Chul Yoo[1]
Jia Zhao[1]
Manoj Pal[1]
Katherine Hersberger[1]
Christian G. Huber[2]
Diane M. Simeone[3]
David G. Beer[3, 4]
David M. Lubman[1, 3, 4]

[1]Department of Chemistry,
 University of Michigan,
 Ann Arbor, MI, USA
[2]Department of Chemistry,
 Instrumental Analysis and
 Bioanalysis,
 Saarland University,
 Saarbrücken, Germany
[3]Department of Surgery,
 The University of Michigan
 Medical Center,
 Ann Arbor, MI, USA
[4]Comprehensive Cancer Center,
 University of Michigan
 Medical Center,
 Ann Arbor, MI, USA

Research Article

Automated integration of monolith-based protein separation with on-plate digestion for mass spectrometric analysis of esophageal adenocarcinoma human epithelial samples

A unique approach of automating the integration of monolithic capillary HPLC-based protein separation and on-plate digestion for subsequent MALDI-MS analysis has been developed. All liquid-handling procedures were performed using a robotic module. This automated high-throughput method minimizes the amount of time and extensive labor required for traditional in-solution digestion followed by exhaustive sample cleanup and analysis. Also, precise positioning of the droplet from the capillary HPLC separation onto the MALDI plate allows for preconcentration effects of analytes for improved sensitivity. Proteins from primary esophageal Barrett's adenocarcinoma tissue were prefractionated by chromatofocusing and analyzed successfully by this automated configuration, obtaining rapid protein identifications through PMF and sequencing analyses with high sequence coverage. Additionally, intact protein molecular weight values were obtained as a means to further confirm protein identification and also to identify potential sequence modifications of proteins. This simple and rapid method is a highly versatile and robust approach for the analysis of complex proteomes.

1 Introduction

The immense complexity of the human proteome [1] presents a great challenge that requires the analysis of large numbers of proteins [2]. Therefore, it is important that proteomic methodologies should be simple, auto-mated, and versatile for rapid and reproducible analysis [3]. 2-DE [4] is still the most widely utilized method for large scale proteomics applications, where in-gel digestion of each protein spot is analyzed by sub-sequent ESI- or MALDI-based MS analysis. Despite development of robotic systems for gel-spot picking and excision [5–7], 2-DE coupled with MS remains difficult and time-consuming due to extensive sample

Correspondence: Dr. David M. Lubman, The University of Michigan Medical Center, 1150 West Medical Center Dr, MSRB1 RMA510B, Ann Arbor, MI 48109, USA
E-mail: dmlubman@umich.edu
Fax: +1-734-763-6199

Abbreviations: CF, chromatofocusing; **NPS**, nonporous; **OG**, *n*-octyl-D-glucopyranoside; **PTM**, post-translational modification

cleanup and digestion of a few hundred to several thousands of spots that may be present in a sample representing the human proteome.

Recently, the integration of HPLC with offline fraction collection for MALDI-MS analysis has received attention [8, 9] and has been applied to several systems [10–13]. LC/MALDI has also proven useful for unique applications and has been demonstrated to be useful in quantitative studies [14] and in the characterization of post-transla-tional modifications (PTMs) [15] of proteins. In experi-ments utilizing LC/MALDI and applied to complex bio-logical mixtures, proteins have been exclusively analyzed by enzymatic digestion of whole cell lysates [16–20] using sequential coupling of ion exchange and RP-HPLC separations. Although this shotgun proteomics approach is useful for a comprehensive analysis on a global scale [21], the results can be misleading, where the complexity of samples may result in false positive identifications when only a small number of peptides are matched [22]. It was previously shown that proteome analysis performed at the level of intact proteins [23] through 2-D liquid phase

separations can help avoid such problems, where each of the protein fractions was further examined in detail and compared with intact molecular weight (MW) analysis [24–27]. This concept was used to develop a novel method that integrates monolithic HPLC separation of intact proteins with on-MALDI plate tryptic digestion [28] for rapid identification of proteins. It has been recently applied to a human breast cancer cell line [29], where lengthy experimental procedures required by in-solution digestion were greatly minimized. Also, unique characteristics of the monolith [30–32] as a separation medium, including high recovery and rapid separation speed [33, 34], were found to be ideally suited for the purpose of LC/MALDI analysis.

