Rapidly Switchable Matrix-Assisted Laser Desorption/Ionization and Electrospray Quadrupole-Time-of-Flight Mass Spectrometry for Protein Identification

Andrew N. Krutchinsky, Wenzhu Zhang, and Brian T. Chait
Rockefeller University, New York, New York, USA

We describe a new interface for a prototype quadrupole–quadrupole-time-of-flight (TOF) mass spectrometer (Centaur, Sciex) that allows rapid switching between electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) modes of operation. Instrument performance in both modes is comparable (i.e., resolution ~10,000 FWHM, mass accuracy <10 ppm, sensitivity ~1 fmol) because the ion source is decoupled from the TOF mass analyzer by extensive gas collisions in the quadrupole stages of the instrument. The capacity to obtain side-by-side high quality ESI and MALDI mass spectra from a single proteolytic mixture greatly facilitates the identification of proteins and elucidation of their primary structures. Improved strategies for protein identification result from this ability to measure spectra using both ionization modes in the same instrument and to perform MS/MS on singly charged as well as multiply charged ions. Examples are provided to demonstrate the utility and performance of the modified instrument. (J Am Soc Mass Spectrom 2000, 11, 493–504) © 2000 American Society for Mass Spectrometry

The hybrid quadrupole–quadrupole-time-of-flight (QqTOF) mass analyzer is a powerful tool for biological research [1]. The high accuracy, resolution, and sensitivity of such instruments in both single-stage (MS) and tandem (MS/MS) modes is particularly useful for the rapid, unambiguous identification of small quantities of protein [2, 3].

The original QqTOF instruments were designed to operate with the electrospray ionization (ESI) source, and utilized two approaches for identifying proteins. The first involves separation of in-gel digested proteolytic peptides by microcapillary high performance liquid chromatography (HPLC), followed by on-line analysis of the eluting peptides by rapidly alternating MS and MS/MS analysis [2, 4–7]. The second approach utilizes nanoelectrospray [8] to analyze unfractionated mixtures of proteolytic peptides, wherein a single-stage MS spectrum is obtained followed by MS/MS analysis of selected peptide ions [3, 9–12].

It has been shown previously that a quadrupole-TOF mass analyzer can operate successfully with a MALDI ion source [1, 13, 14]. Because the MALDI ion source is decoupled from the TOF analyzer by gas collisions in the quadrupoles, the performance characteristics (mass accuracy, resolution, and sensitivity) are similar to those obtained with the ESI ion source configuration. However, it is important to note that MALDI mass spectra of peptides are dominated by singly charged ion species in contrast to the multiply charged ions that are characteristic of ESI mass spectra. Thus, the QqTOF fitted with a MALDI source provides the possibility of obtaining singly charged peptide maps and performing MS/MS structural studies on these singly charged ion species.

Here, we describe the design and performance of a rapidly switchable ESI/MALDI ion source for a prototype QqTOF instrument (Centaur, Sciex). We discuss the advantages of having both modes of ion production rapidly available on the same instrument, especially with respect to optimized strategies for protein identification. Examples of successful protein identifications under analytically challenging conditions are provided.

Experimental

New Interface for the QqTOF Mass Spectrometer

The Sciex prototype QqTOF instrument (Centaur) was modified by the addition of an ion source interface that enables us to operate in either MALDI or ESI modes, with the possibility to change quickly between these two modes. Because the detailed description of the mass analyzer can be found elsewhere [1, 3], we will

Address reprint requests to Brian T. Chait, Rockefeller University, 1230 York Avenue, New York, NY 10021. E-mail: chait@rockvax.rockefeller.edu

1044-0305/00/$20.00
PII S1044-0305(00)00114-8

Received December 17, 1999
Revised January 28, 2000
Accepted January 29, 2000

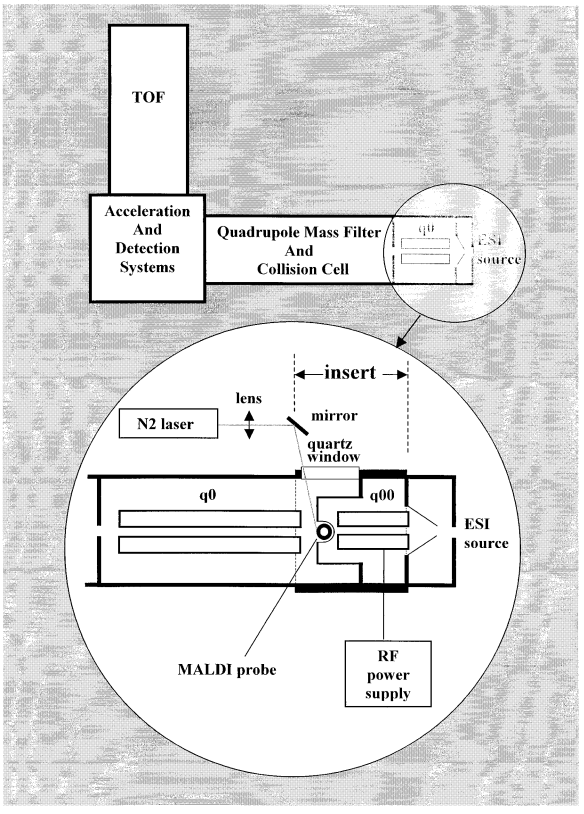

Figure 1. Schematic diagram of the prototype QqTOF mass spectrometer (Centaur, Sciex, Canada). The new interface is shown.

here focus primarily on changes that we have made to incorporate our new interface (Figure 1).

Electrospray Mode

The core of the modification is an insert between the first quadrupole (q0) and the skimmer, where we have added a small quadrupole (q00). This additional quadrupole acts as the first collisional ion guide [15, 16]. The quadrupole is driven by an independent rf power supply, incorporating a 1.6 MHz crystal oscillator-controlled sine wave generator, a rf power amplifier (Model 240L ENI, Rochester, NY), and a transformer wound on a low power loss toroidal core (Model T400-2, Micrometals, CA). Additional details on the circuitry can be found in [17].

When operating in the ESI mode, the MALDI probe is removed (see Figure 1) and ions that are produced in the standard nanoelectrospray ion source (Protana, Denmark) are introduced into the instrument through the standard orifice plate and skimmer into q00, and then through a small gap (0.8 cm) into the original q0-quadrupole ion guide. The offset voltages on q00, q0, and the electrode that separates the two quadrupoles are shown in Figure 2. The remaining operating conditions are similar to those recommended by the manufacturer. Figure 3 provides an example of the performance of the instrument operating in the ESI mode with our standard six component calibration peptide mixture. Only three of the six components were observed in the ESI mass spectrum, presumably because of sample losses during handling and/or signal suppression effects in the nanospray ionization process [18]. Careful measurements performed prior to and after implementation of the new interface revealed that the modification did not result in any loss in sensitivity or resolution in the ESI mode of operation.

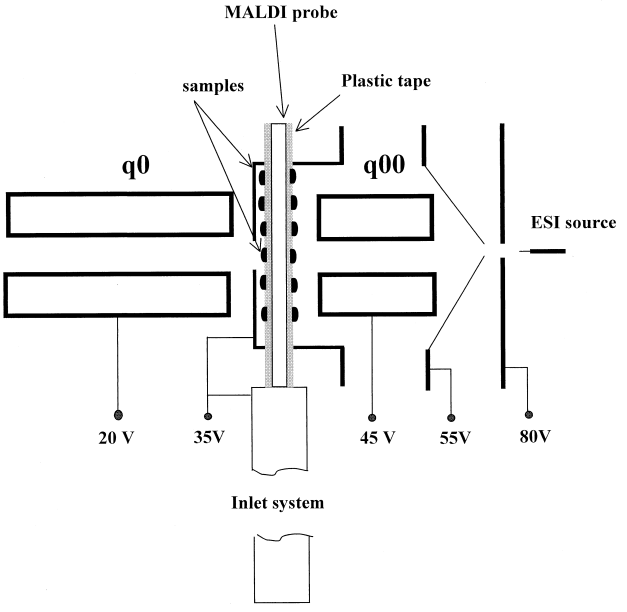

Figure 2. Distribution of voltages on the main elements of the ion optics in the new interface.

