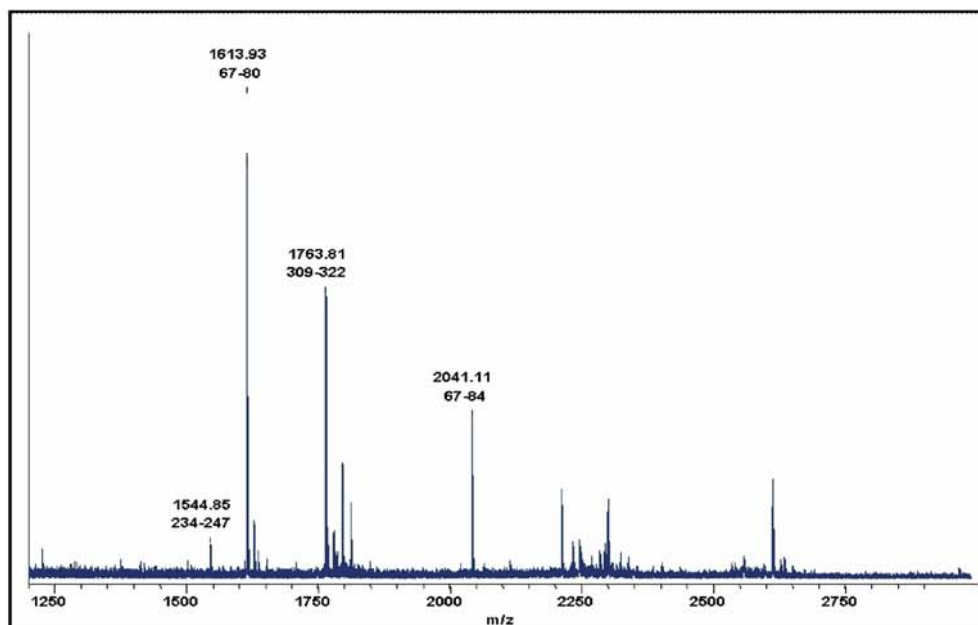

Fig. 5 2D gel with the picked spots in circles. *Green*: annotated by PMF; *blue*: annotated by MS/MS; *red*: identification failed. Spots discussed in this paper are labeled
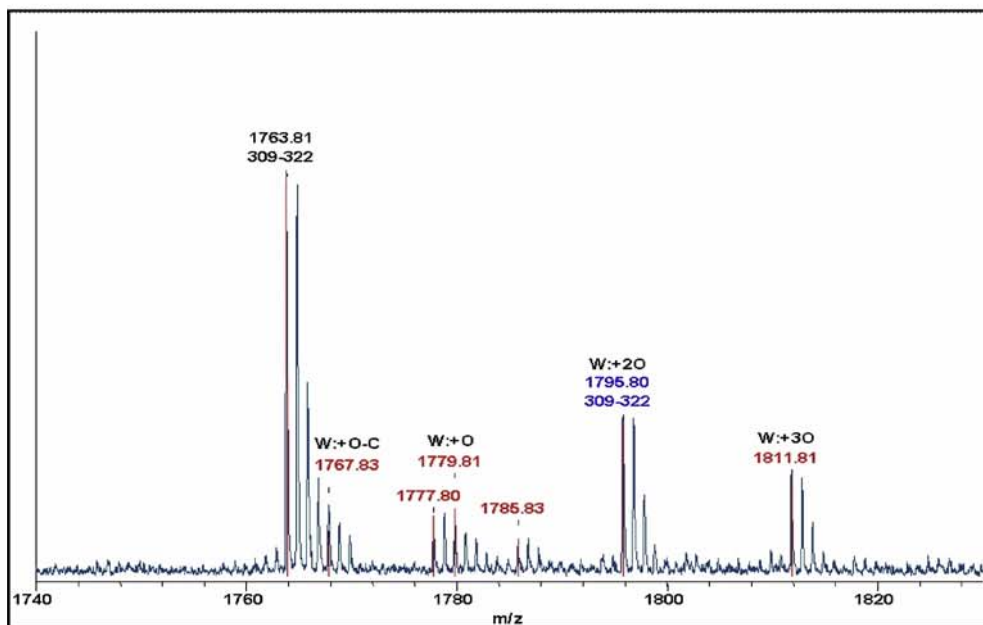
Fig. 6 PMF of spot 10 (**a**). Sequence coverage map of the combined dataset consisting of that PMF and of 3 LIFT-TOF/TOF spectra (**b**). The fragment ion information contributed by the MS/MS data is shown as two rows of red bricks in the gray peptide bars: the upper row shows matching b-ions; lower row, y-ions. (**c**) Detail of the spectrum that highlights the oxidation pattern in the Trp-containing peptide 309–322. Oxidation products appear in peptide signals (*m/z* 1767.83: +4 Da, *m/z* 1779.81: +16 Da, *m/z* 1795.80: +32 Da and *m/z* 1811.81: +48 Da, corresponding to kynurenine and the singly to triply oxidized Trp-residue, respectively. Matched ions in *black*, unmatched *red*, and assigned Trp-oxidized peptide in *blue*

