Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Chromatography A, 1020 (2003) 11–26
JOURNAL OF CHROMATOGRAPHY A
www.elsevier.com/locate/chroma

Using accurate mass electrospray ionization–time-of-flight mass spectrometry with in-source collision-induced dissociation to sequence peptide mixtures

Jon D. Williams [a,*], Michael Flanagan [b], Linda Lopez [b], Steve Fischer [b], Luke A.D. Miller [a]

[a] *Discovery Research, GlaxoSmithKline, 5 Moore Drive, Research Triangle Park, NC 27709-3398, USA*
[b] *LC-MS Products, Pharmaceutical Solutions Unit, Agilent Technologies, 5301 Stevens Creek Blvd., P.O. Box 58059, Santa Clara, CA 95052-8059, USA*

Abstract

Although data-dependent LC–MS–MS with database searching has become *au courant* for identifying proteins, the technique is constrained by duty-cycle inefficiency and the inability of most tandem mass analyzers to accurately measure peptide product ion masses. In this work, a novel approach is presented for simultaneous peptide fragmentation and accurate mass measurement using in-source collision-induced dissociation (CID) on electrospray ionization (ESI)–time-of-flight (TOF) MS. By employing internal mass reference compounds, mass measurement accuracy within ±5 ppm for tryptic peptide precursors and ±10 ppm for most sequence-specific product ions was consistently achieved. Analysis of a complex solution containing several digested protein standards did not adversely affect instrument performance.

Keywords: Interfaces, LC–MS; Instrumentation; Mass spectrometry; Peptide sequencing; Orthogonal acceleration time-of-flight mass spectrometry; Molecular mass; Peptides

1. Introduction

A revolution in biological mass spectrometry began in the mid-1980s with the application of electrospray ionization (ESI) [1] and matrix-assisted laser desorption ionization (MALDI) [2] to the analysis of biopolymers. In concert with the development of protein, genome, and EST databases, focus shifted from detailed characterization of purified, recombinant proteins to systematic identification of proteins implicated in biological processes [3,4]. Lee et al. [5] identified intact proteins of the yeast large ribosomal subunit using accurate mass measurements to search a yeast database. In more sophisticated experiments, amino acid sequence information was obtained using collision-induced dissociation (CID) of intact protein ions [6–8]. Archived protein sequences were queried directly using the combination of accurate protein mass measurement and partial sequence tag information [8].

* Corresponding author. Tel.: +1-919-483-5270; fax: +1-919-483-3411.
E-mail address: Jon.D.Williams@gsk.com (J.D. Williams).

0021-9673/$ – see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2003.07.019

To obtain more useful sequence data, the approach most commonly employed involves enzymatically digesting proteins, then using chromatography to separate the resultant peptides prior to mass analysis [9]. Proteins are subsequently identified by searching databases with the acquired mass spectra [10]. Peptide mass fingerprinting [11] and sequence-specific peptide fragmentation [12] algorithms are routinely applied. In peptide mass fingerprinting, the peptide molecular masses are used to query protein sequence strings. Database searching algorithms include Mascot (www.matrixscience.com [13]) and MS–FIT (http://prospector.ucsf.edu [14]), among others [15]. Alternatively, peptide ions are selected and collisionally dissociated to generate a sequence-specific set of product ions. Both the peptide molecular mass and tandem mass spectrometry (MS–MS) spectra are required to take full advantage of the more sophisticated method for database-aided peptide identification available through algorithms such as Sequest [16] or Mascot MS–MS Ions Search [13].

Peptide mass fingerprinting is commonly accomplished using MALDI coupled to a time-of-flight (TOF) mass spectrometer [17]. Although MALDI–TOF is a rapid and relatively simple way of obtaining molecular mass data, the ± 10 ppm mass accuracy of a typical high-performance TOF (at full width, half maximum (FWHM) with 10,000 resolution [14]) may be insufficient to obtain an unambiguous protein assignment, especially if the sample is a complex protein mixture and there are less than three hits to a particular protein in a database [18]. When protein mixtures are evaluated, the combination of molecular mass and peptide sequence information collectively contribute to confidence in protein identifications [14]. Although LC–MS–MS has emerged as the method of choice, MALDI, using new instrumentation such as TOF–TOF [19], and atmospheric or low-pressure MALDI with tandem mass spectrometers [20,21], can also generate product ions suitable for database searching.

Protein identification by LC–MS–MS is not an especially high throughput or simple process, often due to the temperament of the components employed. The peptides are first separated chromatographically, often by ion exchange prior to nanoscale reversed-phase HPLC [22]. Eluting peptides are fed to a nanoflow ESI source for ionization [23,24] and dissociated in the collision cell of a tandem mass spectrometer. Ions are automatically selected for fragmentation based upon a user-determined criterion, such as intensity level, then proteins are identified by integrated data processing and database searching. The data-dependent capabilities of modern tandem mass spectrometers [25,26] provide the linchpin for this approach. Articles recently published on deciphering the proteomes of malaria parasites using quadrupole ion trap [27,28] and quadrupole-TOF [29] illustrate the power of this technique.

Dynamic exclusion is also used in combination with data-dependence to minimize acquiring redundant mass spectral data during an experiment. The instrument can be programmed to (1) fragment only multiply-charged species, (2) ignore an ion that has been analyzed and look at successively less abundant ions co-eluting during a given chromatographic peak, and (3) dynamically select the collision energy as a function of the *m/z* and precursor ion charge state. There are, however, some disadvantages to typical data-dependent operation. Optimal collision energy, for example, depends also upon the peptide structure, not just the mass and charge [30]. Therefore, excessive or mediocre dissociation may result in poor quality product ion spectra that yield inferior database searching scores.

Often, when doubly- and triply-charged ions of the same peptide co-elute, they are sequenced in separate MS–MS experiments. Although this may aid in the location of post-translational modifications, it also increases data redundancy, significantly slowing analytical throughput. Indeed, there is less of an opportunity to acquire data when spending extensive intervals in MS–MS mode, where the scan time and duty-cycle of the instrument are the least efficient. This is especially true if narrow chromatographic peaks are being examined by the instrument, such as when capillary electrophoresis [31], capillary electrokinetic chromatography [32], or reversed-phase monolithic columns [33] are used. Because tandem operation targets a specific ion species to be dissociated, the product ion spectrum will exclude internal mass reference ions (co-injected ions that appear in the same MS spectrum as analyte ions), so mass measurement accuracy may not be sufficient to provide unambiguous product ion mass assignments.

Instrument efficiency may be improved by either selecting multiple precursor ions and simultaneously dissociating them [34] or eliminating the precursor mass selection step altogether [35–37]. Li et al. [34] successfully demonstrated data-dependent selection of multiple precursor ions and simultaneous dissociation of the precursor ions using either infrared multi-photon dissociation or sustained off-resonance irradiation collisional induced dissociation using an FTMS. Without mass selection prior to precursor dissociation, improvements in duty-cycle efficiency have been accomplished by collisionally dissociating peptides in the ESI source [35,36] or in the orifice-skimmer cone at the back of the drift tube [37]. In the former experiments, low mass resolution quadrupole mass analyzers were utilized [35,36] while in the latter experiment a unique ion mobility quadrupole time of flight instrument was employed [37].

This manuscript describes the utilization of a versatile, low-cost, commercial prototype time-of-flight mass spectrometer to accurately measure the mass of parent ions along with precursor ions generated by collisional dissociation in the ionization source. By sequencing all the precursors that enter the source, this instrument increases efficiency because coeluting ions do not have to be individually targeted for CID in separate operations and more data points may therefore be obtained for a given chromatographic peak. Fragmentation efficiency for all compounds is improved by rapidly cycling the collision energy. Alternatively, the source may be tuned to optimally dissociate a specific precursor ion.

Accurately measuring precursor and product ion masses using as few as two scans can also provide greater confidence in peptide sequence determinations. One of the keys to enhanced mass accuracy in atmospheric pressure ionization (API)–TOF mass analyzers is the use of internal reference masses. We have observed that employing two electrospray nebulizers to simultaneously and independently deliver sample and reference mass compounds to the mass analyzer avoids ionization suppression that can occur when using spiked internal calibrants.