Automation of the sample-handling process is an important issue to be considered in proteomics in order to develop a truly high-throughput methodology facilitating the analysis of large numbers of proteins [35]. An automated integration of HPLC with offline fraction collection for MALDI-MS analysis has been attempted previously in several studies for the analysis of protein digests [10, 36]. In the present studies, we focus on the advancement of this method by automation of all liquid-handling procedures. The proteins from human esophageal adenocarcinoma [37, 38] a cancer type that is demonstrating an alarming increase in incidence, were analyzed to demonstrate the versatility and applicability of this method using whole tissue proteins. A selected set of pH-fractionated Barrett's cancer tissue samples was analyzed for protein identification. HPLC/ESI-TOF MS was performed to compare intact protein MW for further confirmation.

2 Materials and methods

Proteins from esophageal tissue samples were separated by chromatofocusing (CF) for pH fractionation. One of the fractions was selected for further separation by monolithic capillary RP-HPLC connected to the modified nano-plotter (GeSim), equipped with xyz-robotic unit for automatic offline peak collection. The fractions at 30 s intervals were collected directly on the MALDI plate pre-coated with trypsin for on-plate digestion and subsequent MS analysis. A simplified LC/MALDI configuration is described in Fig. 1, where all liquid-handling procedures were automated in this experiment. The same pH fraction was also analyzed by online nonporous (NPS) RP-HPLC/ESI-TOF MS to obtain accurate intact protein MW values for comparison with theoretical MW values.

2.1 Sample preparation

Patient consent was received according to guidelines set forth by the institutional review board (IRB) of the University of Michigan. Tissues were collected and analyzed from patients undergoing esophagectomy at the University of Michigan Health System between 1991 and 2001. Patients received no preoperative radiation or chemotherapy. Each esophageal adenocarcinoma tumor specimen was immediately frozen in liquid nitrogen and stored at −80°C. Cryostat sectioning of all tumors were performed and only portions of tumors containing >80% tumor cellularity were utilized for subsequent protein isolation.
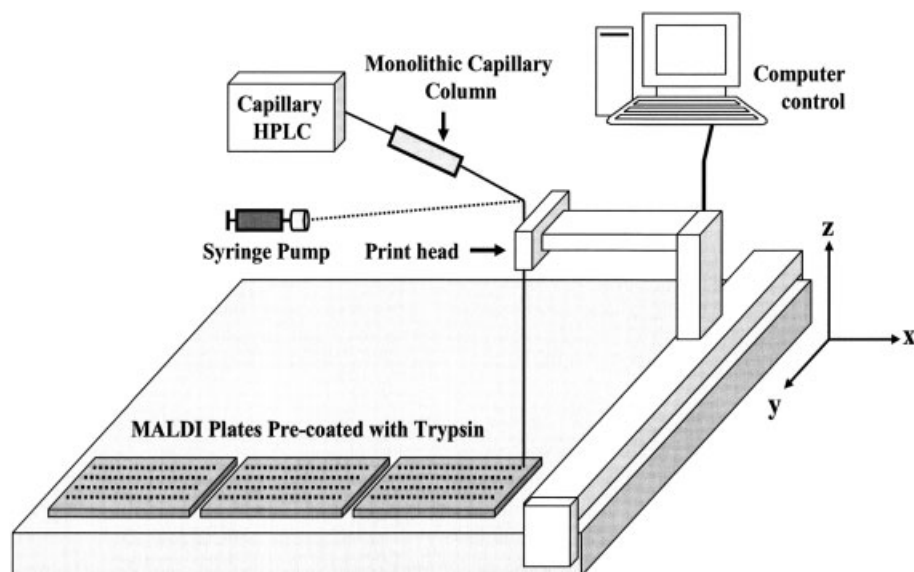

Figure 1. A simplified description of automated LC/MALDI configuration constructed by modifying nano-plotter (not in scale) for all liquid-handling procedures.