a

b

c

Fig. 7 De novo sequencing result from LIFT spectrum of *m/z* 1795.80 (spot 10) with consensus sequence tags (*green*); calculation result (**a**) without assumption of protein modifications, (**b**) with the optional modification Trp+2Ox, the correct sequence that contains double oxidized tryptophan (W) is checked in the respective selection box. Numbers are sequence scores as provided by the software

flow control system (WARP, Bruker). A maximum of four precursor ions per sample were chosen for MS/MS analysis. Protein database searches, through Mascot, using combined PMF and MS/MS datasets were performed via BioTools 2.2 with the RapiDeNovo extension (Bruker). This software allowed in-depth analyses, for example mutation screening, analysis of protein modifications, and de novo sequencing. Homology searches that were based on the most likely de novo sequences were done using MS BLAST (http://dove.embl-heidelberg.de/Blast2/msblast.html) [49].
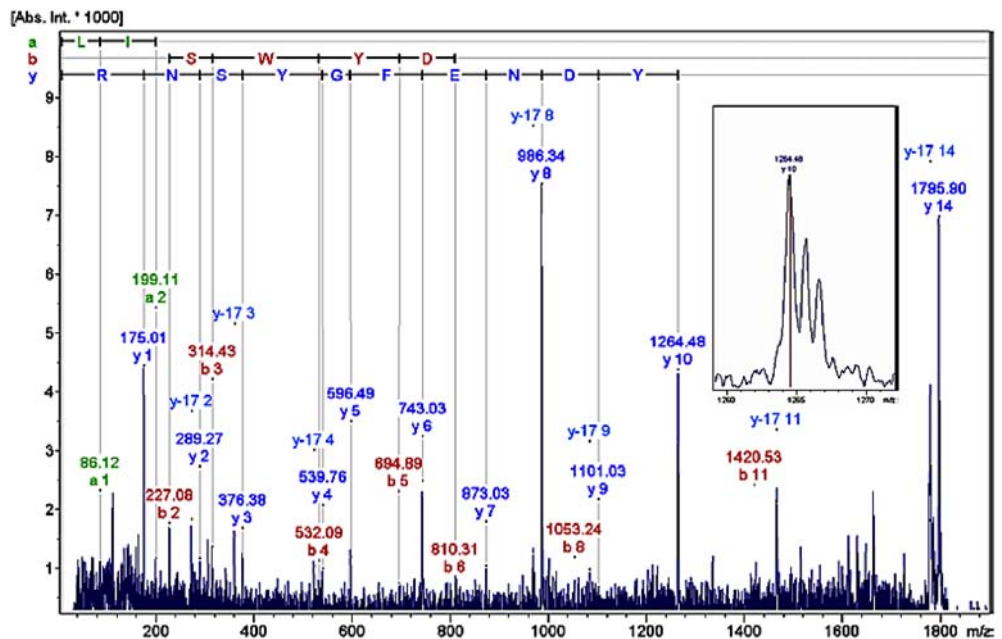
Recombinant *E. coli* thioredoxin was obtained from Sigma. Reflector ISD spectra were acquired as described [34, 35]. The peptide TDFEAAINTR was provided by H.E. Meyer (Medical Proteome Center University of Bochum, Germany).