The objective of this study is to devise a more efficient, automated method to quantitate peptides and identify proteins. Three critical steps that need to be developed to fully realize this goal include: (1) the ability to simultaneously generate product ions and accurately measure the *m/z* ratios of precursor and product ions with a high duty-cycle instrument such as ESI–TOF, (2) accurate precursor ion quantitation using a global isotope label [38,39], and (3) development of a sophisticated database search engine to interpret and report data. This work discusses the successful adaptation of this instrumentation and demonstrates its ability to generate high quality product ion spectra on both individual and relatively complex mixtures of peptides.

2. Experimental

2.1. Sample preparation

Bovine serum albumin (BSA), and catalase tryptic digests were obtained from Michrom Bioresources (Auburn, CA, USA) and used as delivered. Enolase and α-casein, obtained from Sigma (St. Louis, MO, USA), and PPARδ, obtained internally at GSK (Research Triangle Park, NC, USA), were solvated with 10 mM Tris–HCl (pH 8.0) to form 1 ml solutions at concentrations of 3, 6, and 5 pmol/μl, respectively. The samples were digested overnight at 37 °C with modified sequencing grade trypsin from Roche Diagnostics (Indianapolis, IN, USA) at an enzyme to substrate ratio of 1:10. Formic acid was purchased from EMD Chemicals (Gibbstown, NJ, USA) and acetonitrile was obtained from Honeywell Burdick & Jackson (Muskegon, MI, USA). Eighteen megaohm water was generated by a Barnstead/Thermolyne Nanopure System (Dubuque, IA, USA), model D4741. Stock solutions of 3–6 pmol/μl were prepared using 0.1% formic acid and were diluted to 100 μl prior to use. Aliquots of 1.2–1.8 μl, corresponding to 1 pmol, were loaded onto a column using the 6-port Rheodyne micro-injection valve contained within an Agilent Technologies micro Well Plate Sampler module (Little Falls, DE, USA).

2.2. HPLC

An Agilent Technologies 1100 Series capillary LC system (Santa Clara, CA, USA) equipped with a binary capillary pump provided solvent delivery and separations. Protein digests were loaded onto a Zorbax XDB C_{18} column (Agilent, Little Falls, DE, USA),

150 mm × 0.3 mm, with 5 μm particle diameter and 300 Å pore size, maintained at 30 °C. After equilibrating in mobile phase A (water with 0.1% (v/v), formic acid) flowing at 4.5 μl/min, mobile phase B, consisting of acetonitrile with 0.1% (v/v) formic acid, was used to elute peptides from the column via the following gradient: wash and equilibrate at 3% B for 5 min, step to 15% B and then ramp to 60% B over 20 min, hold 5 min, ramp to 80% B over 20 min. The column was then re-equilibrated in mobile phase A.

2.3. TOF calibration

An Agilent Technologies ES Tuning Mix (P/N G2421A), with the addition of three proprietary compounds, provided mass reference ions for a 10-point baseline calibration of the mass axis. A 5–10 μM solution containing two reference mass compounds was introduced through a second capillary nebulizer at 200 nl/min using an Upchurch Scientific (Oak Harbor, WA, USA) Series 6500 precision dispense pump. The two reference masses (*m/z* 622.02896 and 1521.97148) were used to recalibrate the mass axis during the ESI–TOF analysis. These ions were resistant to fragmentation at all of the fragmentor voltages used in the experiments, so it was possible to precisely recalibrate both the precursor and product ion mass scales.

2.4. Mass spectrometry

Fig. 1 schematically depicts the layout of the prototype LC–API–TOF–MS instrument used for these experiments. Details include dual pneumatically-assisted capillary sprayers, components of the ESI source, transport optics, TOF analyzer, and multi-stage vacuum system. Ions were simultaneously generated from the sample eluting from an HPLC by the first capillary nebulizer and from reference compounds, introduced via the second capillary nebulizer. The

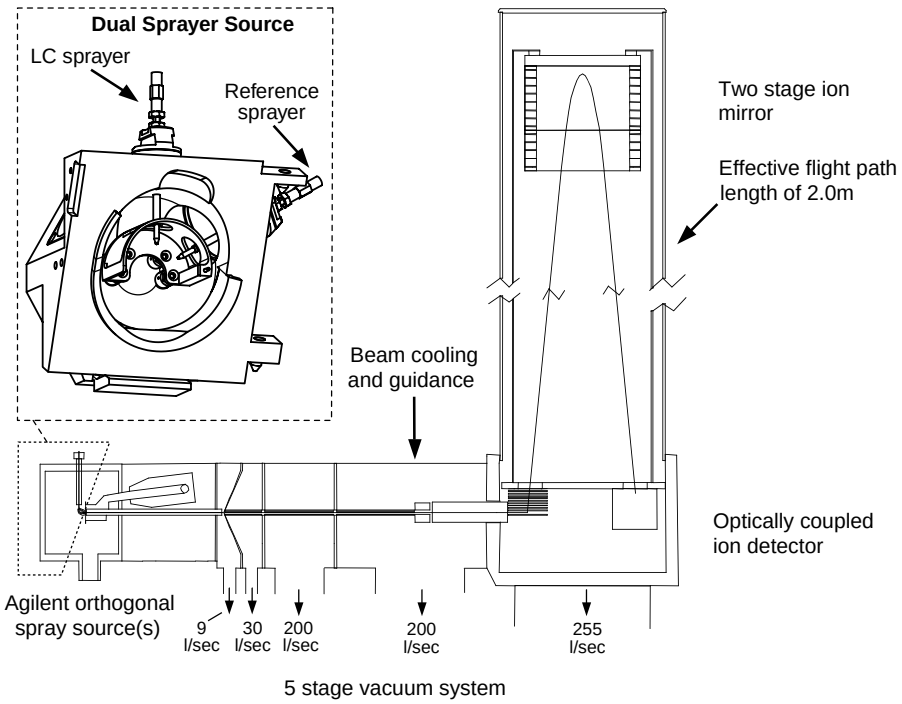

Fig. 1. Schematic diagram of the agilent LC–API–TOF–MS system prototype. Ions generated using an orthogonal spray enter the 5-stage vacuum system through a heated glass capillary. A skimmer and dual octopole decluster the ions and focuses the ion beam into a lensing system to shape the beam prior to injection into a pulsing grid. Ions are pulsed into the TOF mass analyzer and accelerated, then travel through a field-free region prior to reflection with a two-stage ion mirror. Ions are detected via an optically-coupled detector.

sprayers were positioned 7.5 mm in front of the spray nozzle shield and 15 mm apart at an angle of 60° between the respective sprayer needle tips. The nebulizers were held at ground potential. Each nebulizer spray needle was fabricated from 31 gauge hypodermic needle tubing with a nominal 50 μm i.d. The needle tips were tapered to allow: (1) for its adjustment within the exit tip of the ES nebulizer body and (2) flow of clean dry nitrogen (5–10 psig) through the ES nebulizer body to pneumatically-assist nebulization of the sprays (1 psi = 6894.76 Pa). Ions were then drawn into a gold-plated, glass dielectric capillary inlet (0.6 mm aperture) by a combination of voltage (−4.0 kV) and pressure differentials. Heated nitrogen drying gas (325 °C) flowing at 6 l/min around the capillary inlet was used to desolvate the droplets and aid the formation of gas-phase ions. The fragmentor voltage (the voltage differential between the capillary exit and the skimmer) was cycled between 140 V (to focus precursor ions into the skimmer) and over 300 V to collisionally dissociate precursor ions. An Edwards Vacuum Products E2M30 rough pump (Crawley, UK) was used to maintain a vacuum (400 Pa) in this region. An Edwards Vacuum Products Model EXT200/200/30 splitflow turbomolecular pump was used to provide differential pumping of the API–TOF vacuum manifold. Fig. 1 details the pumping speeds at the different instrument stages.

An octopole ion guide confined precursor or product ions in the second vacuum stage and transported them to a second ion guide traversing the third and fourth vacuum stages. The octopole radiofrequency (RF) was 450 V_{p-p} at a frequency of 5 MHz, ensuring transmission of a broad mass-to-charge range (70–3000 u). The ion beam was shaped in the fourth vacuum stage using a dc quadrupole lens. The flattened ion beam then passed through two slits before entering the ion pulser in the TOF analyzer vacuum stage (operating at a vacuum of 3.3×10^{-5} Pa using a separate 255 l/s Edwards turbomolecular pump). Positive ions were injected into the two-stage ion pulser and orthogonally accelerated into the flight tube using a −6 kV flight potential. The other end of the flight tube was attached to a two-stage reflectron, or ion mirror, that focused ions onto the detector. One-dimensional wire grids were used in both the ion pulser and mirror to optimize transmission and maintain beam focus.