MALDI Mode

The insert (Figure 1) has an inlet system, which allows rapid (<30 s) introduction of a MALDI probe. This MALDI probe inserts into the gap in the electrode that separates the q00 and q0 quadrupoles (Figure 2). The cylindrical tip of the probe (diameter 0.5 cm) accommodates ~40 samples on its surface. The beam of a nitrogen laser (Model VSL-337 Laser Science, MA) operating at a repetition rate of 20 Hz is introduced through a quartz window at an angle of ~30° to the surface of the sample. (Recently, we have replaced this laser by a Model VSL-337 ND-S operating at 30 Hz.) The laser beam is focused by a lens (f = 25 cm) to a ~0.2–0.3 mm diameter spot. The power density of the laser radiation in the spot was varied between 10^7 and 10^8 W/cm^2 through fine control of the focal diameter of the laser beam. No discernable increase in the level of metastable decay was observed under these relatively high fluence laser conditions. The position of each sample was manually adjusted at the entrance of q0 by translating and rotating the probe. Both the sample and laser spots were monitored by a video camera. Desorbed ions were introduced directly into q0. No volt-

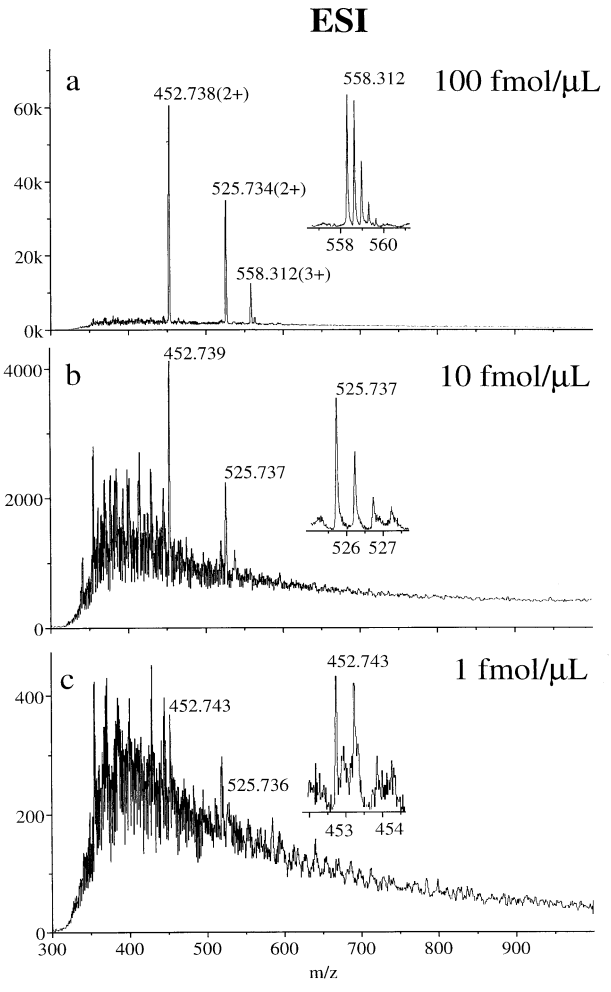

Figure 3. Nanoelectrospray spectra of three different concentrations of a six peptide calibration mixture. The components of the mixture are des-Arg[1]-bradykinin (monoisotopic molecular mass 903.460 Da), [Arg[8]]-vasotocin (1049.453 Da), peptide with sequence CGYGPKKKRKVGG (1376.770 Da), neurotensin (1671.909 Da), ACTH fragment 1–24 (2931.579 Da) and glucagon (3480.614 Da). Spectra acquisition time = 1 min. A two-point calibration was obtained using the singly and doubly charged ions of des-Arg-bradykinin with 2200 V applied to the microchannel plate (MCP) detector. The detection voltage was then lowered to 1900 V and the spectra shown were collected. Lowering the voltage on the MCP favors detection of multiply charged ions [19, 20]. However, note that high MCP voltage (2200 V) was used for collecting the ESI-MS/MS spectra to ensure efficient detection of the singly charged fragment ions.

age changes were required when changing from ESI to MALDI modes (Figure 2).

Sample Preparation

A calibration stock mixture of des-Arg-bradykinin, vasotocin, peptide with sequence CGYGPKKKRKVGG, neurotensin, ACTH 1-24, and glucagon (all obtained from Sigma), was prepared at a concentration 100 fmol/μL per component in water/methanol/acetic acid (1/1/0.1 v/v/v).

Proteins of interest were separated on commercial precast minigels (Novex, CA) or in a few instances, on homemade gels. Bands of interest were excised from the gels and destained according to the manufacturer's protocol for Zn stained gels (Bio-Rad, CA) or using 25 mM ammonium bicarbonate in 50% methanol for Coomassie stained gels. The destained gel pieces were washed in 10% acetic acid/50% MeOH for 6–18 h and finally in pure water or 50 mM ammonium bicarbonate for ~1 h. Destaining and washing were performed at 4 °C. In-gel digestion was carried out as described [21, 22] using either modified or unmodified sequencing grade trypsin (Boehringer Mannheim) for 2–12 h. After digestion, the supernatant was removed and placed in a separate polypropylene tube. Two further extractions were performed (each with ~15–20 μL of 5% formic acid in water for 15 min [23]) and these were combined with the original supernatant. (Currently, we use a simplified one-step extraction procedure wherein we add a higher concentration of formic acid directly to the gel/supernatant.) The resulting proteolytic peptide sample was purified using a Millipore ZipTip pipette. In cases where the amount of protein on the gel was low (0.1–1 pmol), the purification was performed using columns containing 100–300 nL of Porous 20 R2 resin (PerSeptive Biosystems, MA) prepared in Geloader tips (Ultra Micro Pipette, Eppendorf, Germany) as described [24]. Elution of peptides bound to these minicolumns was effected by addition of 1.0–1.5 μL of water/methanol/acetic acid solution (35/60/5 v/v/v). For nanospray, the eluent was guided directly into the nanospray needle. [Prior to sample loading, the internal volume of the nanospray needles (obtained from Protana, Denmark) was washed with 10–20 μL of pure acetonitrile.] For MALDI, 1 μL of sample was mixed with a 1 μL solution of DHB (Aldrich) saturated in water/acetonitrile/TFA (70/30/0.1 v/v/v) at room temperature. Any unused portion of the eluted sample was retained for further analysis either in MALDI or ESI modes, as deemed necessary.

Preparation of Samples for MALDI

Samples for MALDI analysis were prepared by depositing 1 μL of sample solution in DHB matrix on the cylindrical sample probe surface (Figure 2). We tested several different materials as sample support substrates (metal, glass, and a variety of plastics) and found certain plastic surfaces to have superior properties, yielding high signal-to-noise ratios for MALDI. In this regard, "Scotch Magic Tape" (3M, MN) proved to be the best, likely because it is relatively hydrophobic (preventing spreading of the sample solution) and porous (encouraging strong binding of the matrix to the substrate) and because it does not produce a significant background ion signal of its own. Figure 4 shows the MALDI spectrum desorbed from a "magic tape" surface of the same six peptide mixture used to obtain the ESI spectrum shown in Figure 3. In contrast to the nanospray result, the MALDI spectra reproducibly ex-

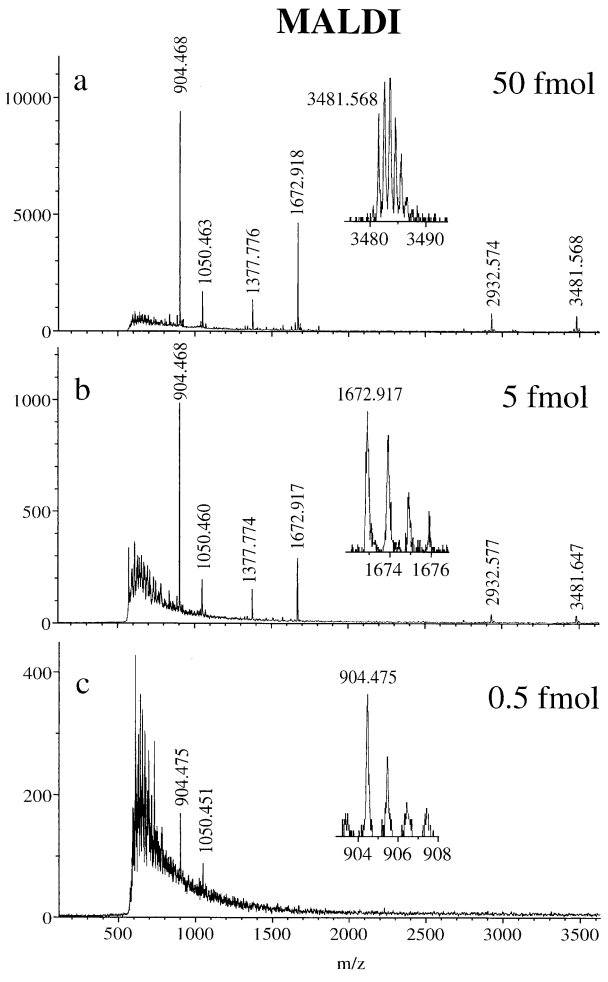

Figure 4. MALDI spectra of the six peptide calibration mixture discussed in the legend to Figure 3. The amount of sample applied to the probe is indicated. Spectra acquisition time = 1 min.

hibited ion signals from all six components of the peptide mixture (for sample amounts ≥5 fmol). No effects of surface charging were apparent during the measurement of the MALDI spectra.