Results and discussion

Automated analysis

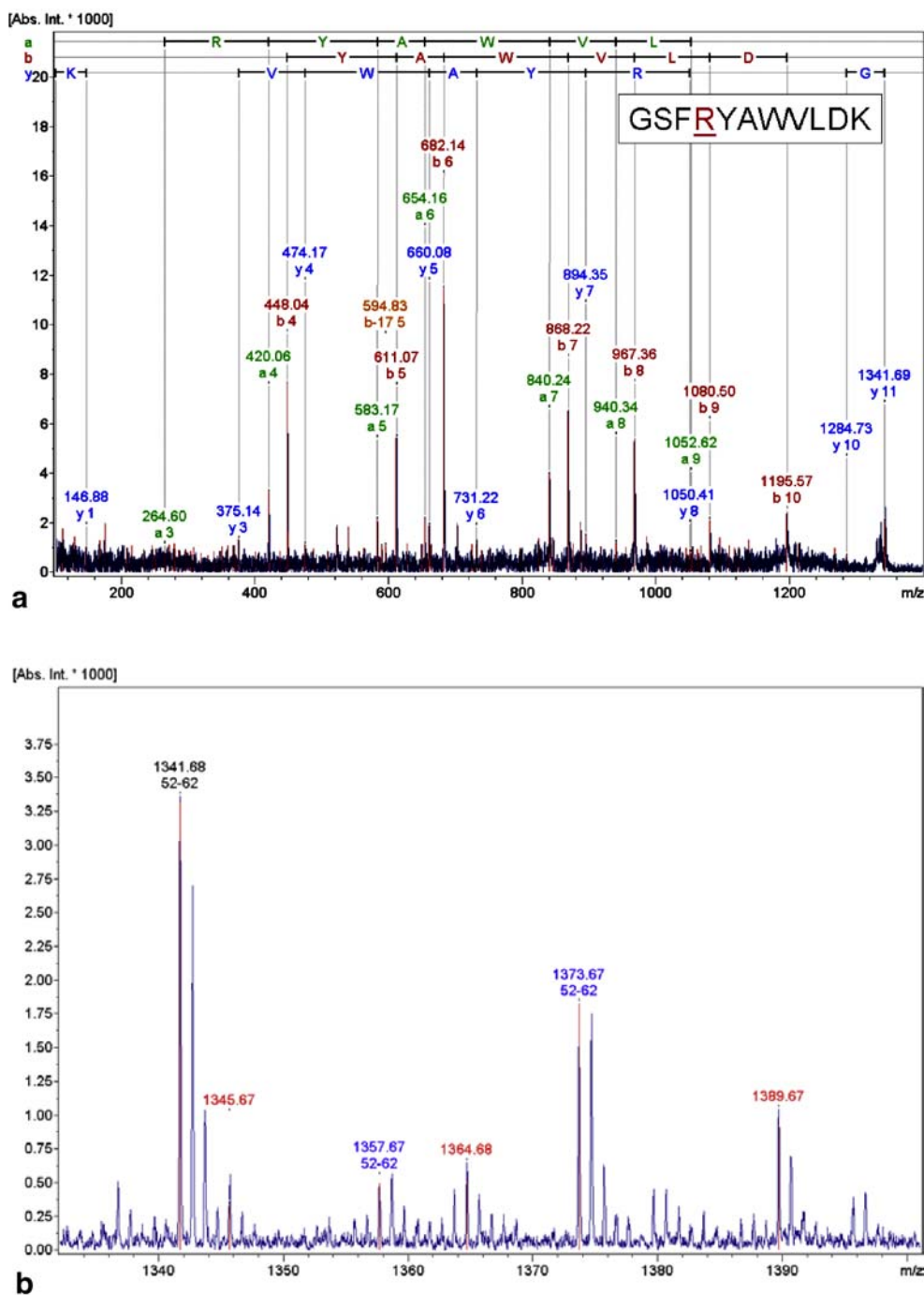Ninety-six spots were analyzed from a region of the 2D gel that covered the MW range 30–100 kDa and the pI range 5–10. Sixty-three spots were identified by the PMF alone and additional 11 on the basis of the automatically acquired MS/MS spectra that were used for combined searching of the PMF and up to four MS/MS spectra (Fig. 5). The success rate of identification therefore accumulated to 66% after PMF and 77% after TOF/TOF, which is the result of fully automated sample processing, spectra acquisition, and database searching. From such gels even 95% success rates have been reported after the TOF/TOF step (90% after TOF), but sophisticated multiple acquisition schemes were employed in that case that involved acquiring PMFs prior and after sample recrystallization on an AnchorChip target [51]. Here, only a standard sample preparation was used, which stresses the importance of sample preparation to get the utmost out of the same samples.

Most notably, the majority of the spots selected for additional MS/MS analyses were weakly stained spots presumably at lower abundance levels. Exceptions to the general observation are spots 7 and 8, to be discussed later. Therefore, the analytical sensitivity is of prime importance for increased success rate of the overall process utilizing MS/MS, if not only the stronger spots were to be MS/MS analyzed. In model studies we determined the mere instrument sensitivity for protein identification by PMF under conditions of full automation in the range of 200 attomoles of digest [52] on the target and for LID-LIFT spectra 1–10 femtomole [53]. However, the sensitivity for the workflow, i.e. protein amount in a gel plug is around 50–100 fmol for phosphorylase b, including MS/MS analysis [54].

Interactive analysis of unmatched MS/MS data

A few examples of interactive analyses are described now that revealed some greater insight into structural details. Spot 10 and the very intense spot 2, both allowed identifi-

Fig. 8 LIFT-TOF/TOF spectrum of *m/z* 1795.80 (spot 10) and assignment of the identified sequence 309–322: LISWYDNEFGYSNR, which contains doubly oxidized tryptophan (W)

960

cation of glyceraldehyde-3-phophate dehydrogenase (Fig. 5).
Since the identification from the weak spot 10 was unsafe
in the PMF, the workflow control software automatically
decided to acquire 4 MS/MS spectra. Three of these spec-
tra confirmed the assignment and the previous Mascot
score of 73 increased to 165, making it a unique and safe
assignment (Fig. 6). Only one of the auto-selected pep-
tides, *m/z* 1795.80, did not match.