The 25 mm diameter (active area) microchannel plate ion detector (Burle Electro-Optics, Sturbridge, MA, USA) amplified secondary electrons generated from ion impact and accelerated them to a scintillator. The resulting light was focused onto a high-speed photomultiplier tube photocathode to allow for its isolation from the flight tube voltage. The potential placed across the photomultiplier tube determined the final gain adjustment of the detector. Total "cycle time" of this prototype instrument was approximately 1.4 s, limited by the data registration speed capacities of the data acquisition board used in this prototype. Signal was acquired using a fast analog-to-digital converter (ADC), which has a much greater dynamic range than more common time-to-digital converter (TDC) devices that do not register data from low-level counts below the discriminator threshold and thus, require dead time corrections [40]. The ADC did not require any dead time corrections and was configured to record up to 10,000 transients per second.

3. Results and discussion

Examined in this work was the mass measurement accuracy of precursor and product ions generated by a prototype LC–API–TOF–MS equipped with an ESI source capable of in-source CID. Although peptide sequencing using in-source CID was previously described [36], a low resolution, single quadrupole mass spectrometer was employed. Because quadrupole mass spectrometers are typically operated at unit resolution, the charge states of precursor and product ion masses cannot always be unambiguously assigned. On the other hand, time-of-flight mass spectrometers having high resolution capabilities and high duty-cycle can correctly identify ion charge states and accurately measure ion masses, when the mass axis is re-calibrated using an internal reference mass.

In a TOF–MS, mass resolving power (R) is defined as:

$$R = \frac{M}{\Delta M} \tag{1}$$

where M is peak mass in u and ΔM the full width, half maximum peak width. Eq. (1) can also be written as:

$$R = \frac{t}{2\,\Delta t} \tag{2}$$

where t is the peak time of flight and Δt the peak width at FWHM. The higher resolution exhibited by TOF–MS instruments equipped with a reflectron results from the small spread in ion arrival times that contributes to the peak width. The advantage of high mass resolving power is analogous to that of high chromatographic resolution—more species can be resolved and detected. The LC–API–TOF–MS is able to perform accurate mass measurements since ion arrival times can be precisely measured and the mass axis can be corrected using internal reference masses.

We exploited the high resolution of the TOF to correctly assign charge states and obtain accurate mass measurements for precursors and product ions so peptide sequences could be unambiguously assigned. We first analyzed tryptic digests of proteins individually, then combined several protein digests together to model a simple proteomics experiment. Operating with reduced fragmentor voltage allowed peptide molecular masses to be obtained, elevated fragmentor voltage generated product ions.

3.1. Single protein analysis

Tryptic digests of several purified proteins were studied to evaluate the ability of the LC–API–TOF–MS to obtain precise molecular masses of precursor and product ions present in simple peptide mixtures. Plotted in Fig. 2a is a typical total ion chromatogram (TIC) for trypsinized α-casein. The chromatographic resolution was sufficient to ensure ions from only one or two peptides were analyzed by LC–API–TOF–MS at any given time. The chromatographic peak width was typically 0.347 min at the base, and 0.130 min at half height, corresponding to 20.8 and 7.8 scans, respectively. A two-step fragmentor method was employed. First, the fragmentor voltage was set to 140 V and peptide molecular masses were measured. Next, the voltage was rapidly switched to 375 V, leading to extensive dissociation. Since double the full width, half height peak width represents approximately 95% of the peak elution time, about seven dual-voltage fragmentation scans were possible over the course of each peak.

Example MS spectra obtained at low and high fragmentor voltages are presented in Fig. 2b and c, respectively. At the lower fragmentor potential (140 V), the dominant peaks resulted from the doubly- and triply-charged protonated molecules, corresponding to the [23–37] α-casein tryptic fragment (Fig. 2b). No other peptides were seen to co-elute at 21.7 min. Reference ions at *m/z* 622 and 1522 were detected at reduced abundance, but had sufficient signal for recalibrating the mass scale. Lower abundance peaks, in this case, resulted mainly from chemical noise. Depending upon the peptide structure, however, fragmentation products may be observed when particularly labile peptides are ionized and mass analyzed.

The fragmentor voltage was switched to 375 V to generate fragment ions (Fig. 2c). Reference masses were detected in addition to theoretically expected product ions, enabling mass axis recalibration and accurate measurement of product ion masses. Unlike a typical MS–MS experiment, where a precursor ion is isolated and analyzed individually, in this experiment the radiofrequency voltage and frequency of the transmission octopole is set to allow a broad range of masses into the mass analyzer. Hence, low-level chemical noise and the doubly-charged (*m/z* 880) and singly-charged (*m/z* 1759) precursor ions were also present in Fig. 2c. The entire peptide sequence could be deduced (see Table 1) from the y- and b-type product ions that were detected (product ions are identified following the Roepstorff–Fohlman nomenclature [41]), suggesting that the fragmentor potential was sufficient to generate a high-quality fragmentation spectrum for this peptide. Fragmentation yield (defined here as $\sum$product ion abundances/$\sum$precursor ion abundances and expressed as a percent) was found to be substantially higher using in-source CID, as compared to linear tandem MS–MS instruments equipped with a collision cell. We analyzed the enolase tryptic peptide [289–311] by in-source CID ESI–TOF and calculated the fragmentation yield to be ∼60%. Table 1 summarizes the calculated mass measurement accuracy for precursors and product ions of the α-casein tryptic peptide [23–37]. Precursor ion masses were reported within ±5 ppm of the theoretically expected values at either fragmentor voltage setting (140 or 375 V). The measurement precision of the product ions was slightly more variable, but still within ±10 ppm of the expected value.

Further examples of the quality of product ion spectra acquired at the high fragmentor potential using the two-step fragmentor technique are shown in Fig. 3. In Fig. 3a and b, complete sequence information could be