Computer Resources

Data acquisition and computer control of the instrument were performed by commercial software (LC2Tune and TOFMA, Sciex) installed on a Power Macintosh G3 computer. Data were analyzed on the Macintosh computer (TOFMA) or on a Pentium III PC (Dell Dimension, XPS T550). The program "MoverZ" (Proteometrics, NY) installed on the PC allowed rapid processing of the peptide mapping and MS/MS data. The processed data could be automatically transferred to a set of protein identification tools resident at URL http://prowl.rockefeller.edu. These tools include "ProFound" [25], a search engine that identifies proteins by comparison of peptide mapping data with data in a protein/DNA database and "PepFrag" [26], a search engine that utilizes of MS and MS/MS data to identify proteins or EST fragments from various databases.

Calibration

All the spectra were "externally" calibrated using the standard calibration mixture described above. By external calibration, we mean a calibration spectrum acquired prior to obtaining data on the samples of interest. The same standard calibration mixture was used for MALDI and ESI. The calibration obtained in MS mode was used to calibrate both MS and MS/MS spectra, and was observed to hold steady (within <10 ppm) for at least 30 min. The calibration constants obtained using MALDI-MS were indistinguishable from those obtained with ESI-MS, so that they could be used interchangeably.

Results and Discussion

The spectra shown in Figures 3 and 4 demonstrate that the above-described modifications to the QqTOF mass spectrometer allow us to obtain high quality ESI and MALDI mass spectra from as little as 1 fmol of peptide sample at a FWHM resolution of 10,000. The mass accuracy using external calibration is always <20 ppm, and given sufficient counting statistics it is frequently <5 ppm (see below). The resolution and mass accuracy are independent of the mode of ionization. These relatively high performance properties together with the ability to perform MS/MS measurements on both MALDI and ESI ions make the instrument well suited for the solution of a variety of challenging bioanalytical problems.

Mass Accuracy

One facile approach to protein identification involves the combination of MALDI peptide mapping with database searching [27]. The QqTOF mass spectrometer proves to be an excellent choice for protein identification by peptide mapping because the efficacy of this method increases rapidly with the accuracy of the peptide mass measurement [26, 28].

Figure 5 shows a MALDI spectrum of tryptic peptides obtained from in-gel digestion of ~1 pmol of an unknown protein (apparent molecular mass 178 kDa) from *Saccharomyces cerevisiae*. The masses of all monoisotopic peaks above a threshold corresponding to a signal-to-noise ratio ~2 were determined by the program "M/Z" (Version 9.3.2.0) and these masses were automatically fed to the protein search engine "ProFound" (Version 4.6.0). The search was performed on all *S. cerevisiae* proteins contained within the NCBInr database, allowing for one missed tryptic cleavage in any given peptide, partial oxidation of methionine residues, and a mass uncertainty of ±10 ppm. The majority of the larger peaks that do not correspond to known contaminants (trypsin self-digestion, keratin, etc.) are labeled. The top-ranked candidate was identified to be Kem1p (probability ~1.00), a nuclease with 5′–3′ exonuclease activity for single-stranded RNA and

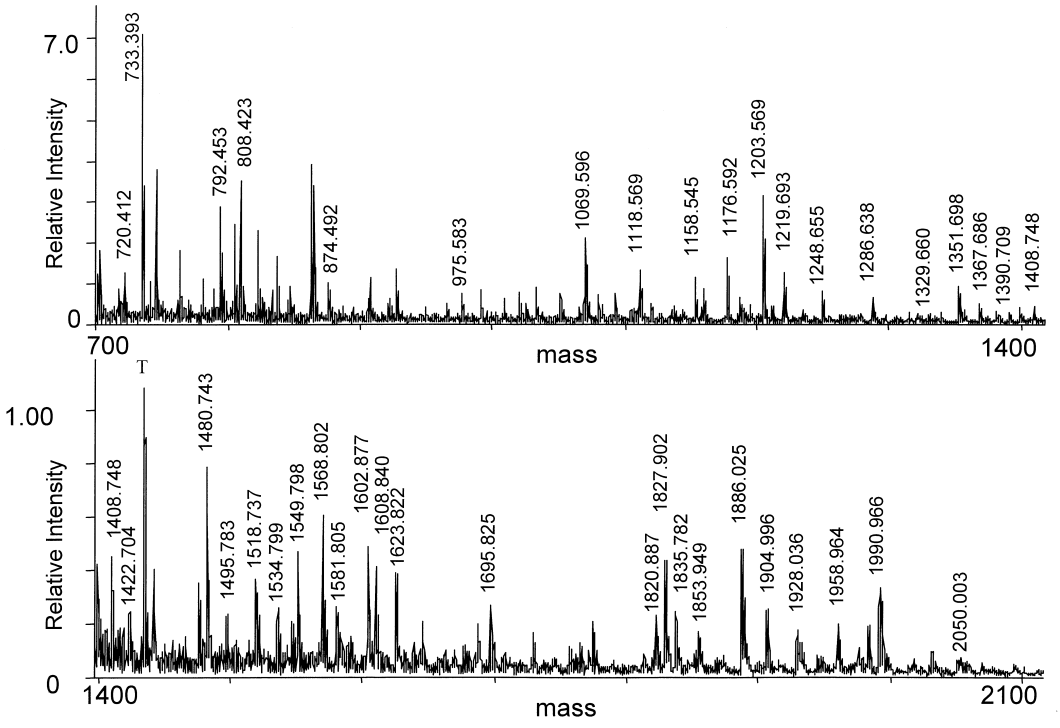

Figure 5. MALDI-MS tryptic map of a ~178 kDa protein from *S. cerevisiae*. T refers to a trypsin self digestion product.

DNA, whose primary known role is to degrade decapped mRNA [29]. The next highest ranked candidate, Blm3p, was observed with a probability fully 106 orders of magnitude lower than that observed for Kem1p. This very high level of discrimination, which provides identification with extremely high confidence, is often observed when the quality of the MS data is good and the coverage of the protein by the tryptic map is adequately high (~30% in this case). Comparison of the measured and the computed tryptic peptide masses for the top-ranked protein (Table 1) shows that 75% of the measured peak masses fall within 5 ppm of the computed values and the remaining 25% within 10 ppm. This level of mass accuracy is typical of MALDI spectra obtained using external calibration with our set of standard peptides (Figure 2). Such external calibrations were observed to remain stable (i.e., systematic errors <10 ppm) for more than 30 min during acquisition of the MALDI spectra.

Sensitivity

In many practical applications where one wishes to identify a protein, the amount of sample available is severely limited. Figure 6a shows the MALDI-MS tryptic map obtained from an unknown 70 kDa SDS-PAGE protein band isolated from *Streptococcus pneumonia* for which <5 fmol of sample was available subsequent to extraction from the gel. The amount was estimated by comparison of the signal response with that of a series of peptide standards.