Without further structural assumption, automated de novo
sequencing yielded the consensus sequence tag YDNEF-
GYSNR (Fig. 7a). An MS BLAST search with the 41 top

sequences gave a safe assignment to glyceraldehyde 3-phos-
phate dehydrogenase and a matching sequence of the pep-
tide 310–323 LISWYDNEFGYSNR that was 32 Da too
low compared to the molecular ion. That mass increase of
32 Da is typical for doubly oxidized peptides and investi-
gating the peptide sequence revealed a Trp residue that is
prone to protein oxidation, in particular of "double oxida-
tion" [50]. Repeating the de novo calculation with op-
tional double oxidation of Trp gave the correct sequence
in the answer list (Leu/Ile ambiguity unresolved, Fig. 7b) and
a good match between the LIFT spectrum of *m/z* 1795.80

Fig. 10 LID- (**a**) and CID-LIFT-TOF/TOF (**b**) spectra of the synthetic peptide TDFEA-AINTR. The sequence was determined automatically using the LID spectrum, except for the Leu/Ile ambiguity. The w_4-ion m/z 458.03 in the CID spectrum allowed assignment of Ile-7 as the only possibility
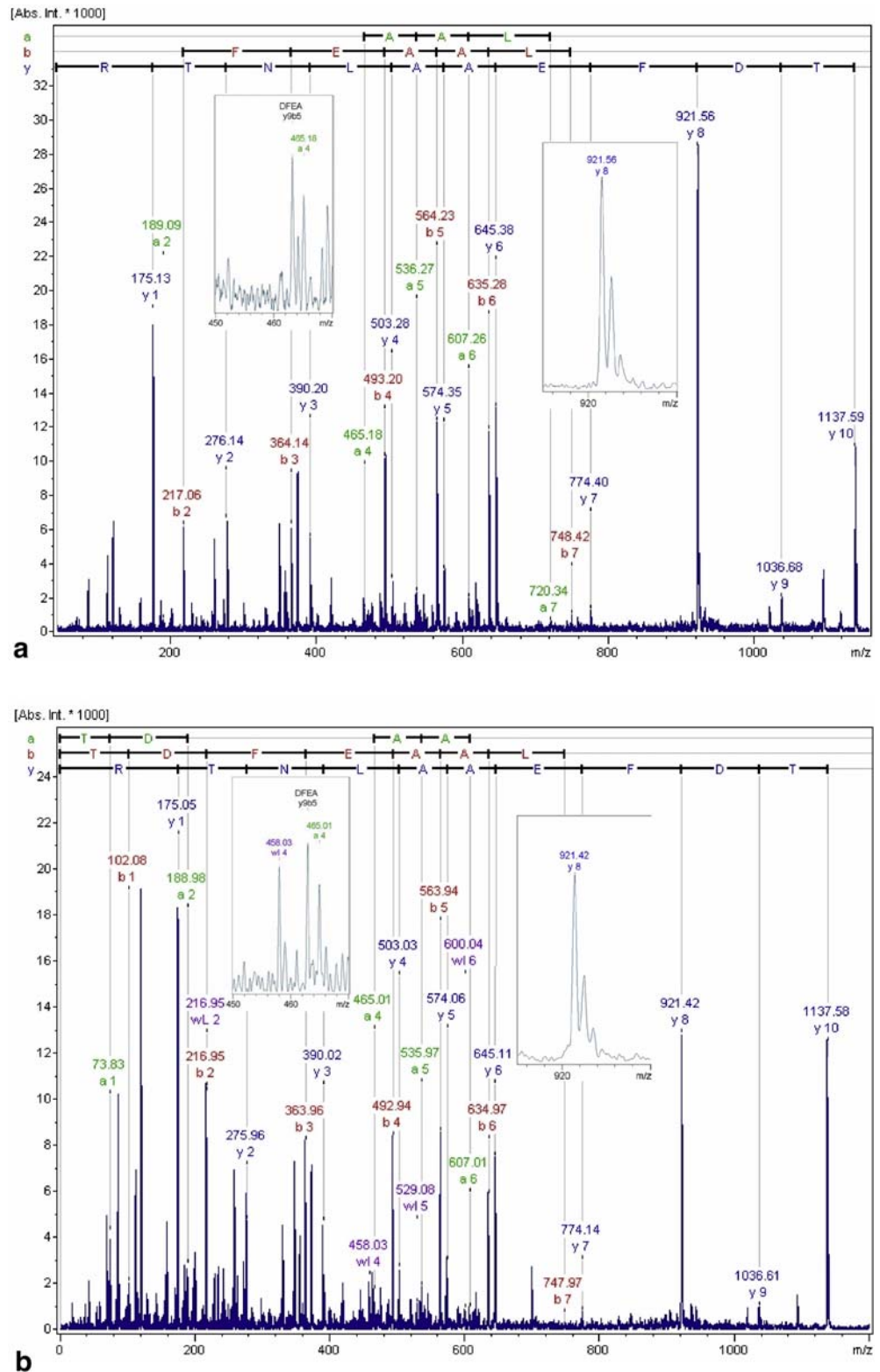
to peptide 310–323 with double oxidation of Trp (Fig. 8). In addition, the pattern of adjacent peaks (1763.81, +4, +16, +32 and +48) corresponds essentially to the pattern recently described as typical for tryptophan oxidation [28]. The +32 Da ion at m/z 1795.80 corresponding to double oxidation is the most intense oxidation product in the spectrum. The +4 Da ion at m/z 1767.83 was also observed by Bienvenut et al. [28] and was suggested to come from kynurenine. Remaining other peaks do not agree between the two examples (e.g. m/z 1777.80) and are likely

not linked to tryptophan oxidation. A major difference between the result from Bienvenut and the presented data is that the extent of oxidation is significantly higher for proteins bound to a blotting membrane, indicating a much stronger susceptibility to oxidation compared to the in-gel digestion protocol employed here. In several cases of Trp-containing peptides these patterns were reproduced and typically the relative intensity of the double oxidized peptide was in the range of 30–60% compared to the native peptide (see also Fig. 9b).