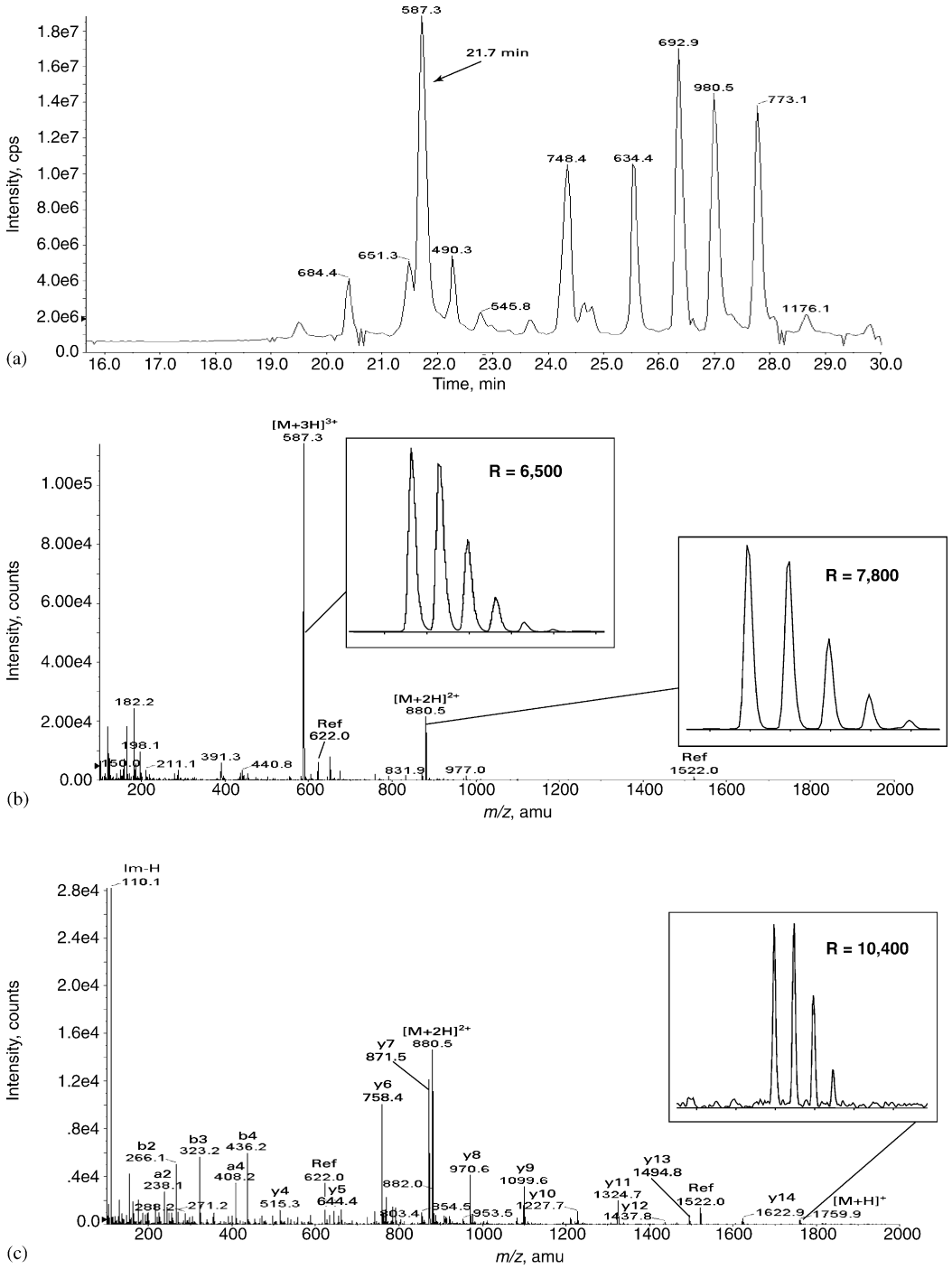

Fig. 2. (a) Total ion chromatogram (TIC) from a tryptic digest of α-casein. (b) The precursor mass spectrum of the tryptic peptide fragment corresponding to amino acids [23–37] reveals that only one peptide species elutes at the 21.7 min retention time. (c) The product ion spectrum resulting from in-source CID of the precursors. Sequence coverage of this peptide was complete.

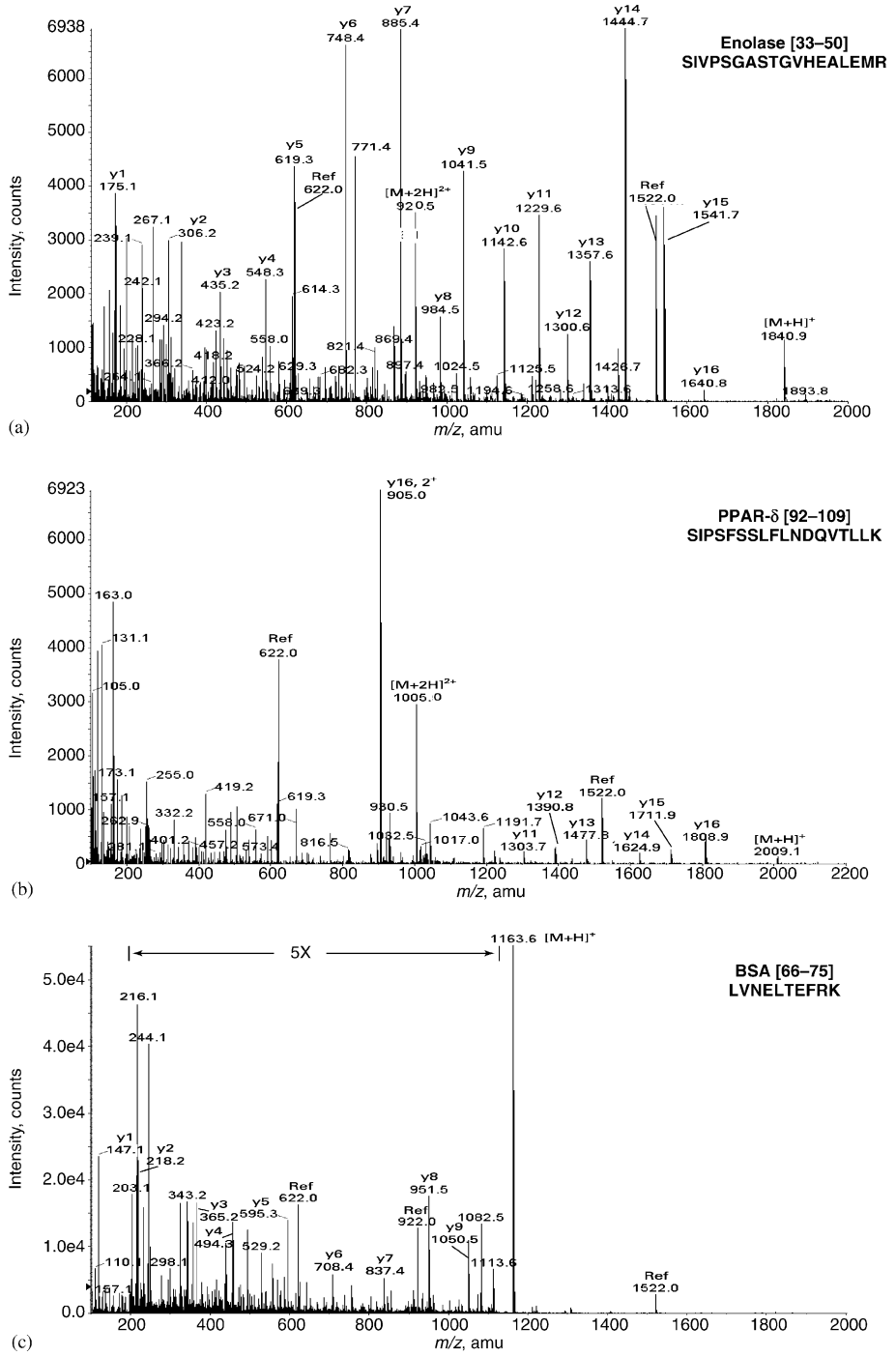

Fig. 3. A particular fragmentor voltage (375 V in this case) can effectively dissociate some peptides, such as (a) fragment [33–50] from enolase and (b) fragment [92–109] from PPARδ. However, less intense product ions are observed in the product ion spectrum of (c) fragment [66–75] from BSA.

Table 1
Mass measurement accuracy of precursor and product ions from α-casein [23–37] tryptic peptide

Identity	Measured (u)	Theoretical (u)	Accuracy (ppm)
3+	587.3226	587.3199	−4.7
2+	880.4799	880.4761	−4.3
1+	1759.9418	1759.9450	1.8
y14	1622.8740	1622.8860	7.4
y13	1494.8205	1494.8275	4.7
y12	1437.8065	1437.8060	−0.3
y11	1324.7190	1324.7219	2.2
y10	1227.6673	1227.6692	1.5
y9	1099.6086	1099.6106	1.9
y8	970.5679	970.5680	0.1
y7	871.4995	871.4996	0.1
y6	758.4151	758.4155	0.5
y5	644.3713	644.3726	2.0
y4	515.3302	515.3300	−0.5
y3	401.2876	401.2871	−1.3
y2	288.2048	288.2030	−6.4
y1	175.1200	175.1190	−5.5
b2	266.1259	266.1248	−4.1
b3	323.1479	323.1462	−5.3
b4	436.2316	436.2303	−3.0
b5	533.2843	533.2831	−2.2
b6	661.3408	661.3416	1.2
b7	790.3841	790.3842	0.1
b8	889.4516	889.4526	1.1
b9	1002.5327	1002.5367	4.0

derived for enolase [33–50] and PPARδ [92–109] from the spectra. Fig. 3c, corresponding to the BSA-derived tryptic peptide [66–75], produced a lower quality spectrum. For BSA [66–75], the most prominent ions detected were the Y1 ion (*m/z* 147) and the $[M + H]^+$ species (*m/z* 1163). The BSA [66–75] $[M+2H]^{2+}$ precursor ion at *m/z* 582 is not observed in Fig. 3c. However, the doubly-charged enolase [33–50] and PPARδ [92–109] precursor ions (*m/z* 921 and 1005 in Fig. 3a and b, respectively) were still detected in the product ion spectra. These data suggest that the fragmentor potential far exceeded the potential necessary to dissociate the BSA [66–75] $[M + 2H]^{2+}$ species and a lower fragmentor potential should have been used to generate higher intensity sequence-specific fragment ions. Given the many instances when high quality fragmentation spectra were achieved with the fragmentor potential set at 375 V, and other instances when this potential was too excessive to produce good fragmentation spectra, we decided to step the fragmentor potential in four discrete steps: 140, 225, 300, and 375 V. This "multi-step" experiment was implemented to conduct analyses of more complex mixture of peptides.