14 ion peaks, which were distinct from the background observed with a blank gel control, were selected for the tryptic mapping database search. The results are shown in Table 2 together with the search parameters. The top-ranked protein candidate, choline binding protein A (*S. pneumoniae*), had a probability score = 0.99, whereas the second-ranked protein candidate had a probability score = 0.008. This level of discrimination may not be sufficient for certain purposes. For example, were the identification to be used as the basis for initiating a costly, large-scale biological investigation, we would normally insist upon an additional check of the correctness of the identification. In the present instrument, this check can be provided by collision induced dissociation (CID) of selected singly charged MALDI peptide ions. As an example, we show the CID spectrum of the weak ion at *m/z* 1406.746, obtained over a data acquisition period of 5 min during which a total of 2124 ions were detected (Figure 6b). Although this is a modest number of ions, it proved sufficient to accumulate 5–30 ions for each fragment peak labeled in the spectrum. A relatively narrow mass window (2 Da) was used for parent ion selection to reduce the contribution of the chemical noise. The resulting signal-to-noise ratio was sufficiently high to permit identification of fragment peaks with as few as 3 counts/fragment ion. The labeled peaks in Figure 6b were used to search the NCBInr database with the protein identification tool PepFrag [26]. Of the 13 fragment ion masses supplied to the search routine (together with the mass of the parent ion measured with high accuracy in the single-stage MS

Table 1. Tryptic peptides from the MALDI-MS map shown in Figure 5

Measured mass (Da)	Computed mass (Da)	Error (Da)	Residues start–to	Missed cuts	Peptide sequence[a]
720.412	720.417	−0.005	841–846	0	LFNSLK
733.393	733.391	+0.002	798–802	0	WPYLR
792.453	792.449	+0.003	1155–1161	0	SSFALLR
808.423	808.426	−0.004	181–186	0	IM*NFIR
874.491	874.491	−0.000	722–729	0	YGLLPNAK
975.583	975.586	−0.003	1014–1022	0	HINVGIPVK
1069.596	1069.592	+0.004	1284–1292	0	AHDLLNFIK
1118.569	1118.564	+0.005	981–989	0	YYPSYIVSK
1158.545	1158.541	+0.004	1130–1139	0	YAANIEGHER
1176.592	1176.592	−0.000	146–154	0	NLQYFIHDK
1203.569	1203.571	−0.002	1056–1064	0	QTFPDFFFR
1219.693	1219.696	−0.003	304–314	0	GAFPVLLQTFK
1248.654	1248.653	+0.001	652–661	0	LIEAM*QPYLR
1286.638	1286.636	+0.002	1130–1140	1	YAANIEGHERK
1286.638	1286.646	−0.008	363–373	0	QLENISLEGER
1329.660	1329.659	+0.001	1118–1129	0	TSIAAVEDHIM*K
1351.698	1351.706	−0.009	990–1000	0	NM*HLHPLFLSK
1351.698	1351.695	+0.002	80–91	0	IFYMAIDGVAPR
1353.695	1343.682	+0.013	970–980	1	ERLQMDHQAVK
1367.686	1367.690	−0.004	80–91	0	IFYM*AIDGVAPR
1390.709	1390.708	+0.001	960–971	1	FLDSEPTIGKER
1408.748	1408.744	+0.004	1105–1117	0	FIAVSLESDSLTK
1422.704	1422.710	−0.005	777–788	1	TNNVTLEDFSKR
1480.743	1480.755	−0.013	827–839	0	FGFITKPAETQDK
1495.783	1495.785	−0.002	79–91	1	KIFYM*AIDGVAPR
1518.737	1518.738	−0.001	423–434	1	DLDM*KDHLEFLK
1534.799	1534.787	+0.012	709–721	0	EYITIPLDSSEIR
1549.798	1549.813	−0.015	1040–1052	0	GWEYSNLTLNLLK
1568.802	1568.808	−0.006	915–928	0	GPIPIEEEFPLNSK
1581.805	1581.806	−0.001	764–776	0	QQSM*VLQITDIYK
1602.877	1602.872	+0.005	1223–1237	0	GLGLDASFLLNITNR
1608.840	1608.850	−0.010	827–840	1	FGFITKPAETQDKK
1608.840	1608.850	−0.010	826–839	1	KFGFITKPAETQDK
1623.822	1623.829	−0.007	435–448	0	EFAFDLGLFITHSK
1695.825	1695.821	+0.004	315–329	0	EALLHTDGYINEHGK
1820.887	1820.905	−0.006	146–160	1	NLQYFIHDKISNDSK
1827.902	1827.904	−0.002	676–690	0	DLIYSFNPQVDNLYK
1835.782	1835.766	+0.016	263–276	0	EIADEM*QFEYNFER
1853.949	1853.951	−0.002	913–928	1	ERGPIPIEEEFPLNSK
1886.025	1886.033	−0.008	861–878	0	IGPM*EAIATVFPVTGLVR
1904.996	1904.997	−0.001	406–422	0	LSPDLPDEEIPTLELPK
1928.036	1928.043	−0.007	390–405	0	LIGSIKPWLM*EQLQEK
1958.964	1958.983	−0.018	1068–1085	0	VGNDIPVLEDLFPDTSTK
1990.965	1990.974	−0.009	163–180	0	EVQIIFSGHEVPGEGEHK
2050.003	2050.010	−0.007	485–501	1	YQNAIIVEDKEELETEK

[a]M* denotes methionine sulfoxide.

mode), 12 were uniquely identified as y- or b-series fragments of a peptide from the choline binding protein A (lower portion of Table 2). Even more conclusive confirmation of the identity of band as the choline binding protein A was obtained through subsequent MS/MS of a second peptide ion peak at *m/z* 988.540 (not shown). This example demonstrates that a highly confident identification can be obtained even when only a few fmol of digest is available for analysis.

The fragmentation pattern observed in the present instrument is strongly influenced by the energy of collisions of the parent ions with the gas molecules in the collision cell. To obtain the optimal fragmentation of singly charged MALDI peptide ions for the purpose of protein identification, we set the collision energy according to the empirical rule $E_{CID} \sim M/20$ eV, where E_{CID} is the CID energy and M is the peptide mass. The collisional gas was Ar and the pressure in the collision cell was a few mtorr. This choice of parameters produces relatively soft fragmentation, dominated by y- and/or b-type fragments. Such predictable fragmentation allows us to specify the expected ion series, and thus greatly facilitates automatic protein identification based on the MS/MS spectral data [26].

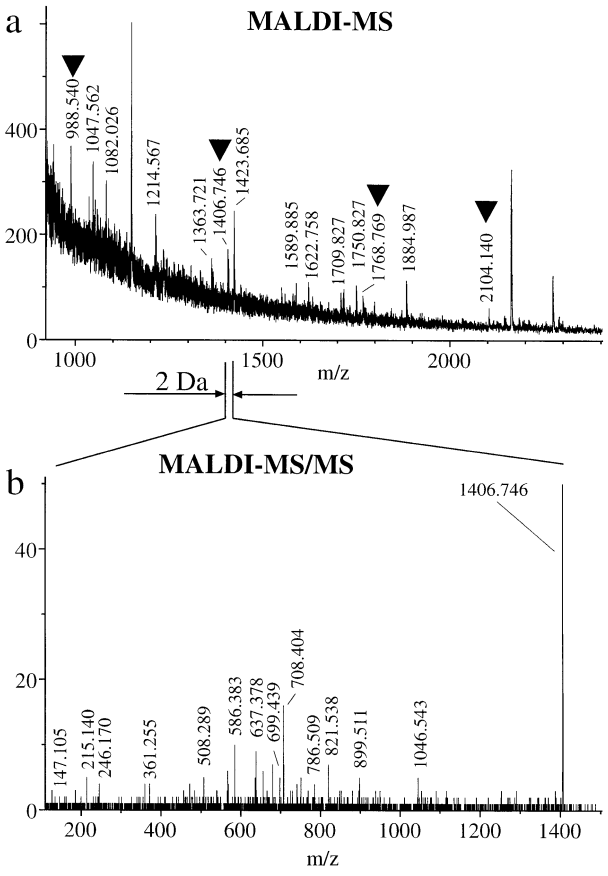

Figure 6. (**a**) MALDI-MS tryptic map obtained from ≤5 fmol of a ~70 kDa protein from *S. pneumonia*. (**b**) MALDI-MS/MS spectrum of the singly charged *m/z* 1406.746 ion from the tryptic map.