2.2 Cell lysis and buffer exchange

A 200 μg of tissue sample was lysed with 2 mL of lysis buffer which consisted of 7.5 M urea, 2.5 M thiourea, 4% *n*-octyl-D-glucopyranoside (OG), 10 mM tris(2-car-boxyethyl)phosphine, 10% glycerol, 50 mM Tris (all from Sigma, St. Louis, MO), and 40 μL protease inhibitor solution (one tablet in 1 mL PBS buffer, Roche, Indianapolis, IN). Samples were then homogenized, vortexed frequently for 1 h at room temperature, and centrifuged at 30 000 rpm for 70 min at 4°C. The collected supernatant was subjected to buffer exchange against CF start buffer using a PD-10 G-25 column (Amersham Biosciences, Piscataway, NJ). Bradford protein assay kit (BioRad, Hercules, CA) was used to quantify the amount of proteins in mixtures. The buffer-exchanged protein mixtures were stored at −80°C until further use.

2.3 CF separation

CF separation was performed on an HPCF-1D column (2.1 mm id × 250 mm length, Beckman Coulter, Fullerton, CA) using a Beckman System Gold HPLC. Prior to sample loading, the column was equilibrated with a start buffer containing 25 mM Bis–Tris propane (Sigma), 6 M urea, and 1% OG, where its pH was adjusted to pH 7.4 with saturated iminodiacetic acid solution (Sigma). Proteins (*ca.* 4.5 mg) were loaded to the elution buffer at pH 4.0 containing 10% Polybuffer 74 (Amersham Pharmacia), 6 M urea, and 1% OG at a flow rate of 0.2 mL/min. A linear pH gradient was generated so that proteins eluted off according to their p*I* for detection at 280 nm. Accurate pH was measured online by a postdetector pH electrode (Lazar Research Laboratories, Los Angeles, CA) for fraction collection at every 0.2 pH unit in the range of pH 7.0–4.0.

2.4 Online NPS-RP-HPLC/ESI-TOF MS for intact protein molecular weight determination

Fractions obtained from CF were subjected to NPS-RP-HPLC separation using an ODSIII-E column (4.6 mm id × 33 mm L, Eprogen, Darien, IL) packed with 1.5 μm NPS silica. The column was maintained at 60°C by a column heater (model 7971, Jones Chromatography, Resolution Systems, Holland, MI) to perform separation at a flow rate of 0.5 mL/min, where 40% was split into an ESI-TOF MS (LCT Premier, Waters/Micromass, Milford, MA). Formic acid (0.5%, Sigma) was added postsplitter using a syringe pump (Harvard Apparatus, Holliston, MA). The remainder of the flow was monitored at 214 nm (detector model 166, Beckman Coulter) for offline peak collection using a fraction collector (model SC100, Beckman Coul-

ter) controlled by an in-house acquisition program. A mobile phase system of A and B was comprised of 0.1% TFA (J. T. Baker, Phillipsburg, NJ) in purified DI water (Millipore RG system, Bedford, MA) and ACN (Sigma), respectively. The gradient profile was as follows: 5–15% B in 1 min, 15–25% B in 2 min, 25–31% B in 3 min, 31–41% B in 10 min, 41–47% B in 3 min, 47–67% B in 4 min, and 67–100% B in 1 min. The capillary voltage for electrospray was set at 3200 V, sample cone at 35 V, extraction cone at 3 V, and reflectron at 750 V. Desolvation temperature was maintained at 330°C and source at 130°C with desolvation gas flow of 650–800 L/h. Intact protein MW was obtained by automatic deconvolution using Protein Trawler software (BioAnalyte, Portland, ME).

2.5 Monolithic capillary HPLC for protein separation

The preparation of copolymerized styrene and divinylbenzene monolithic capillary columns (200 μm id × 60 mm L) was performed according to procedures described elsewhere [33]. A capillary pump (Ultra-Plus II MD, Micro-Tech Scientific, Vista, CA) was used for separation. The capillary column was directly mounted on a microinjector with a 500 nL internal sample loop (Valco Instruments, Houston, TX). The capillary HPLC separation was controlled at 60°C with an in-house column heater. The flow from the solvent delivery pump was split pre-column to generate a flow rate of *ca.* 2.5 μL/min through the monolithic columns. A mobile phase system of two solvents was used, wherein solvents A and B were composed of 0.05% formic acid in HPLC-grade water (Fisher Scientific, Hanover Park, IL) and ACN, respectively. A linear gradient of 0–100% B in 18 min was applied and held at 100% B for 3 min.