3.2. Protein mixture analysis

To further test the capabilities of the LC–API–TOF–MS system prototype, a peptide mixture containing tryptic peptides generated from BSA, α-casein, enolase, and catalase proteins was prepared for analysis. As expected, the number of species greatly exceeded the resolving capacity of the HPLC column and each chromatographic peak consisted of multiple co-eluting species. Fig. 4 shows the chromatographic complexity (refer to the TIC in the inset of Fig. 4a) and its effect on the acquired mass spectra. Fig. 4a plots the mass spectrum corresponding to the 21.7 min retention time peak, obtained when the fragmentor potential was set at 140 V. Five peptide species were readily identifiable from this spectrum.

The ions at *m/z* 582 and 1163 were assigned as the doubly- and singly-protonated BSA [66–75] tryptic fragment precursors. The ions at *m/z* 651 and 976 were assigned as the triply- and doubly-protonated phosphorylated α-casein [119–134] tryptic fragment precursors. Finally, the ions at *m/z* 722 and 1082.5 were assigned as the tryptic fragment precursors of triply- and doubly-protonated trypsin [50–69]. The other peptide species identified, α-casein [121–134] and α-casein [23–37] were detected at more abundant levels in the spectra acquired immediately preceding and following, respectively, those shown in Fig. 4. So, there are at least five precursor ions present in this chromatographic peak that would have required individual interrogations if this was a standard data-dependent MS–MS experiment, assuming singly-charged ions were excluded from being selected as candidate precursor ions. Two of these experiments would have yielded redundant information, since only the charge state differs for the α-casein [119–134] and trypsin [50–69] precursor ions. However, using the technique presented here, only three scans were necessary to obtain sufficient sequence information for all of these species.

Fig. 4b–d are the mass spectra recorded as the fragmentor potential (V_{frag}) was increased to 225, 300, and

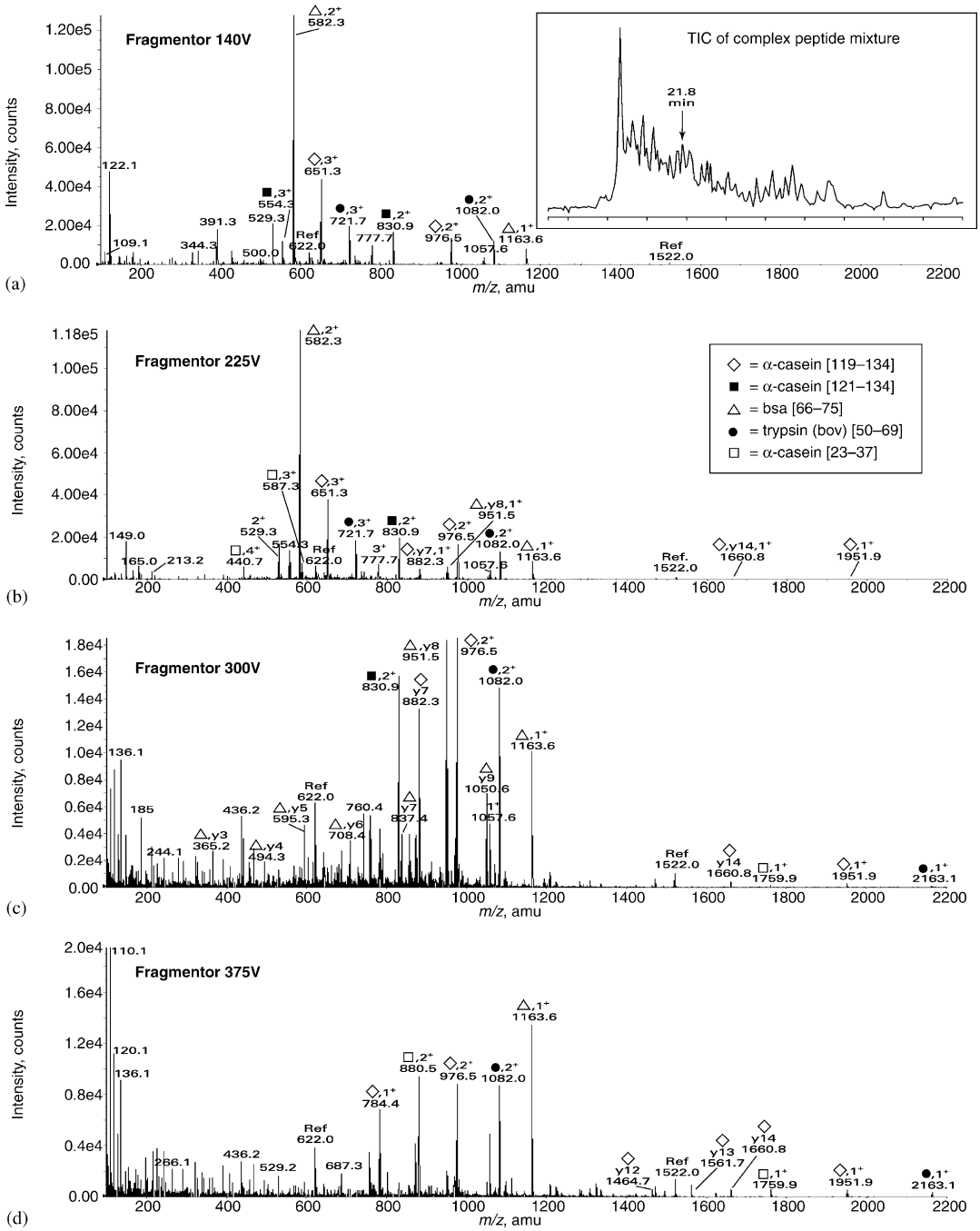

Fig. 4. Mass spectra of peptides generated from trypsinized BSA, catalase, enolase, and α-casein. The fragmentor was stepped from (a) 140 V, to (b) 225 V, to (c) 300 V, and finally, to (d) 375 V to improve fragmentation efficiency of co-eluting peptides. Peptides from BSA [66–75], (△); α-casein [119–134], (◇); bovine trypsin [50–69], (●); α-casein [121–134], (■); and α-casein [23–37], (□) appear or change intensity at individual fragmentor potentials that produce a more complete picture of the peptide sequence. The TIC for the peptide mixture is shown in the inset.

375 V, respectively. Many differences in product ion intensities were observed in these spectra. For example, minimal fragmentation was observed with V_{frag} set to 225 V, and the spectrum is practically identical to that of the precursor ion spectrum given in Fig. 4a. This fragmentor potential provided insufficient energy to induce efficient fragmentation of the peptide molecular species. As V_{frag} was increased to 300 V, and then 375 V, a greater number of abundant fragment ions were generated and singly charged protonated molecular ions were detected from each of the three abundant peptide species. There was a striking difference between these spectra, especially the ∼10 fold difference in the relative abundance of the BSA [66–75] y9, y8, and y7 fragment ions. Note that these ions were also detected at low levels in the "single protein" product ion spectrum acquired using V_{frag} of 375 V (Fig. 3c).

The improvement in the quality of the spectrum with V_{frag} at 300 V suggests that the reduced voltage setting was more effective in inducing fragmentation of BSA [66–75]. Complete sequence information for this species was derived from these data. A significant number of sequence specific product ions were assigned for the α-casein [119–134] tryptic fragment. Moreover, the extent of fragmentation information was sufficient to directly assign the phosphorylation site to Ser-130. Only a few product ions were assigned to the trypsin [50–69] autolysis fragment. However, given the accurate mass measurement of the precursor ions and that five consecutive y-type ions were detected, this species would likely be found by a database search engine that was designed to take advantage of data acquired in this fashion. By cycling through the three elevated fragmentor potentials, a more comprehensive collection of product ion spectra was generated than if the fragmentor had been operating at a single elevated potential. Parent ion selection was not necessary since (at least in this case) only a few precursor ions, some of which were just different charge states of the same species, actually contributed to the formation of product ions.