MS/MS

In many cases, MS/MS proves complementary to single-stage MS peptide mapping. Thus, MS/MS can definitively confirm the presence of proteins detected in a mapping experiment or detect additional proteins not detected in the original mapping experiment. Figure 7 shows the tryptic map of a mixture of *S. cerevisiae* proteins extracted from a single ~80 kDa SDS-PAGE gel band. The peptide mapping search engine ProFound detected the presence of four different proteins in the tryptic peptide mixture (Figure 7). The automatic procedure that we used to identify protein components of mixtures by MS peptide mapping is described elsewhere [25]. Because the difference between the probabilities for the top- and second-ranked protein mixture was only 3 orders of magnitude, we used MS/MS of selected individual peptides to provide an independent test for the presence/absence of the four putative proteins. Thus, a tryptic ion peak was chosen for MS/MS evaluation from each of the four probable candidates as indicated in Figure 7. Figure 8 shows the four resulting MALDI MS/MS spectra, each of which was collected over a period of 1–2 min (i.e., sufficient time to accumulate 5–1000 ions for each fragment peak). The results of database searches (NCBInr fungi) with this MS/MS data are provided in Table 3. The searches were performed with the algorithm PepFrag [26] using the peaks labeled in the spectra, allowing mass errors of ±0.1 Da for the fragment ions and ±0.03 Da for the parent ions. In all such searches, we made no attempt to assign fragment ions prior to the search. Rather, we set up the PepFrag routine to automatically search for exclusively y- and/or b-type fragments. This simple strategy

Table 2. (a) MS tryptic mapping data search using ProFound[a] and (b) MS/MS data search using PepFrag[b]. Both searches identified choline binding protein A (*Streptococcus pneumoniae*) mass = 75 kDa

(a) Measured mass (Da)	Computed mass (Da)	Error (Da)	Residues start–to	Missed cuts	Peptide sequence
987.532	987.534	−0.002	164–172	1	NEGTIKQAK
987.532	987.534	−0.002	185–192	1	LENIKTDR
987.532	987.534	−0.002	347–354	1	LENIKTDR
1405.738	1405.733	0.005	135–146	0	TLELEIAEFDVK
1767.761	1767.754	0.007	403–417	1	AEKTDDQQAEEDYAR
2103.132	2103.131	0.001	426–444	1	LTQQQPPKTEKPAQPSTPK

. . . TNTYK<u>TLELEIAEFDVK</u>VKEAE . . .
Residues: 135–146

(b) Measured value (*m/z*)	Ion type	Computed value (*m/z*)	Measured value (*m/z*)	Ion type	Computed value (*m/z*)
147.10	**y1**	147.11	699.44	**b6**	699.39
215.14	**b2**	215.14	708.40	**y6**	708.36
246.17	**y2**	246.18	786.51		
361.25	**y3**	361.21	821.54	**y7**	821.44
508.29	**y4**	508.28	899.51	**b8**	899.47
586.38	**b5**	586.31	1046.54	**b9**	1046.54
637.38	**y5**	637.32			

[a]Database: NCBInr: Bacteria; mass tolerance 10 ppm; coverage of protein sequence: 10%.
[b]Database: NCBInr: Bacteria; *m/z* of parent peptide: 1406.746 ± 0.03, charge state = 1+; expected types of fragment ions: b, y; error: ±0.1 *m/z*.

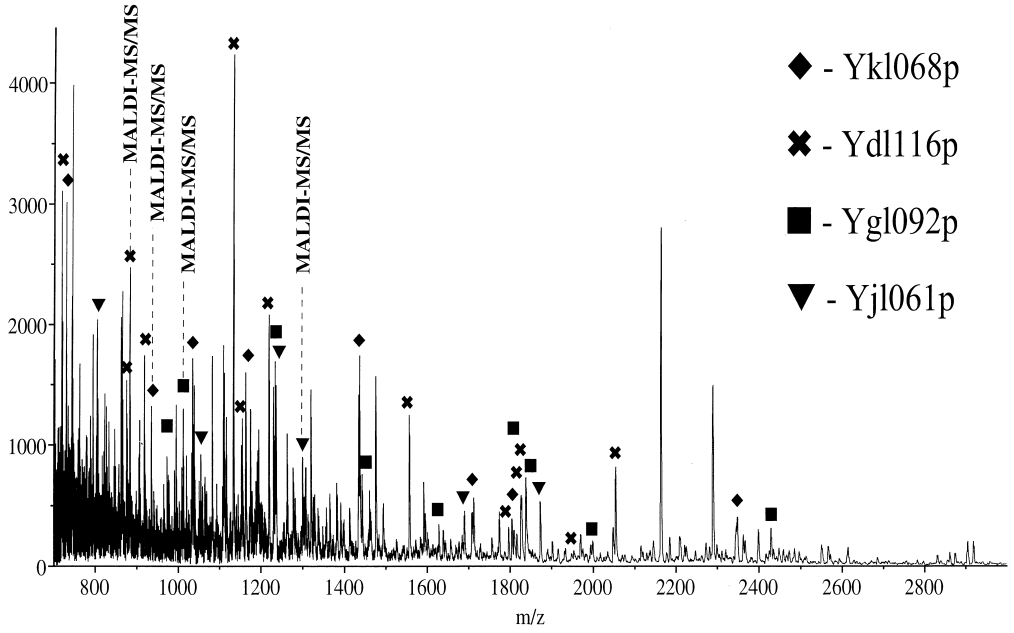

Figure 7. MALDI-MS tryptic map of a mixture of *S. cerevisiae* proteins contained within a single ~80 kDa band. The presence of four proteins in the band was detected by ProFound. Peaks selected for MS/MS (spectra shown in Figure 8) are indicated.

proves fast and effective. Each search identified a unique protein when the number of matches between the experimental and computed *m/z* values were maximized, confirming the proteins that were identified by MS tryptic peptide mapping. Thus, the MS/MS data provided greatly increased confidence of the presence of the four previously identified proteins (Figure 7).

Combined ESI and MALDI Analysis

At present, identification of human proteins can be challenging because a complete database of human proteins is not yet available. When the protein to be identified is represented in the database, the MS-mapping procedure proves rapid and effective. Conversely, when the protein to be identified is not in the database, it is necessary to resort to an alternative strategy in which MS/MS data from selected proteolytic peptide are used to search the expressed sequence tag (EST) database [26, 30]. This identification strategy is more laborious and time consuming than peptide mapping because the search is performed on a large number of EST sequences (at present ~2.7 million human, mouse, and rat ESTs multiplied by the 6 possible reading frames), because the EST sequence error rate is high, and because the coverage by overlapping ESTs of any given cDNA is frequently incomplete. The success of this procedure depends to a large extent on the number of different peptides from the unknown protein that yields informative MS/MS data. The larger the number, the more likely we are to find a matching EST. In this context, the combined use of ESI and MALDI MS/MS proved very useful. The high accuracy of the QqTOF peptide mass spectral data also improves the chance of finding unique EST matches. Finally, the success of the procedure depends on the possibility of accurately stitching together overlapping ESTs to produce a sufficiently long stretch of contiguous cDNA sequence for use in homology searches or cDNA cloning.

Here we present an example of the identification of an unknown human protein (which we designate p42) through combined use of the ESI- and MALDI-MS/MS modes, demonstrating the utility of the present ion source combination within a single high-performance instrument. The MALDI spectrum of tryptic peptides extracted from a ~42 kDa gel band purified from human HeLa cell nuclear extract is shown in Figure 9a. Although a peptide mapping search using ProFound revealed that several of the less intense peaks arose from heterogeneous nuclear ribonucleoproteins C1/C2 and keratin (the latter a ubiquitous contaminant), the higher intensity peaks above *m/z* 1000 were not assigned by the search. To identify the origin of these unassigned peaks, we performed MS/MS experiments on selected singly charged MALDI ion species. We also measured the spectrum of the same mixture of tryptic peptides using ESI, from which we collected additional MS/MS spectra of multiply charged ion species.

MS/MS of the MALDI-MS peak at *m/z* 943.573 (Figure 9a) confirmed the peptide mapping identification of C1/C2 heterogeneous nuclear ribonucleoprotein (not shown). The MS/MS spectrum of the intense unassigned peak at *m/z* 1335.712 (Figure 10a) yielded a series of fragment masses that when supplied to the PepFrag search routine did not yield any protein or EST candidates. Nevertheless, the quality of the fragmenta-

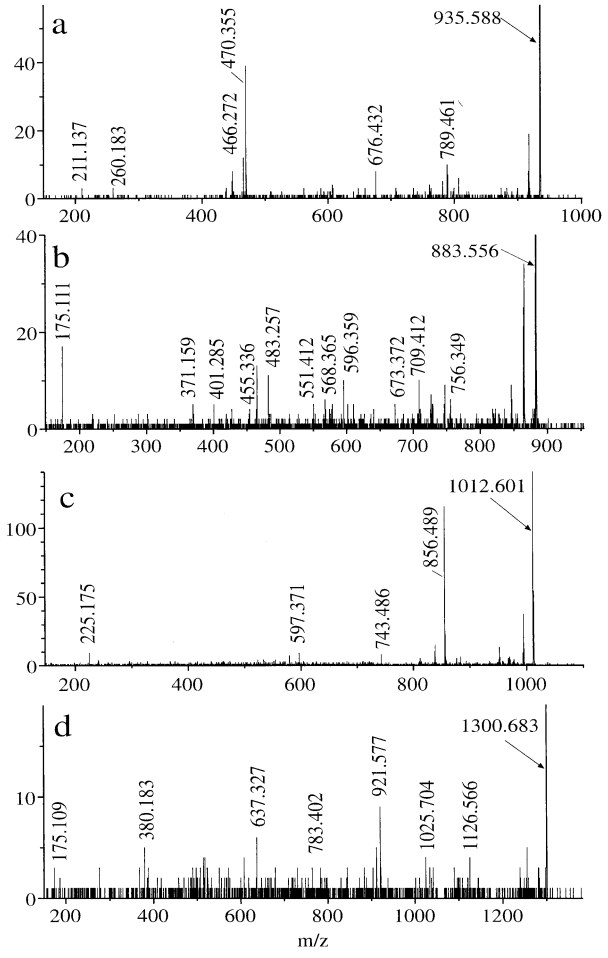

Figure 8. MALDI-MS/MS spectra of selected parent ions from the tryptic map shown in Figure 7. The indicated peaks were supplied to the search tool PepFrag (search results provided in Table 3).