2.6 Automated integration of protein separation and on-MALDI plate enzymatic digestion

A syringe pump was interfaced with an xyz-module of Nano-Plotter 2.0 piezoelectric pipetting system (GeSiM, Großerkmannsdorf, Germany) by a capillary tubing in order to precoat the MALDI plate with 0.5 μL of TPCK-modified trypsin solution of 0.15 μg/μL (Promega, Madison, WI) into each spot. Prefractionated esophageal tumor proteins at pH 5.6–5.8 were loaded onto a monolithic capillary column for separation, where the connecting capillary tubing off the column was interfaced directly to the print head of the nano-plotter. An instrument controller program written in nano-plotter language (NPL) was used to control the print head for precise movements along the *x*, *y*, and *z* axes for automatic control over the designated

fraction deposition time of 30 s intervals onto the MALDI plate. Fraction collection was performed in a real-time mode with proper calibration. Following the protein collection onto each trypsinized spot, 0.5 μL of 50 mM ammonium bicarbonate (Sigma) was automatically added to the top layer of each spot using a syringe pump and the plate was maintained at room temperature for digestion for *ca.* 10 min. Then, 0.5 μL of 0.1% TFA was added to each spot to halt digestion, followed by addition of 0.5 μL of α-cyano-4-hydroxycinnamic acid (α-CHCA, Sigma) matrix solution prepared by diluting saturated α-CHCA with 60% ACN/0.1% TFA at a 1:4 ratio. The internal standards of angiotensin I, adrenocorticotropic hormone (ACTH) fragment 1–17, and ACTH fragment 18–39 (all from Sigma) were included to have a final concentration of 50 fmol in every spot.

2.7 MALDI-TOF MS analysis and database searching

The MALDI-TOF MS analysis was performed on a Tof-Spec2E (Waters/Micromass) equipped with delayed extraction in reflectron mode using a 337 nm Nd:YAG laser as the ionization source. The delay time was set at 520 ns, the extraction voltage at 1:1 to the source voltage at 20 kV, and the pulse voltage at 2300 V. Each spectrum was internally calibrated and monoisotopic peptide masses were obtained using MassLynx software version 4.0 (Waters-Micromass) for submission to the MS-Fit search engine using Swiss-Prot database for protein identification. The search was carried out under the species of *Homo sapiens* at <100 ppm of mass tolerance with no limitation set for MW and p*I*. One missed cleavage was allowed and the possible modifications included N-terminal Gln to pyroGlu, oxidation of Met, N-terminal acetylation, and phosphorylation at S, T, and Y. Protein identification was accepted as a match by filtering according to the following parameters: MOWSE score of $>10^3$ and sequence coverage of >20% and comparison with intact MW values.

2.8 MALDI-TOF/TOF MS analysis and database searching

The MALDI-TOF/TOF MS analysis was performed using a 4800 MALDI TOF/TOF analyzer (Applied Biosystems, Foster City, CA) with a 384-well plate. A CID voltage of 2 kV was used throughout the runs. S/N threshold of 50 and 30 were used for selecting the precursor and fragment peaks, respectively. The MASCOT available online at www.matrixscience.com was used for database searching under Swiss-Prot for *Homo sapiens*. Mass tolerances of 50 ppm and 0.3 Da were used for precursor

and fragment ions, respectively, by allowing one missed cleavage. The same variable modifications were allowed as in PMF analysis.

3 Results and discussion

3.1 Design of automated integration of monolithic LC-based protein separation and on-plate tryptic digestion

A number of different approaches have been developed to interface HPLC and MALDI-MS. It has recently been shown that the use of sheath gas flow around the emerging droplet of LC effluent assisted solvent evaporation, but prevented the oxidation of analytes [18]. In another study, a pull-down deposition method, applying an electric field to the analyte-collection device, has been demonstrated to effectively deposit a droplet from the HPLC [36]. This method can be particularly useful where very small droplets are generated through nanoscale LC to reduce droplet adhesion to the connecting tubing. However, it may complicate the experimental platform with the use of high voltage applied across the entire analyte-collection device, where electrochemical reactions may adversely affect MS spectra. Others reported a pressure pulse-driven dispenser device [39] and a heated droplet interface [40].