3.3. Mass measurement accuracy

The key to unambiguously sequencing peptides, either using a database or de novo methods, relies on spectrometer mass measurement accuracy. Although quadrupole-TOFs can measure precursor ion masses with ±10 ppm accuracy, the mass measurement accuracy for product ions is typically reduced by a factor of two to three times in typical peptide sequencing experiments. This clearly introduces ambiguity in assigning product ion identifications. The high-resolution LC–API–TOF–MS system prototype described here is capable of improved mass measurement accuracy for product ions because the product ions are generated as received in the source, rather than from an isolated precursor in a collision cell. This allowed us to use internal reference masses to precisely recalibrate the product ion mass axis. The total signal intensity of the product ions is substantially higher by in-source CID, since product ions can arise from multiple charge states of the same peptide precursor. Experiments were designed first to assess the mass measurement accuracy of single peptides, and then to investigate the effect of multiple co-eluting precursors on the mass precision.

Peptide ions from tryptic digests of purified proteins were chromatographically separated such that only one peptide would enter the mass analyzer at any given time. The mass accuracies of the precursor ions detected under these conditions are plotted for α-casein, enolase, BSA and catalase in Fig. 5a. Out of 49 precursors measured, 88% were measured with an absolute error within ±5 ppm. Only six precursors, corresponding to 12% of the total number of ions, were measured with a precision between ±5 and 10 ppm. In no case was the precursor absolute mass measurement error greater than 10 ppm.

The data clearly show that mass measurement accuracy for precursor ions is typically within ± 3.4 ppm (1σ). For confident sequence assignment based on molecular mass information only, the error tolerance (if automated interpretation is performed) should be 3σ, which is the 99.7% confidence limit [42]. In this case, the error tolerance for peptide identification, using only molecular mass data to conduct a database search, would then be ±10 ppm. Clearly, the LC–API–TOF–MS system prototype is well within this confidence limit for peptide mass fingerprinting applications.

Slightly degraded performance is graphically demonstrated in Fig. 5b, where the mixture of tryptic digests resulted in several co-eluting peptides in each chromatographic peak. The mass measurement

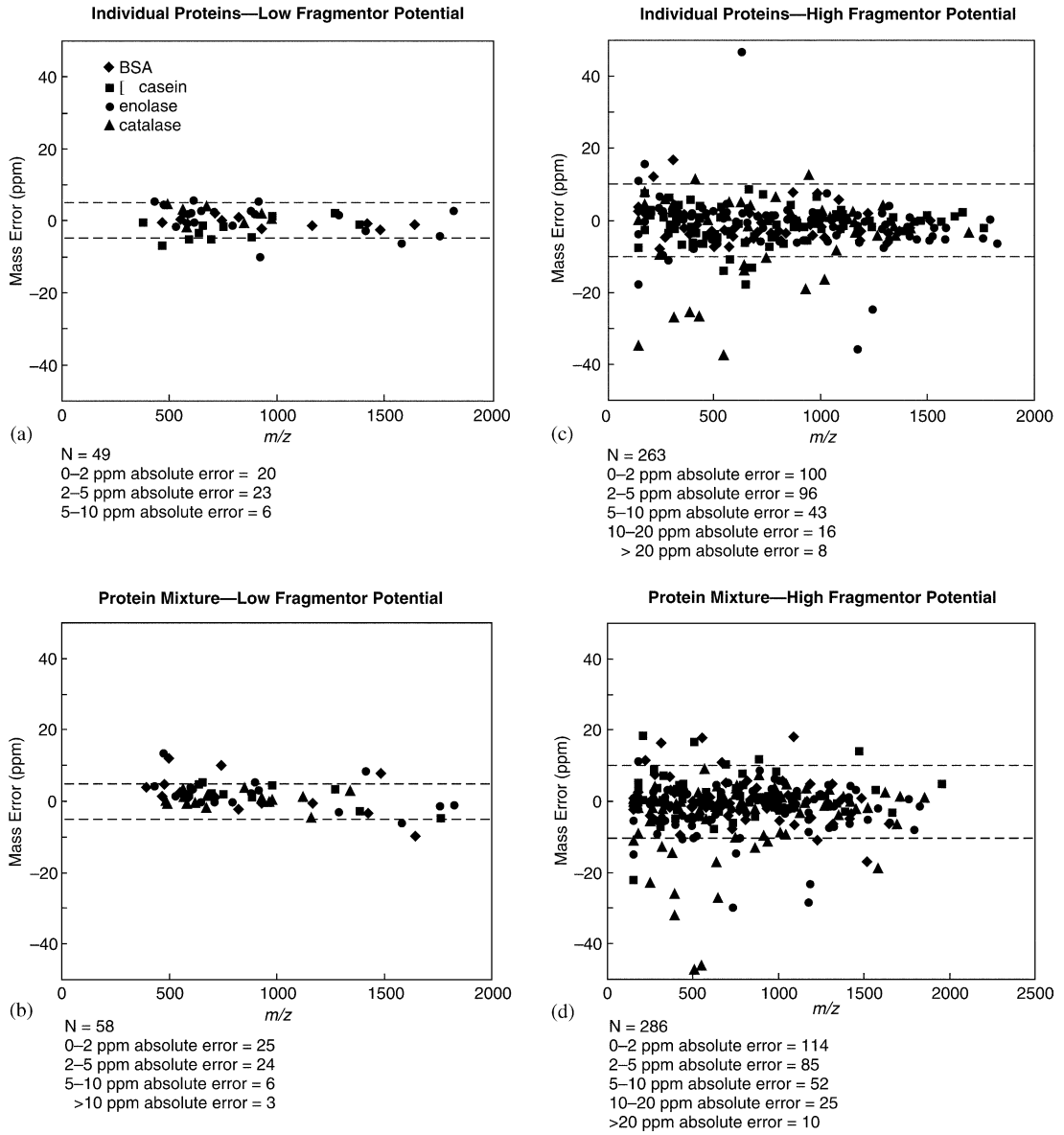

Fig. 5. Scatter plots of mass measurement accuracies for (a) precursor ions measured from tryptic digests of α-casein, (■); enolase, (●); BSA, (◆); and catalase (▲); (b) precursor ions from four digested protein mixture; (c) product ions generated by the precursors in (a), and (d) product ions generated from precursors in (b). The average error in (a) was 0.23 ppm with a standard deviation of ±3.4 ppm, that in (b) was 1.9 ppm with a standard deviation of ±4.0 ppm. The error in (c) was −1.0 ppm with a standard deviation of ±7.3 ppm, and that in (d) was −1.3 ppm with a standard deviation of ±4.0 ppm.

accuracy for all precursor ions was within ± 4.0 ppm (1σ). There was a slight mass-dependent bias (1.9 ppm), as determined by the average mass measurement accuracy. This bias was only 0.2 ppm when single peptides were analyzed individually (Fig. 5a). However, the increased bias could be attributed to a majority of the least accurate measurements all having positive deviations from their theoretical values.

The mass measurement accuracy of the product ions varies more widely than that of the precursor ions. As is seen in Fig. 5b, out of the 263 product ions surveyed, ~75% of the product ion mass assignments were within ±5 ppm, 16% were between ±5 and 10 ppm, 6% were between ±10 and 20 ppm, while only 3% had absolute measurement errors greater than ±20 ppm. The larger deviations in mass accuracy are caused primarily by the lower signal intensity of the fragment ions, when compared to the signal strength of the precursor ions, and interference with unresolved species isobars, such as those from chemical noise or other peptide fragments. Similar patterns are reflected in Fig. 5d (out of 249 product ions, 65% were within ±5 ppm, 20% were between ±5 and 10 ppm, 11% were between ±10 and 20 ppm, 4% were greater than ±20 ppm), suggesting the presence of product ions from multiple precursors does not adversely affect mass measurement accuracy. In addition, it should be noted that the fragmentor voltage (375 V) employed in this experiment was likely too high for the catalase peptides. These labile peptides did not dissociate to sequence-specific product ions; they tended to dissociate to individual amino acid residues, leading to inferior spectra. Analysis of these product ions at other voltages could possibly eliminate the larger spread attributed to catalase.