Table 3. PepFrag search results using MS/MS data in Figure 6

a. Ykl068p (*Saccharomyces cervisiae*)
mass = 99927.7 Da
NITHPLLK = 935.568 (*m/z*)

211.14	470.35 → **y4** = 470.33
260.18 → **y2** = 260.20	676.43 → **b6** = 676.38
466.27 → **b4** = 466.24	789.46 → **b7** = 789.46

b. Ydl116p (*Saccharomyces cervisiae*)
mass = 83582.9 Da
LPASNILR = 883.537 (*m/z*)

175.11 → **y1** = 175.12	568.36
371.16	596.36 → **b6** = 596.34
401.28 → **y3** = 401.29	673.37 → **y6** = 673.40
455.34	709.41 → **b7** = 709.42
483.26 → **b5** = 483.26	756.35
551.41	

c. Ygl092p (*Saccharomyces cervisiae*)
mass = 145570.9 Da
RLPTELQR = 1012.590 (*m/z*)

225.17	856.49 → **y7** = 856.49
597.37 → **b5** = 597.34	
743.49 → **y6** = 743.40	

d. Yjl061p (*Saccharomyces cervisiae*)
mass = 82033.9 Da
IHEQQFELTR = 1300.665 (*m/z*)

175.11 → **y1** = 175.12	921.58 → **y7** = 921.48
380.18 → **b3** = 380.19	1025.70
637.33	1126.57 → **b9** = 1126.55
783.40 → **b6** = 783.38	

tion data was sufficient to determine a 9-residue stretch of amino acid sequence (VSE[I/L]EAA[K/Q]K). This partial sequence was used to search the protein and EST databases with the sequence search routine ProteinInfo [31]. Although no match was found among known proteins, an EST was identified with an identical partial sequence but with an N-terminal extension (underlined sequence: MATFVSELEAAKK). Assuming that the terminal residue corresponds to the initiator Met, that this Met is removed in the mature p42 protein, and that the adjacent Ala residue is acetylated (as is often the case for terminal Ala residues [32]), we calculate a mass of 1334.708 Da for this hypothetical peptide in close agreement with the measured mass of 1334.704 Da of the peptide. This data provides convincing evidence that we have determined the N-terminal sequence of the unknown protein. Additional evidence for the accuracy of our assignment was provided by comparison of the fragmentation of the *m/z* 1207.625 ion, having a mass 128.087 Da lower than the putative *m/z* 1335.712 N-terminal ion. These two peptides yielded closely similar fragmentation patterns, indicating their common sequence (with the exception of the C-terminal Lys residue.)

Final confirmation of our assignment was obtained by MS analysis of the sample using ESI-MS. Comparison of Figure 9a with 9b reveals that many of the same proteolytic peptides are represented in both the ESI and MALDI spectra (Table 4) although the charge states and relative abundances of ions arising from a given peptide tend to be different. Thus, for example, the putative N-terminal peptide, which produced a singly charged ion at *m/z* 1335.712 by MALDI-MS, gave rise to the corresponding doubly charged species at *m/z* 668.360 in the ESI-MS map (Figure 9b). MS/MS of this doubly charged ESI ion produces a spectrum very different from that obtained from the singly charged MALDI ion (compare Figures 10a, b). The stretch of sequence that can be readily deduced from this ESI-MS/MS spectrum is considerably shorter than that deduced from the MALDI-MS/MS data, but provides complementary information. Thus, although it was not possible for us to identify a unique EST from this short stretch of deduced sequence, the spectrum was useful in that it confirmed the four residue stretch at the N-terminus that was not determined in the MALDI-MS/MS measurement.

Because individual ESTs generally only provide a small portion of a given protein's sequence, it is usually necessary to splice together several different ESTs to

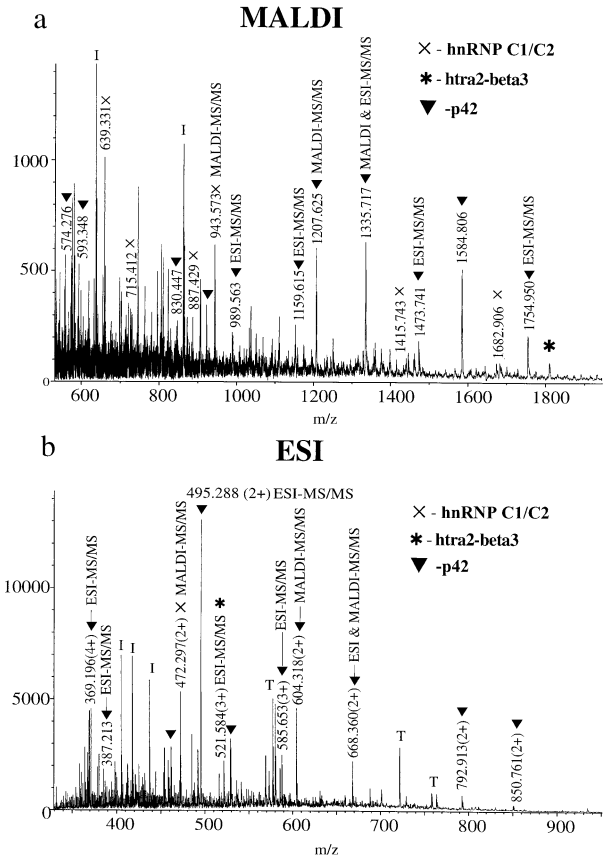

Figure 9. (**a**) MALDI-MS tryptic map of proteins present in a 42 kDa gel band obtained from human HeLa cell nuclear extract. (**b**) ESI-MS tryptic map of the same sample. The presence and identities of two known proteins were established: (i) heterogeneous nuclear ribonucleoproteins C1/C2 (hnRNP C1/C2) and (ii) serine/arginine-rich pre-mRNA splicing factor, htra2-β3 [34]. A third protein (p42) was detected and its probable sequence determined from the dbEST database by a combination of MALDI-MS/MS and ESI-MS/MS (see the text). I designates unknown impurities and T trypsin self-digestion products.

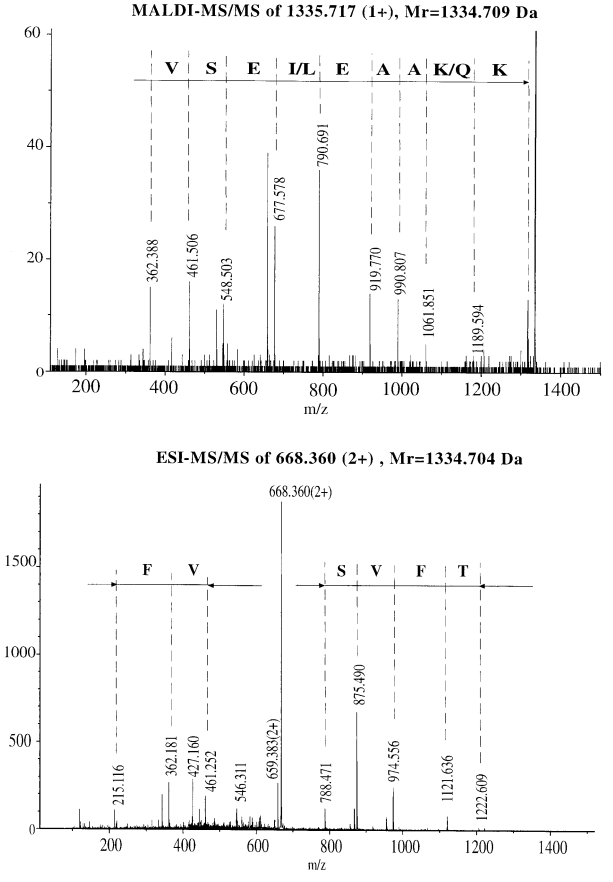

Figure 10. (**a**) MALDI-MS/MS of the singly charged tryptic fragment at *m/z* 1335.717. The observed ladder of b-type fragments defines a 9-residue stretch of the amino acid sequence, i.e., . . . VSE[I/L]EAA[Q/K]K. (**b**) ESI-MS/MS of the doubly charged ion at *m/z* 668.360, arising from the same tryptic fragment as in (**a**). A short stretch of sequence can be deduced from the ladder of y-type fragments, i.e., . . . TFVS, which is complementary to the sequence deduced from the MALDI-MS/MS spectrum.