Figure 1 describes the schematic of the automated configuration to integrate a monolithic capillary HPLC with on-plate digestion of intact proteins for subsequent MALDI-TOF MS analysis. This experimental configuration involves a continuous liquid deposition through contact to the MALDI plate for simplicity, which also proved to provide sufficiently high precision for the purpose of this study. Given the flow rate used for HPLC separation in our experiment, *ca.* 1.2 μL from the HPLC is deposited on each spot. Unlike nano-LC applications, this effluent volume is considered to be sufficiently large so that droplet adhesion to the connecting capillary is less of a concern, eliminating the need for a sophisticated platform for liquid collection.

Figure 2 shows a very close proximity between the interface of LC effluent and the MALDI plate, which measures *ca.* 0.1 mm. Compared to the protein collection and liquid deposition performed manually in our previous study [29], this automated device can ensure highly precise positioning of liquid deposition into a very small area of the spot surface. This provides a method enhancing protein concentration for highly efficient digestion, therefore, detection for improved sensitivity for MS analysis. This added advantage can be particularly useful for analyzing proteins present in low abundance or limited amounts of sample.
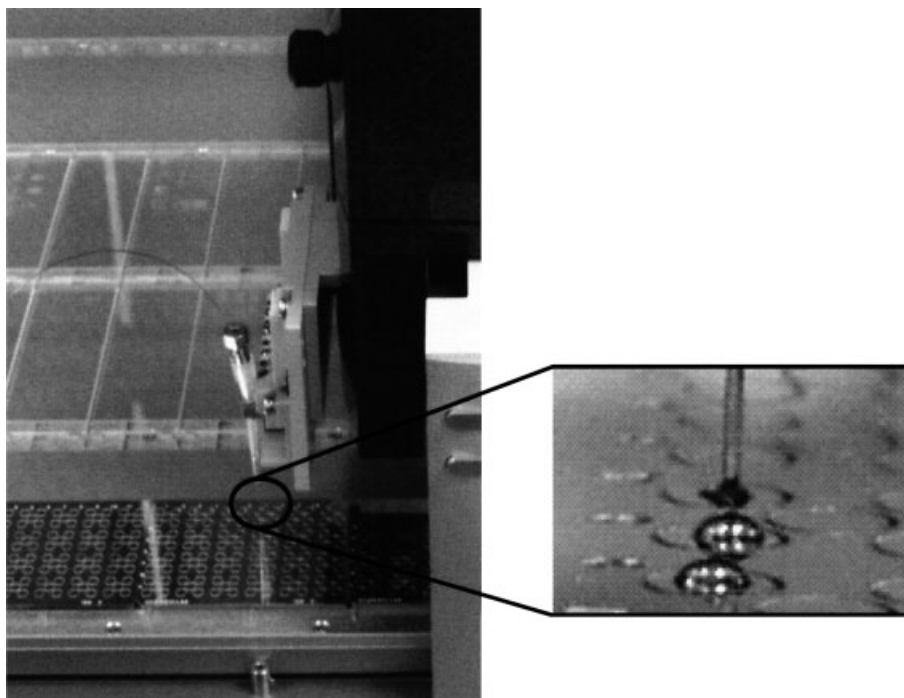
Figure 2. Automated LC/ MALDI interface where proteins separated by capillary mono-lithic HPLC are deposited di-rectly onto the MALDI target plate precoated with trypsin.

The optimal time interval for analyte collection will be de-pendent on the complexity of the samples [17]. Although a 30 s interval was used for protein collection time in this study, our automation device can be programed for shorter analyte deposition time down to several seconds or less that may be necessary for the analysis of very complex proteomes. The flexibility of the automated sys-tem allows multiple numbers of the MALDI plates of varying sizes to be accommodated, which can further enhance throughput of this method.