This example demonstrates that the sensitivity and dynamic range of the LID-LIFT technology provides insight in structural details even from rather weak spots and the weaker peptides in the PMF at conditions of full automation. Also, the automatically obtained MS/MS spectrum from m/z 1795.80 (Fig. 8) did not contain detectable levels of the oxidation product ions at a distance of ±16 Da and lower (ions presented on Fig. 6c) confirming a suitable TIS selectivity and calibration on real-life samples under conditions of automation. In this case, actually, a TIS resolution of 150 would have been sufficient to suppress the satellite ions.

Spots 7, 8, and 9 were only unsafely assigned to elongation factor α–1 based on the PMFs, although they are quite intense. Therefore, two MS/MS spectra were acquired from spot 8 during the automated run and m/z 1404.73 identified the protein safely (combined Mascot score 96). However, m/z 1341.68 did not yield a match to the sequence database. Therefore, that spectrum was matched to the EF α–1 sequence locally in the software, tolerating single amino acid exchanges in the sequence. The best match was identified by the BioTools software as peptide 52–62 GSFKYAWVL, revealing a K55R exchange (Fig. 9a). Whether this exchange is due to a DNA sequencing error or a polymorphism cannot be decided based on these results. However, the corresponding sequence of the New Zealand caddis fly was found to display exactly that sequence variant in a search without species restriction indicating a sequence error in the database. Interestingly with respect to discussion of tryptophan oxidation, that peptide also contains Trp and the same modification pattern of a dominating +32 Da ion and minor +4, +16 and +48 ions was found in the PMF (Fig. 9b).

As independent further confirmation of the presence of R[55], a RapiDeNovo analysis of this LIFT-TOF/TOF spectrum followed by a MS BLAST search of 255 calculated sequence candidates directly provided 48 matches to EF α–1 of various species. Although the N-terminal tripeptide sequence GAY... was incorrectly determined due to the absence of meaningful fragment ions, the correct main body of the sequence allowed homology searching with highly significant protein identification and to suggest the presence of the correct R[55] residue in contrast to the database entry.

This example shows how important software support is in elucidating structural details and how helpful the knowledge about typical features in PMF data can be to confirm computer generated peptide structures. It also indicates the significance of database errors and protein modifications – such as preparation artifacts – for the success of PMF-based protein identification.

High-energy CID-LIFT

For all proteomics routine measurements we currently use the LID-LIFT approach. However, the increased information content of high-energy CID spectra seems attractive for all cases where a maximum of information needs to be accumulated, e.g. the differentiation of the isobaric Leu/Ile and the almost isobaric Gln/Lys residues. In the course of a de novo-sequencing project of a synthetic peptide, LID and CID spectra were acquired using standard parameter sets (Fig. 10). The LID spectrum provided the full and correct sequence leaving only the Leu/Ile ambiguity at position 7: TDFEAA(L/I)NTR. The CID spectrum allowed the assignment of Ile due to the diagnostic w-ion m/z 458.03. The comparison between both spectra reveals insight into the information content typically provided by LID and CID spectra and allows to draw conclusions on their usefulness in proteomics studies.

In Table 1 the total fragment ion numbers present in the spectrum and assigned to the sequence are shown. Although the number of fragments increased by almost 40% using a collision gas, the number of assigned ions increased only by one. The number of assigned backbone fragments, internal (y–b)-ions and i-ions basically remained unchanged but one additional w-type ion was observed in the CID spectrum, which allowed assignment of Ile-7. Beyond that important additional information, however, the major effect of high-energy CID compared to LID seems to be the increased number of ions which are not assigned. In fact, they obscure the analysis, since they can currently not be used by the de novo sequencing software or by search programs such as Mascot, reducing the overall specificity of the analysis. In addition, the increased number of fragmentation channels in fact decreases the intensity of the backbone cleavage products a, b, and y that are most critical to protein identification. However, the increased intensity of low MW fragments such as the immonium ions (i-ions) and abundant diagnostic w-ions allow a "pinpoint analysis" on a more detailed level that is not accessible by LID and the low-energy CID spectra typical for ion traps or Q-TOFs. Based

Table 1 Comparison of the number of ions in various categories of the LID-LIFT vs. CID-LIFT spectra of the synthetic peptide TDFEAAINTR

	LID	CID	Δ(CID-LID)
Total	102	138	36
a, b, y ions	21	20	−1
a-17, b-17, y-17 ions	12	12	0
Internal ions (yb)	17	17	0
Immonium ions	4	5	1
Diagnostic w-ions	0	1	1
Assigned	54	55	1
Not assigned	48	83	35