specify the protein sequence. In the present example, we acquired several additional ESI-MS/MS spectra (as indicated in Figure 9) to assist us in the task of EST assembly. We observed that fragmentation of peptide ions having charge states 2+, 3+, 4+ yielded data that proved accessible to rapid, automatic EST searches via PepFrag. Such automatic searches require only that the charge state of the parent ion be specified together with the expected type of fragmentation (here, b- and y-type ions). The high resolution of the present instrument allows unambiguous specification of the charge state. Figure 11a provides an example of the fragmentation of a triply charged ion (*m/z* 585.653) from the ESI-MS map shown in Figure 9b. This spectrum yields information on 13 out of the 15 residues in the peptide, conclusively identifying a second EST from the gene coding for p42. Additional ESTs corresponding to the p42 protein were identified by ESI-MS/MS (data not shown).

The ESTs found by the above procedure were assembled into a contiguous stretch of sequence using the mass spectrometric data to guide correct assembly. In addition to the MS/MS data, the peptide mapping data (Figure 9) was compared to a theoretical digest of this assembled EST sequence to search for additional matches. The assembled sequence of p42 is shown in Figure 12 with the portions identified by MS indicated. Mass spectrometric measurement confirmed ~30% of this sequence. This use of mass spectrometry [33] to assist in accurate assembly of ESTs should prove extremely useful for correcting the many errors prevalent in extant EST assemblages (e.g., TIGR Unique Gene Indices, www.tigr.org).

The high quality of the data available from the present instrument enables us to readily observe details such as phosphorylation. Thus for example, the low intensity ESI-MS triply charged ion peak at *m/z* 521.584 obtained from the p42 band (Figure 9a) yielded an MS/MS data (Figure 11b) that did not produce a hit using PepFrag. However, inspection of the spectrum revealed two doubly charged peaks (*m/z* 655.297 and 606.306) spaced apart by 97.982 Da, indicative of the

Table 4. Comparison of p42 tryptic peptides in the ESI and MALDI-MS maps

ESI				MALDI		
	m/z	(*n*+)	Mass	Mass	*m/z*	
	461.748	(2+)	**921.480**	**921.486**	922.494	
ESI-MS/MS	495.288	(2+)	**988.560**	**988.555**	989.563	
ESI-MS/MS	387.213	(3+)	**1158.615**	**1158.607**	1159.615	
ESI-MS/MS	580.307	(2+)	**1158.598**	**1158.607**	1159.615	
	604.318	(2+)	**1206.620**	**1206.617**	1207.625	MALDI-MS/MS
ESI-MS/MS	668.360	(2+)	**1334.704**	**1334.709**	1335.717	MALDI-MS/MS
ESI-MS/MS	369.196	(4+)	**1472.731**	**1472.733**	1473.741	
	528.943	(3+)	**1583.806**	**1583.798**	1584.806	
	792.913	(2+)	**1583.810**	**1583.798**	1584.806	
ESI-MS/MS	585.653	(3+)	**1753.942**	**1753.943**	1754.950	

elimination of the elements of H_3PO_4 (calculated mass 97.976 Da) from a phosphopeptide ion fragment. PepFrag can utilize data obtained from phosphopeptides to identify proteins/ESTs provided that we specify that the peptide is phosphorylated. Using this assumption, PepFrag determined that the peptide originated from human transformer-2-beta isoform 3 [34], and that the specific phosphorylation site is Thr-4 rather than Thr-6 or Tyr-10 (Figure 11b).

In summary, the combined MALDI/ESI analysis provided an extraordinarily rich source of information concerning the proteins present in the 42 kDa gel band. This information included the identity of two known proteins, the probable sequence of a third previously unknown protein (p42), the amino-terminus of p42, and posttranslational modifications in two of the three proteins.

Conclusions

We have designed and implemented a new interface for a prototype QqTOF mass spectrometer that allows rapid (<30 s) switching between ESI and MALDI modes of operation. Instrument performance (resolution ~10,000 FWHM, mass accuracy <10 ppm, sensitivity ~1 fmol) in both modes is comparable because the ion source is decoupled from the TOF mass analyzer by extensive gas collisions in the quadrupole stages. We demonstrate that this facility to obtain both ESI and MALDI spectra from a single proteolytic mixture can greatly assist challenging biological analyses. Examina-

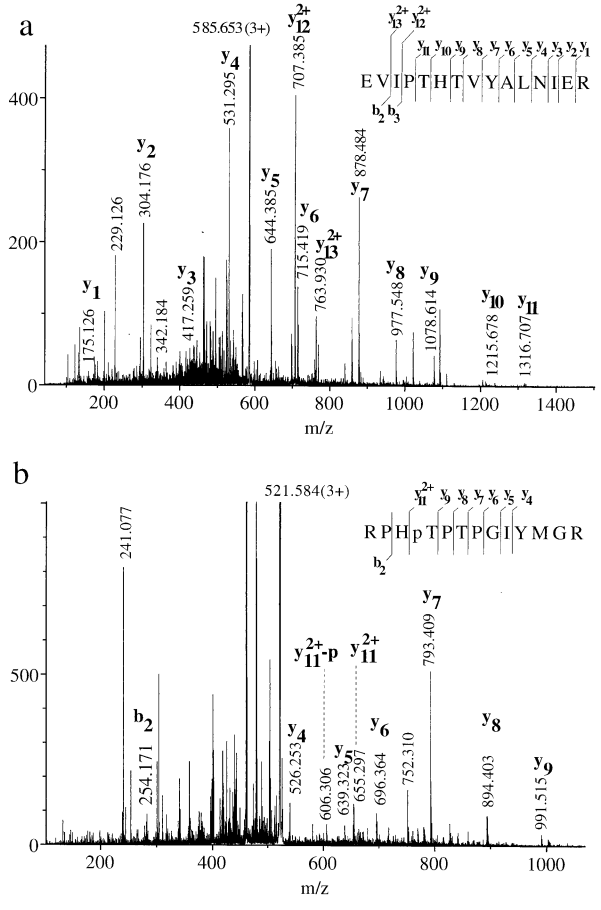

Figure 11. (**a**) ESI-MS/MS spectrum of a triply charged tryptic fragment (*m/z* 585.653) from p42. Fully, 13 out of 15 residues in the peptide were identified from the y-fragment ladder. (**b**) ESI-MS/MS spectrum of the triply charged fragment at *m/z* 521.584. The mass difference of 97.982 between the two doubly charged fragments at *m/z* 655.297 and 606.306 indicates that the peptide is phosphorylated. PepFrag identified the origin of this phosphorylated peptide as htra2-beta3 [34].

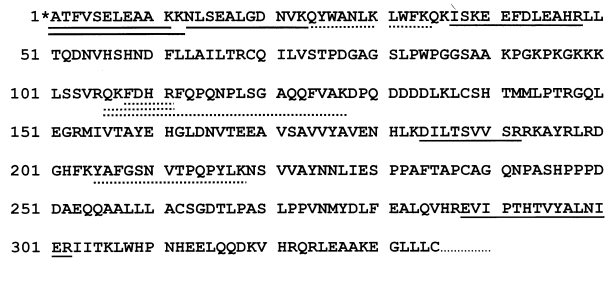

Figure 12. Probable sequence of the protein designated p42. Asterisk designates acetylation of the terminal alanine residue. Solid lines correspond to the fragments determined by MALDI-MS/MS and ESI-MS/MS. Dashed lines correspond to tryptic fragments derived from the probable protein sequence and observed in the MS tryptic maps. The C-terminus of the protein remains unidentified. Of the sequence shown, 106 residues were confirmed by MS or MS/MS.

tion of more than 100 protein gel bands by proteolysis followed by mass spectrometry has convinced us of the considerable added value of the MALDI-MS mode of operation in that the resulting singly charged peptide maps are particularly straightforward to analyze. Furthermore, MS/MS spectra of these $(M + H)^+$ species can be performed subsequent to the rapid mapping analysis, and continued until either a definitive answer is obtained or until we run out of sample (after typically 5–10 MS/MS analyses). Advantages of the MALDI mode of operation include the tolerance of MALDI for impurities, buffers, and additives; the long-term stability of samples on the MALDI probe (allowing MS/MS analysis days/weeks subsequent to the MS mapping analysis); and the ease of sample production (i.e., no HPLC required and no manipulations of delicate nanospray needles required). A primary advantage of the ESI mode of operation is its ability to produce multiply charged peptide ions, which frequently yield rich and informative MS/MS spectra. We envisage that optimized strategies for combining the MALDI and ESI modes of operation will prove particularly effective for identifying proteins and elucidating their primary structures.