3.2 Identification of human esophageal cancer tissue proteins

In our previous study, human breast cancer cell line pro-teins prefractionated by a Rotofor device of preparative scale were chosen as a model to show the applicability of LC/MALDI to biological mixtures [29]. However, the use of Rotofor device is less appropriate when the samples to be analyzed are limited in quantity, as it is often the case for human tissue samples. Recently, the usefulness of CF [41] as a first-dimensional fractionation has been suc-cessfully demonstrated for the analysis of various human cancer proteomes [24, 25]. In addition to reducing the complexity of samples, it is important to note that experi-mental pI values of proteins can be monitored online by CF so that the presence of possible PTM can be asses-sed by comparing to theoretical pI values [42]. In this

experiment, we extended the application of the LC/ MALDI scheme to human cancer tissue samples, where prefractionation was performed by CF.

Proteins fractionated in the pH range of 5.6–5.8 from an esophageal adenocarcinoma tissue sample was sub-jected to monolithic HPLC separation interfaced with on-plate digestion for subsequent MALDI-TOF MS and MALDI-TOF/TOF MS analyses by automating all liquid-handling procedures, where a limited amount of CF-frac-tionated sample was required due to the low loading ca-pacity of the monolithic capillary column [43]. Table 1 shows a list of proteins identified through both PMF and intact protein MW analyses in each of the collection times, where the LC system used a delay time of *ca.* 6.5 min. A monolith-based protein separation was performed for the collection time of *ca.* 20 min. As discussed in a recent work [29], a monolithic capillary column of the small id used in the experiment may not be the most suitable means to separate many proteins present in complex mixtures with high resolution and high efficiency. How-ever, a somewhat lower resolution of protein peaks obtained from HPLC separation does not pose a serious problem for reliable protein identification, since sub-sequent MS analysis can resolve the proteins.

Figure 3 illustrates a MALDI-TOF MS spectrum obtained for the splicing factor, proline- and glutamine-rich, a large protein of >70 kDa, that was identified with 43% of sequence coverage, where its identified tryptic peptide

Table 1. List of proteins identified by automated interfacing of monolithic HPLC with MALDI-TOF MS through on-plate digestion for Barrett's esophageal adenocarcinoma samples prefractionated by CF at pH 5.6–5.8

Collection time (min)	Protein name	Accession no.	Theoretical		Experi-mental MW	Sequence coverage (%)	MOWSE score
			MW	p*I*			
6.5–7	ATP synthase beta chain, mitochondrial precursor	P06576	56 560 (51 736)[a)	5.3	51 636	36	1.397e + 04
9–9.5	Neuron-specific calcium-binding protein hippocalcin	P84074	22 428	4.9	22 574	48	5041
10–10.5	NADH-ubiquinone oxidoreductase B16.6 subunit	QP0J0	16 698	8.0	16 682	45	1.268e + 04
11–11.5	Propionyl-CoA carboxylase beta chain	P05166	58 206	7.6	57 695	30	3182
12.5–13	Paraplegin	Q9UQ90	88 136	8.7	88 634	25	6.390e + 05
13.5–14	Triosephosphate isomerase	P60174	26 670	6.4	26 652	40	1.456e + 04
	Heterogeneous nuclear ribonucleoprotein A1	P09651	38 846	9.3	39 461	33	3.354e + 05
14.5–15	Splicing factor, proline- and glutamine-rich	P23246	76 150	9.5	75 490	43	1.113e+12
15–15.5	Aldo-keto reductase family 1 member C3	P42330	36 845	8.1	36 886	52	2.008e + 07
16–16.5	Calgranulin B	P06702	13 242	5.7	13 816	66	8.249e + 04
17–17.5	Lymphoid-enhancer-binding factor 1	Q9UJU2	44 201	6.9	44 549	20	1.104e + 04
18–18.5	HLA class I histocompatibility antigen	Q29718	40 421	5.7	40 195	35	1.815e + 04
18.5–19	Aldehyde dehydrogenase 1A1	P00352	54 862	6.3	54 862	26	6.096e + 04
19–19.5	39S Ribosomal protein L45	Q9BRJ2	35 351	9.1	35 214	26	4949
19.5–20	Tumor necrosis factor receptor superfamily member 11B precursor	O00300	46 041	8.7	45 799	32	1.943e + 04
	Cytokine-inducible SH2-containing protein	Q9NSE2	28 663	6.5	28 635	47	1.415e + 04
21–21.5	MAP kinase-interacting serine/threonine kinase 1	Q9BUB5	51 343	6.3	50 686	38	2.607e + 04
22.5–23	Spindlin (ovarian cancer-related protein)	Q9GZT2	27 077	5.8	27 246	30	1.059e + 05
23.5–24	Transgelin-3	Q9UI15	22 473	6.8	22 878	38	6968
24.5–25	Alcohol dehydrogenase	P14550	36 573	6.3	36 531	25	1.415e + 04
25–25.5	Cyclin H	P51946	37 644	6.7	37 533	43	2496
26.5–27	Peptidyl-prolyl *cis–trans* isomerase	Q13451	51 213	5.7	50 686	31	4.837e + 04