Fig. 11 Reflector ISD spectrum of *E. coli* thioredoxin (MW=11674.3 Da). (**a**) Full spectrum containing the singly and doubly protonated molecular ions, metastable fragmentation products (LID) and the ISD ions, which are assigned in the expanded view (**b**) – the c-ion gap from c_{31}–c_{36} is due to the internal disulfide bridge. (**c**) Close-up view with the more intense c-ions and less intense y-ions at isotopic resolution
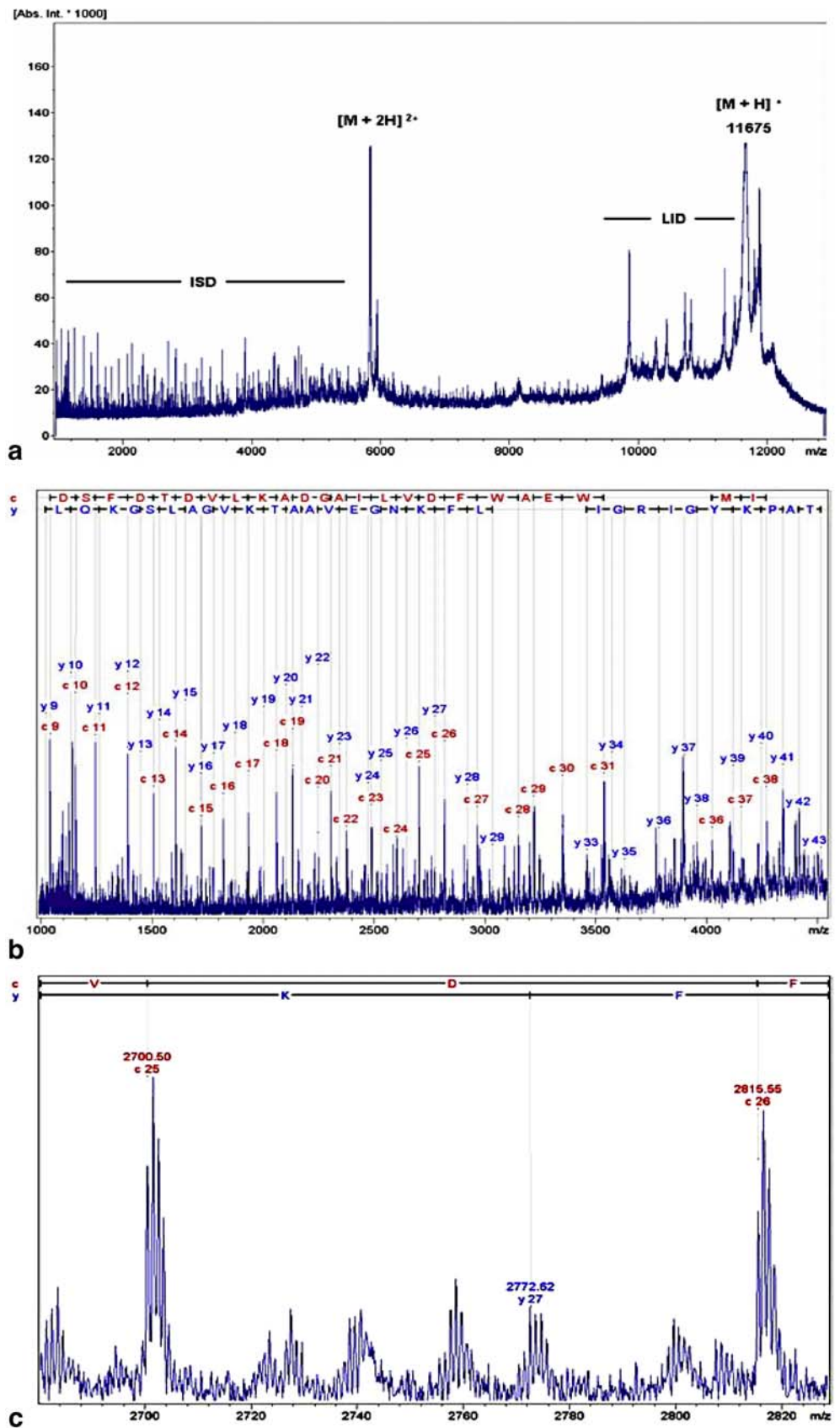
on such evidence we routinely use LID for protein identification, since it provides results with higher confidence and use CID as a second analytical step only where additional information is required.

Reflector ISD for the Characterization of Intact Proteins

For larger proteins, the MS/MS information provided by CID and LID is limited to the major backbone cleavage sites, i.e., C-terminal of Asp or N-terminal of Pro residues [55]. Therefore, these methods provide useful data to map known proteins, but the low resolution, mass accuracy, and sequence coverage does not permit using them for protein identification. Such LID ions of thioredoxin are also abundantly present in the reflector ISD spectrum in the m/z range 8000–11500 (Fig. 11a). More interestingly, in reISD, the m/z range 1000–4500 contains monoisotopically resolved fragment ions (Fig. 11c) that provided significant sequence readout based on c-ions and the less intense y-ions (Fig. 11b). Average absolute mass accuracy achieved was 0.11 Da and 49 of the 107 backbone cleavage sites (46%) provided sequence information. Therefore, such data provide c-ion tags that can be utilized to search for homology (BLAST) on intact proteins without the need of digesting them. Sequence readout can typically start from the 10th residue from either side, which renders reISD a very powerful technique for "near N-terminal" and "near C-terminal" sequencing [34]. However, protein crosslinks like the cystein bridge C^{32}–C^{35} in thioredoxin prevent sequence-ion formation, as the respective gap in the c-ion series shows (Fig. 11b). The sequence readout subsequent to C^{35}, therefore, confirms the proper state of the intact disulfide bridge, given a mass accuracy better than 0.2 Da.