In Fig. 5b and d, low mass product ions appear to be measured with diminished accuracy. One factor contributing to errors at low *m/z* values ($m/z < 200$) is poor mass peak shape definition and lack of calibration correction (the lowest reference mass used is at *m/z* 622.02894). This is an inherent problem for time-of-flight mass spectrometers due to the high velocity and small spatial distributions of low mass ions arriving at the detector. A mass error of 0.001 u at *m/z* 200 translates into a 5 ppm error. Although the errors appear to be higher for low mass ions, they are typically well within 0.003 u. By using y_1 (generally lysine or arginine for tryptic peptides) or an immonium ion as an additional reference mass for the low end of the mass spectrum, it may be possible to significantly reduce this error. Schlosser and Lehman adopted this procedure to correct the low mass region of MS–MS spectra obtained on a quadrupole-TOF mass spectrometer [43].

The increased imprecision of mass measurement at low *m/z* is further explained by considering the factors that determine the peak width (pw) of an ion. Since pw is equivalent to Δt, the following expression can be derived from Eq. (2):

$$R = \frac{t}{2\mathrm{pw}} \tag{3}$$

The peak width is primarily dependent upon two general factors, ion arrival time spread and detector circuit pulse width. The ion arrival time spread ($\mathrm{pw_{ms}}$) is related to the system design, machining precision and stability of the electronics. It increases with the square root of mass. In a perfect world, the resolution of TOF–MS would be constant for all masses. However, the detector circuit limits the minimum peak width. For a single ion hitting the detector face, the circuit will give out a pulse of some width ($\mathrm{pw_{det}}$). Since the detector pulse width is not correlated with the ion arrival time spread, the two can be summed using the law of squares:

$$\mathrm{pw}^2 = \mathrm{pw}_{\mathrm{ms}}^2 + \mathrm{pw}_{\mathrm{det}}^2 \tag{4}$$

therefore

$$R = \frac{t}{2(\mathrm{pw}_{\mathrm{ms}}^2 + \mathrm{pw}_{\mathrm{det}}^2)^{1/2}} \tag{5}$$

The detector circuit pulse width will therefore limit the attainable resolution at the low mass end, since lower mass peaks have smaller $\mathrm{pw_{ms}}$ as compared to high mass peaks, hence mass measurement accuracy of ions with *m/z* values below 200 will be reduced as compared with higher mass ions.

As mentioned above, an additional source of error can be attributed to poor ion statistics, the result of an insufficient number of ions reaching the detector. The misshapen mass peaks obtained at low signal strength may not be centroided appropriately following mass correction. Fig. 6 illustrates the effect of signal strength on mass measurement accuracy. As one would expect, the precision of mass measurement is reduced at lower signal intensities. In the analysis of replicate injections, the spread or variability in reported mass values are a function of ion signal intensity, and mechanical and electronic “jitter”. Assuming the ion signal behaves as “shot noise,” the spread of reported mass values would change inversely with the square root of the ion signal intensity. If the jitter is independent of ion signal intensity, then the resultant spread in mass measurements

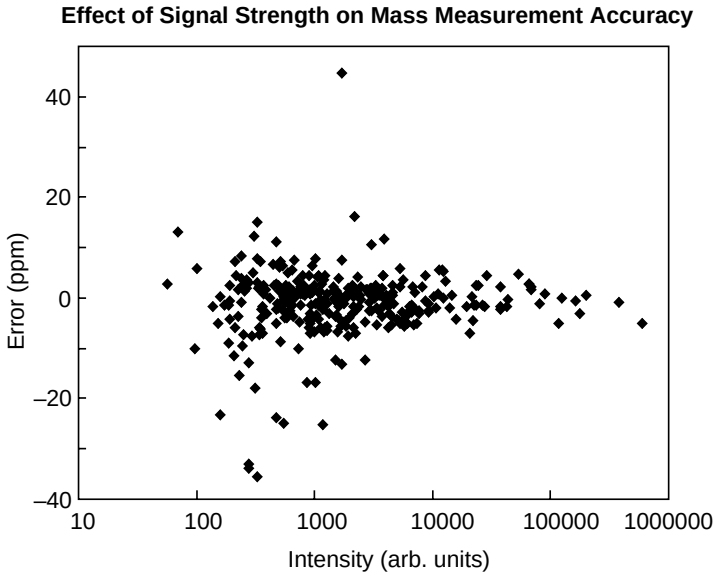

Fig. 6. Scatter plot of mass measurement accuracies for product ions as a function of signal intensity. The decrease in measurement error as signal intensity increases agrees with theoretical predictions (Eq. (6)).

will follow:

$$(\text{mass} - \text{axis spread})^2 = \frac{a}{(\text{ion signal intensity})} + b \cdot (\text{jitter})^2 \quad (6)$$

where a and b are system constants.

Thus, the reported mass-axis spread should improve inversely proportional to the square root of signal intensity, if jitter is constant. The data in Fig. 6 conform to Eq. (6), showing a decrease in the error spread from about ±15 ppm at 100 counts to within ±1 ppm at 10^6 counts. In other experiments using reserpine, we have observed ±5 ppm mass-axis spread (1σ) with 10 pg of sample loaded onto the column, and ±2–3 ppm spread for spectra averaged over 5–10 s (data not shown). Since many proteins of interest are expressed at low copy number in cells, the ability of the LC–API–TOF–MS system prototype to obtain excellent mass measurement accuracy, even at low ion counts, is likely to be very advantageous in proteomics applications.

4. Conclusions

In this work we have demonstrated the ability of the LC–API–TOF–MS system prototype to generate high quality product ion spectra for both single and complex mixtures of peptides. By rapidly switching the fragmentor potential during the experiment, a range of collision energies were imparted to the ions to improve overall ion fragmentation efficiency. Complete sequence coverage was observed for a considerable number of dissociated precursors; immonium ions and substantial “sequence tags” were obtained in most other instances. Thus, this technology platform generates data that is well suited for automated interpretation by both database searching and de novo algorithms.

Varying the collision energy was also helpful in the analysis of peptides from complex mixtures by exploiting structural differences in co-eluting peptides. Although the prototype instrument investigated here required values for the fragmentor potential to be user-input, the development of data-dependent relationships to automate dissociation of specific co-eluting precursors could likely yield further improvements in the quality of product ion spectra for even more effective database searching.

Over the course of many analyses (albeit using standard proteins) the mass measurement accuracy was consistently within ±5 ppm for most precursors and between ±5 and 10 ppm for most product

ions. Poor signal strength was noted as a common source of measurement error, recognizing that misshapen peaks may not have been centroided ideally. Therefore, ion abundance may be a good parameter with which to trigger automated software processing, where measurement precision is crucial for data analysis success. Typical proteomics applications require information from samples at the low fmol level and below. Under microspray ESI conditions, we estimate that the LC–API–TOF–MS system prototype can detect tryptic peptides at 20–50 fmol on-column with $S/N_{pk-pk} \geq 10{:}1$ (based upon extracted precursor molecular ions). It is likely that the development of a dual-spray nanoflow source will enhance this instrument's utility by providing high sensitivity at the low concentrations prevalent in modern applications with "one-pass" convenience.

The combination of reference mass calibrants, a zero dead-time fast ADC, and a photon-based detection system produced excellent mass accuracies for both precursor and product ions. In contrast to conventional tandem mass spectrometers, this platform does not require the precursor-product ion relationship to be specified prior to data processing. Sophisticated software algorithms are currently under development to use the superior mass measurement accuracy of the LC–API–TOF–MS system prototype to determine the identities of both precursor and product ions automatically.

Conventional tandem instruments eliminate chemical noise in the product ion spectrum by selecting a specific precursor to dissociate. In the LC–API–TOF–MS system prototype, every precursor is fragmented and all product ions are detected in a single experiment. Chemical noise will be present throughout the spectra and can appear concurrently at the same expected molecular mass of a product ion. This could result in slightly larger mass measurement errors for some ions, especially at the low end of the mass spectrum where resolution is diminished. A complete characterization of background ions will be crucial to the design of data processing algorithms so that signal from interfering ions can be accounted for and perhaps excluded. Hence, it may be feasible that the effect of chemical contaminants on mass accuracy can be moderated to some extent. The distinctive properties of this approach could prove useful in the development of de novo sequencing or database searching software, where lower mass (or lower abundance) peaks may be allowed a wider deviation in mass measurement accuracy than more abundant higher mass ions.