Acknowledgments

This work is supported by the National Institute of Health (Grant RR00862). We thank Sciex, Canada for the generous gift of the prototype quadrupole-time-of-flight instrument (Centaur). We also thank Igor Chernushevich, Bruce Thomson, and Ron Bonner of Sciex for their excellent technical support and advice, Agneta Tjernberg for valuable contributions, and ProteoMetrics, LLC, New York for the continuous development of software products that were used in the present work. We thank our many colleagues at Rockefeller University for supplying the biological samples.

References

1. Chernushevich, I. V.; Ens, W.; Sanding, K. G. *Anal. Chem.* **1999,** *71*, 452A–461A.
2. Morris, H. R.; Paxton, T.; Dell, A.; Langhorne, J.; Berg, M.; Bordoli, R. S.; Hoyes, J.; Bateman, R. H. *Rapid Comun. Mass Spectrom.* **1996,** *10*, 889–896.
3. Shevchenko, A.; Chernushevich, I.; Ens, W.; Standing, K. G.; Thomson, B.; Wilm, M.; Mann, M. *Rapid Commun Mass Spectrom.* **1997,** *11*, 1015–1024.
4. Morris, H. R.; Paxton, T.; Panico, M.; McDowell, R.; Dell, A. J. *Protein Chem.* **1997,** *16*, 469–479.
5. Annan, R. S.; Martin, L. B.; Huddleston, M. E.; Carr, S. E. *Proceedings of the 46th ASMS Conference on Mass Spectrometry and Allied Topics*, 1998; p 628.
6. Borchers, C.; Parker, C. E.; Deterding, L. J.; Tomer, K. B. *J. Chrom.* **1999,** *854* (1–2), 119–130.
7. Figeys, D.; Lock, C.; Taylor, L.; Aebersold, R. *Rapid Comun. Mass Spectrom.* **1998,** *12*, 1435–1444.
8. Wilm, M. S.; Mann, M. *Int. J. Mass Spectrom. Ion Processes* **1994,** *136*, 167–180.
9. Katahira, J.; Strässer, K.; Podtelejnikov, A.; Mann, M.; Jung, U. J.; Hurt, E. *EMBO J.* **1999,** *18*, 2593–2609.
10. Schrotz-King, P.; Wilm, M.; Andersen, J. S.; Ashmann, K.; Podtelejnikov, A. V.; Bachi, A.; King, A.; Mann, M. *J. Recept. Signal Trusduct. Res.* **1999,** *19*, 659–672.
11. Huang, L.; Shen, M.; Chernushevich, I.; Burlingame, A. L.; Wang, C. C.; Robertson, C. D. *Mol. Biochem. Parasitol.* **1999,** *102*, 211–223.
12. Willems, A. R.; Goh, T.; Taylor, L.; Chernushevich, I.; Shevchenko, A.; Tyers, M. *Philos. Trans. R. Soc. London B Biol. Sci.* **1999,** *354*, 1533–1550.
13. Krutchinsky, A. N.; Loboda, A. V.; Spicer, V. L.; Ens, W.; Standing, K. G. *Rapid Commun. Mass Spectrom.* **1998,** *12*, 508–518.
14. Krutchinsky, A. N.; Loboda, A. V.; Brominski, M.; Standing, K. G.; Ens, W. *Proceedings of the 46th ASMS Conference on Mass Spectrometry and Allied Topics*, Dallas, TX, 1999; p 794.
15. Gerlich, D. In *State-Selected and State-to-State Ion Molecular Reaction Dynamics, Part 1: Experiment*; Ng, C-Y.; Baer, M., Eds.; Advances in Chemical Physics Series, Vol. LXXXII; Wiley: New York, 1992.
16. Douglas, D. J.; French, J. B. *J. Am. Soc. Mass Spectrom.* **1992,** *3*, 398–408.
17. Krutchinsky, A. N.; Chernushevich, I. V.; Spicer, V. L.; Ens, W.; Standing, K. G. *J. Am. Soc. Mass Spectrom.* **1998,** *9*, 569–579.
18. Duelcks, T.; Thierolf, M.; Karas, M. *Discrimination of Analytes by Absorption to the Gas Capillary Wall in Nano-Electrospray Mass Spectrometry*, a poster presented at *47th ASMS conference on MASS Spectrometry and Allied Topics*; Dallas, TX, June 13–17, 1999.
19. Axelsson, J.; Reimann, C. T.; Sundqvist, B. U. R. *Int. J. Mass Spectrom. Ion Processes* **1994,** *133*, 141–155.
20. Donald, L. J.; Chernushevich, I. V.; Zhou, J.; Verentchikov, A.; Poppe-Schriemer, N.; Hosfield, D. J.; Westmore, J. B.; Ens, W.; Duckworth, H. W.; Standing, K. G. *Protein Sci.* **1996,** *5*, 1613–1624.
21. Qin, J.; Fenyö, D.; Zhao, Y.; Hall, W. W.; Chao, D. M.; Wilson, C. J.; Young, R. A.; Chait, B. T. *Anal. Chem.* **1997,** *69*, 3995–4001.
22. Wilm, M.; Shevchenko, A.; Houthaeve, T.; Bret, S.; Schweigere, L.; Fotsis, T.; Mann, M. *Nature* **1996,** *379*, 466–469.
23. Steen, H.; Andersen, J.; Küster, B.; Podtelejnikov, A.; Molina, H.; Mann, M. *Increasing the Throughput of Protein Identification Using Nanoelectrospray QqTOF Mass Spectrometry* a poster presented at *47th ASMS Conference on Mass Spectrometry and Allied Topics*; Dallas, TX, June 13–17, 1999.
24. Gobom, J.; Nordhoff, E.; Ekman, R.; Roepstorff, P. *Int. J. Mass Spectrom. Ion Processes* **1997,** *169/170*, 153–163.
25. Zhang, W.; Chait, B. T., in press.
26. Fenyö, D.; Qin, J.; Chait, B. T. *Electrophoresis* **1998,** *19*, 998–1005.
27. Patterson, S. D. *Anal. Biochem.* **1994,** *221*, 1–15; Kuster, B., Mann, M. *Curr. Opin. Struct. Biol.* **1998,** *8*, 393–400; Yates, J. R. 3rd *Electrophoresis* **1998,** *19*, 893–900.
28. Jensen, O. N.; Podtelejnikov, A.; Mann, M. *Rapid Commun. Mass Spectrom* **1996,** *10*, 1371–1378.
29. http://www.proteome.com
30. Neubauer, G.; King, A.; Rappsilber, J.; Calvio, C.; Watson, M.; Ajuh, P.; Sleeman, J.; Lamond, A.; Mann, M. *Nat. Genet.* **1998,** *20*, 46–50.
31. Fenyö, D.; Zhang, W.; Chait, B. T.; Beavis, R. C. *Anal. Chem.* **1996,** *68*, 721A–726A.
32. Huang, S.; Elliott, R. C.; Liu, P. S.; Koduri, R. K.; Weickmann, J. L.; Lee, J. H.; Blair, L. C.; Ghosh-Dastidar, P.; Bradshaw, R. A.; Bryan, K. M.; et al. *Biochemistry* **1987,** *26*, 8242–8246.
33. Putney, S.; Herlihy, W.; Royal, N.; Pang, H.; Aposhian, H. V.; Pickering, L.; Belagaje, R.; Biemann, K.; Page, D.; Kuby, S.; et al. *J. Biol. Chem.* **1984,** *259*, 14317–14320.
34. Nayler, O.; Cap, C.; Stamm, S. *Genomics* **1998,** *53*, 191–202.