Conclusion

The MALDI LIFT-TOF/TOF mass spectrometer enables peptide mass fingerprints (PMFs) and MS/MS spectra to be acquired from the same sample with high throughput and sensitivity. Close to 80% of the spots selected from a 2D gel were identified in a fully automated way that included the workflow from gel imaging to search result judgement. MS/MS spectra were only acquired from spots whose PMF failed to give unambiguous identification. Such spots typically furnished poor quality mass spectra, underlining the importance of sensitivity in both MS and MS/MS modes of operation. Important factors contributing to the achieved sensitivity were:

1. fully automated spot handling reducing the risk of sample contamination and adsorptive losses,
2. integrated sample preparation on AnchorChip targets known to increase sensitivity by an order of magnitude, and
3. a TOF/TOF instrument designed to utilize the abundant LID ions for MS/MS analysis equipped with a largely gridless ion optics for high signal-to-noise levels.

A single MS/MS spectrum can be acquired in a few seconds. Only a fraction of the sample is used which enables the acquisition of 20 or more MS/MS spectra per sample. The data obtained from this instrument are suitable for detailed analyses such as automated de novo sequencing and protein homology searching. This extends the scope of MS-based library searching to error-tolerant searching and to the identification of proteins across species. Tryptophan oxidation ($+O_2$, +32 Da) was found to be a major artifact of MALDI sample preparation that has been largely ignored during protein identification studies (unlike methionine oxidation), possibly due to the low frequency of tryptophan occurrence in protein sequences.

High-energy CID produced significantly more complex spectra without adding information significant to protein identification. LID appears to provide a higher specificity for protein identification due to a higher ratio of sequence-specific backbone fragments over fragments that cannot be assigned to the sequence and a higher fragment ion yield (Table 1). Ions important for diagnostic purposes are, however, generated by CID making CID a useful secondary analytical step. The most prominent example is differentiating between Leu/Ile, which will become relevant if DNA probes are to be designed to screen cDNA libraries and to eventually sequence the corresponding protein at the DNA level. Another interesting application likely to benefit from side-chain fragmentations provided by high-energy CID on the TOF/TOF instrument is the elucidation of protein modifications. Here, specific differences in the side chain fragmentations pattern could distinguish, say, hydroxyproline from leucine or could facilitate the sequence determination of oligosaccharide component of glycoproteins due to the observation of cleavages across the ring [56]. It is likely that such applications will be highly dependent on high-energy CID in the near future.

Together with the reflector ISD technique, the MS/MS capabilities of the described instrument make it a very versatile tool for solving protein analytical questions far beyond the routine task of protein identification.

References

1. Pandey A, Mann M (2000) Nature 405:837–846
2. Lander ES et al. (2001) Nature 409:860–921
3. Venter JC et al. (2001) Science 291:1304–1351
4. Gygi SP, Rochon Y, Franza BR, Aebersold R (1999) Mol Cell Biol 19:1720–1730
5. http://www.nobel.se/chemistry/laureates/2002/public.html
6. Gevaert K, Vandekerckhove J (2000) Electrophoresis 21:1145–1154
7. Schuerenberg M, Luebbert C, Eickhoff H, Kalkum M, Lehrach H, Nordhoff E (2000) Anal Chem 72:3436–3442
8. Schultz J, Doerks T, Ponting CP, Copley RR, Bork P (2000) Nat Genet 25:201–204
9. Miller C, Gurd J, Brass A (1999) Bioinformatics 15:111–121
10. Choudhary JS, Blackstock WP, Creasy DM, Cottrell JS (2001) Trends Biotechnol 19(10 Suppl):S17–S22
11. Moseley MA (2001) Trends Biotechnol 19(10 Suppl):S10–S16