For quantitation, or under certain chromatographic conditions where peak widths are very narrow, the time required to write data to an external archival computer limits the number of data points that can be collected during an eluting peak. At present, the "cycle time" of the instrument is approximately 1.4 s. Efforts are currently underway to reduce the cycle time by an order of magnitude through hardware improvements. Low cycle time will be especially important when other goals of the project, such as the development of a global isotope label for quantitation, are introduced. Analysis of post-translational modifications has not been addressed in this work and will certainly be a consideration in prospective applications with this instrumentation.

Acknowledgements

The authors thank John Fjeldsted and Bill Russ (Agilent, Santa Clara, CA, USA) for sharing their expertise and insight; Will Old and Dean Thompson (Agilent, Fort Collins, CO, USA) for their help with data processing; and Karen Jonscher and Peter Jonscher of KJ Technical Communications, LLC (Denver, CO, USA) for providing invaluable assistance in the preparation of this manuscript.

References

[1] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Science 246 (1989) 64.
[2] M. Karras, F. Hillenkamp, Anal. Chem. 60 (1988) 2299.
[3] W.H. McDonald, J.R. Yates III, Traffic 1 (2002) 747.
[4] T.C. Hunter, N.L. Andon, A. Koller, J.R. Yates III, P.A. Haynes, J. Chromatogr. B 782 (2002) 165.
[5] S.-W. Lee, S.J. Berger, S. Martinovic, L. Pasa-Tolic, G.A. Anderson, Y. Shen, R. Zhao, R.D. Smith, Proc. Natl. Acad. Sci. USA 99 (2002) 5942.
[6] J.A. Loo, C.G. Edmonds, R.D. Smith, Science 248 (1990) 201.
[7] J.L. Stephenson, S.A. McLuckey, G.E. Reid, J.M. Wells, J.L. Bundy, Curr. Opin. Biotechnol. 13 (2002) 57.
[8] G.A. Valaskovic, N.L. Kelleher, F.W. McLafferty, Science 273 (1996) 1199.
[9] W.J. Henzel, W.H. Bourell, J.T. Stults, Anal. Biochem. 187 (1990) 228.

[10] J.R. Yates III, Electrophoresis 19 (1998) 893.
[11] W.J. Henzel, T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley, C. Watanabe, Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 5011.
[12] D.F. Hunt, J.R. Yates III, J. Shabanowitz, S. Winston, C.R. Hauer, Proc. Natl. Acad. Sci. U.S.A. 83 (1986) 6233.
[13] D.J. Pappin, P. Hojrup, A.J. Bleasby, Curr. Biol. 3 (1990) 153.
[14] K.R. Clauser, P.R. Baker, A.L. Burlingame, Anal. Chem. 71 (1999) 2871.
[15] K. Gevaert, J. Vandekerckhove, Electrophoresis 21 (2000) 1145.
[16] J.K. Eng, A.L. McCormack, J.R. Yates III, J. Am. Soc. Mass Spectrom. 5 (1994) 976.
[17] D.J. Pappin, Methods Mol. Biol. 211 (2003) 211.
[18] F. He, C.L. Hendrickson, A.G. Marshall, Anal. Chem. 73 (2001) 647.
[19] W.V. Bienvenut, C. Deon, C. Pasquarello, J.M. Campbell, J.-C. Sanchez, M.L. Vestal, D.F. Hochstrasser, Proteomics 2 (2002) 868.
[20] A. Wattenburg, A.J. Organ, K. Schneider, R. Tyldesley, R. Bordoli, R.H. Bateman, J. Am. Soc. Mass Spectrom. 13 (2002) 772.
[21] S.C. Moyer, R.J. Cotter, Anal. Chem. 74 (2002) 468A.
[22] M.T. Davis, D.C. Stahl, S.A. Hefta, T.D. Lee, Anal. Chem. 67 (1995) 4549.
[23] M.R. Emmett, R.M. Caprioli, J. Am. Soc. Mass Spectrom. 5 (1994) 605.
[24] D.C. Gale, R.D. Smith, Rapid Commun. Mass Spectrom. 7 (1993) 1017.
[25] D.C. Stahl, K.M. Swiderek, M.T. Davis, T.D. Lee, J. Am. Soc. Mass Spectrom. 7 (1996) 532.
[26] D.C. Stahl and T.D. Lee, in: P. James (Ed.), Proteome Research: Mass Spectrometry (Principles and Practice), Springer, New York, 2000, pp. 55–74.
[27] J.M. Carlton, S.V. Angluoli, B.B. Suh, T.W. Kooij, M. Pertea, J.C. Silva, M.D. Ermolaeva, J.E. Allen, J.D. Slengut, H.L. Koo, J.D. Peterson, M. Pop, D.S. Kosack, M.F. Shumway, S.L. Bidwell, S.J. Shallom, S.E. van Aken, S.R. Rledmuller, T.V. Feldblyum, J.K. Cho, J. Quackenbush, M. Sedegah, A. Shoalbi, L.M. Cummings, L. Florens, J.R. Yates III, J.D. Raine, R.E. Sinden, M.A. Harris, D.A. Cunningham, P.R. Preser, L.W. Bergman, A.B. Valdya, L.H. van Lin, C.J. Janse, A.P. Waters, H.O. Smith, O.R. White, S.L. Salzberg, J.C. Venter, C.M. Fraser, S.L. Hoffman, M.J. Gardner, D.J. Carucci, Nature 419 (2002) 512.
[28] L. Florens, M.P. Washburn, J.D. Raine, R.M. Anthony, M. Grainger, J.D. Haynes, J.K. Moch, N. Muster, J.B. Sacci, D.L. Tabb, A.A. Witney, D. Wolters, Y. Wu, M.J. Gardner, A.A. Holder, R.E. Sinden, J.R. Yates III, D.J. Carucci, Nature 419 (2002) 520.
[29] E. Lasonder, Y. Ishihama, J.S. Andersen, A.M.W. Vermunt, A. Pain, R.W. Sauerwein, W.M.C. Eling, N. Hall, A.P. Waters, H.G. Stunnenberg, M. Mann, Nature 419 (2002) 537.
[30] A.R. Dongre, J.L. Jones, A. Somogyi, V.H. Wysocki, J. Am. Chem. Soc. 118 (1996) 8365.
[31] P. Camileri, in: Proceedings of the Capillary Electrophoresis: Theory and Practice, second ed., CRC Press, Boca Raton, FL, 1998.
[32] R.N. Warriner, A.S. Craze, D.E. Games, S.J. Lane, Rapid Commun. Mass Spectrom. 12 (1998) 1143.
[33] A. Premstaller, H. Oberacher, W. Walcher, A.M. Timperio, L. Zolla, J.-P. Chervet, N. Cavusoglu, A. van Dorssler, C.G. Huber, Anal. Chem. 73 (2001) 2390.
[34] L. Li, C.D. Masselon, G.A. Anderson, L. Pasa-Tolic, S.-W. Lee, Y. Shen, R. Zhao, M.S. Lipton, T.P. Conrads, N. Tolic, R.D. Smith, Anal. Chem. 73 (2001) 3312.
[35] C.J. Barinaga, C.G. Edmonds, H.R. Udseth, Rapid Commun. Mass Spectrom. 3 (1989) 160.
[36] V. Katta, S.K. Chowdhury, B.T. Chait, Anal. Chem. 63 (1991) 174.
[37] Y.J. Lee, C.S. Hoaglund-Hyzer, C.A. Srebalus Barnes, A.E. Hilderbrand, S.J. Valentine, D.E. Clemmer, J. Chromatogr. B 782 (2002) 343.
[38] M. Hamdan, P.G. Righetti, Mass Spectrom. Rev. 21 (2002) 287.
[39] A. Chakraborty, F. Reigner, J. Chromatogr. A 949 (2002) 173.
[40] J. Fjeldsted, B. Russ, W. Frazer, A. Hidalgo, M. Flanagan, E. Darland H. Wollnik, in: Proceedings of the 50th ASMS Conference on Mass Spectrometry and Allied Topics, Orlando, FL, 2–6 June 2002.
[41] P. Roepstorff, J. Fohlman, Biomed. Mass Spectrom. 11 (1984) 601.
[42] T.M. Sack, R.L. Lapp, M.L. Gross, B.J. Kimble, Int. J. Mass Spectrom. Ion Proc. 61 (1984) 191.
[43] A. Schlosser, W.D. Lehmann, Proteomics 2 (2002) 524.