12. Lee H, Griffin TJ, Gygi SP, Rist B, Aebersold R (2002) Anal Chem 74:4353–4360
13. Griffin TJ, Gygi SP, Rist B, Aebersold R, Loboda A, Jilkine A, Ens W, Standing KG (2001) Anal Chem 73:978–986
14. Vandekerckhove J (2002) 5th Siena Meeting From Genome to Proteome: Functional Proteomics, September 2–5, Siena
15. Cramer R, Corless S (2001) Rapid Commun Mass Spectrom 15:2058–2066
16. Kirsch D, Spengler B, Kaufmann R (1993) Proc 41st ASMS Conf on Mass Spectrometry and Allied Topics, San Francisco, May 30–June 4
17. Kaufmann R, Kirsch D, Spengler B (1994) Int J Mass Spectrom Ion Proc 131:355–384
18. Purcell AW, Gorman JJ (2001) J Immunol Methods 249:17–31
19. Sickmann A, Meyer HE (2001) Proteomics 1:200–206
20. Sickmann A, Marcus K, Schaefer H, Butt-Dorje E, Lehr S, Herkner A, Suer S, Bahr I, Meyer HE (2001) Electrophoresis 22:1669–1676
21. Talbo GH, Suckau D, Malkoski M, Reynolds EC (2001) Peptides 22:1093–1098
22. Gevaert K, De Mol H, Puype M, Houthaeve T, De Boeck S, Vandekerckhove J (1998) J Protein Chem 17:560–560
23. Roepstorff P, Fohlman J (1984) Biomed Mass Spectrom 11:601–601
24. Biemann K (1990) Methods Enzymol 193:886–887
25. Gevaert K, Demol H, Martens L, Hoorelbeke B, Puype M, Goethals M, Van Damme J, De Boeck S, Vandekerckhove J (2001) Electrophoresis 22:1645–1651
26. Cornish TJ, Cotter RJ (1993) Rapid Commun Mass Spectrom 7:1037–1040
27. Baldwin M, Medzihradszky K, Juhasz P, Vestal M, Campbell J, Chen X, Burlingame AL (1999) Proc 47th ASMS Conference on Mass Spectrometry and Allied Topics, Dallas, June 13–17
28. Bienvenut WV, Deon C, Pasquarello C, Campbell JM, Sanchez JC, Vestal ML, Hochstrasser DF (2002) Proteomics 2:868–876
29. Medzihradszky KF, Campbell JM, Baldwin MA, Falick AM, Juhasz P, Vestal ML, Burlingame AL (2000) Anal Chem 72:552–558
30. Papayannopoulos IA (1995) Mass Spectrom Rev 14:49–73
31. Reiber DC, Brown RS, Weinberger S, Kenny J, Bailey J (1998) Anal Chem 70:1214–1222
32. Reiber DC, Grover TA, Brown RS (1998) Anal Chem 70:673–783
33. Schnaible V, Wefing S, Resemann A, Suckau D, Bucker A, Wolf-Kummeth S, Hoffmann (2002) Anal Chem 74:4980–4988
34. Resemann A, Schweiger-Hufnagel U, Witt M, Suckau D (2001) Proc 49th ASMS Conference on Mass Spectrometry and Allied Topics, Chicago, May 27–31
35. Suckau D, Cornett DS (1998) Analusis 26:M19–M 20
36. Reilly JP, Colby SM, King TB, US patent 5,504,326; US patent 5,510,613; US patent 5,712,479
37. Vestal ML, Juhasz P, US patent 5,625,184; US patent 5,627,369; US patent 5,760,393; US patent 6,057,543; US patent 6,281,493
38. Holle A, Köster C, Franzen J, US patent 5,654,545
39. Wiley WC, McLaren IH (1955) Rev Sci Instr 26:1150–1157
40. Colby SM, King TB, Reilly JP (1994) Rapid Commun Mass Spectrom 8:865–868
41. Schürenberg M, Franzen J, US patent 6,287,872
42. Gobom J, Schuerenberg M, Mueller M, Theiss D, Lehrach H, Nordhoff E (2001) Anal Chem 73:434–438
43. Köster C, Holle A (1999) Proc 47th ASMS Conference on Mass Spectrometry and Allied Topics, Dallas, June 13–17, MPA 003
44. Franzen J, US patent 5,969,348
45. Holle A, Franzen J (2001) Proc 49th ASMS Conference on Mass Spectrometry and Allied Topics, Chicago, May 27–31, ThPA 007
46. Vlasak R et al. (1996) Rev Sci Instrum 67:68–72
47. Aratari R (1988) Nucl Instr Methods Phys Res B34:493–498
48. Klose J, Kobalz U (1995) Electrophoresis 16:1034–1059
49. Shevchenko A, Sunyaev S, Loboda A, Shevchenko A, Bork P, Ens W, Standing KG (2001) Anal Chem 73:1917–1926
50. Thiede B, Lamer S, Mattow J, Siejak F, Dimmler C, Rudel T, Jungblut PR (2000) Rapid Commun Mass Spectrom 14:496–502
51. Langen H, Berndt P, Suckau D, Schuerenberg M (2002) Proc 50th ASMS Conference Mass Spectrometry and Allied Topics, Orlando, June 2–6, ThPB 036
52. Hufnagel P, Schürenberg M, Ehrhard B, Resemann A, Suckau D (2000) Proc 48th ASMS Conference Mass Spectrometry and Allied Topics, Long Beach, June 2–6, WPH 315
53. La Rotta A, Holle AE, Hillenkamp F (2001) Proc 49th ASMS Conference Mass Spectrometry and AlliedTopics, Chicago, May 27–31, MPA 003
54. Schuerenberg M, Hufnagel P, Suckau D, Bailey S, Wattenberg A, Blueggel M, Gobom J (2002) Proc 50th ASMS Conference Mass Spectrometry and Allied Topics, Orlando, June 2–6, MPA 041
55. Yu W, Vath JE, Huberty MC, Martin SA (1993) Anal Chem 65:3015–3023
56. Spina E, Cozzolino R, Ryan E, Garozzo D (2000) J Mass Spectrom 35:1042–1048