a) The number in parentheses indicates theoretical MW of the truncated form.
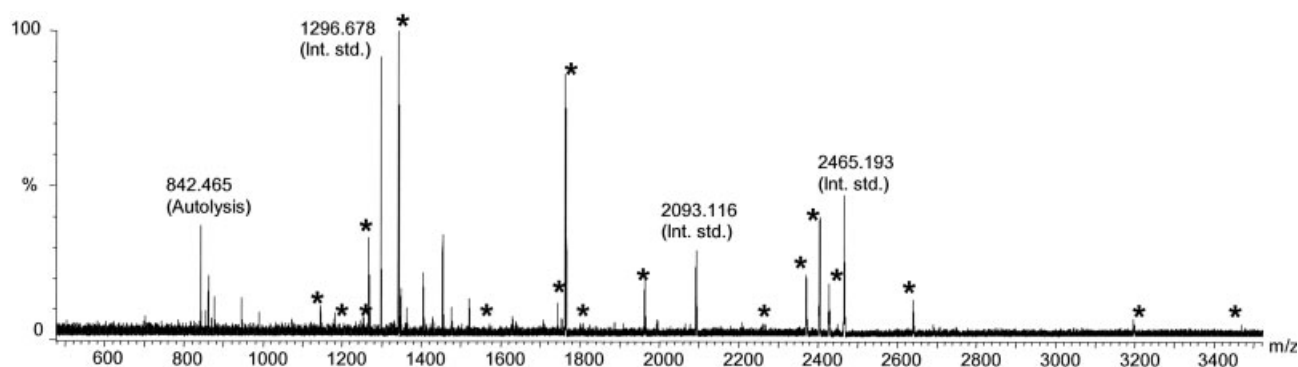

Figure 3. MALDI-TOF MS spectrum (unprocessed) obtained for the splicing factor, proline- and glutamine-rich, from an esophageal tissue sample. * Indicates peptides identified by monolithic LC/MALDI scheme through on-plate digestion. Refer to Table 2 for information regarding each peptide.

sequences are summarized in Table 2. As shown in Table 1, relatively high sequence coverage was obtained for most of the proteins and this is considered to be primarily due to minimal sample transfers associated with direct LC/MALDI interface to prevent sample loss. Also, compared to traditional in-solution digestion, it appears

that protein enrichment effects obtained through a small spot size due to precise control of liquid-handling combined with a large surface area for enzymatic reactions provided by precoated trypsin perhaps resulted in more efficient digestion. Also, reduced trypsin autolysis relative to in-solution digestion may be expected due to trypsin

Table 2. A list of sequenced tryptic peptides identified from splicing factor, proline- and glutamine-rich, where peptides with * are shown in Fig. 3

Exp. $[M+H]^+$	Theor. $[M+H]^+$	Possible modifications	Start	End	Missed cleavages	Database sequence
770.3267	770.3321		594	599	0	EESYSR
830.4011	830.4484		543	548	1	QEELRR
830.4011	830.4484		542	547	1	RQEELR
*869.3300	869.3650		600	606	0	MGYMDPR
886.4747	886.4998		400	407	0	AVVIVDDR
1120.4904	1120.5210		682	693	0	GMGPGTPAGYGR
*1143.5992	1143.6275		366	376	0	FATHAAALSVR
*1170.5725	1170.5036	1Met-ox	600	608	1	MGYMDPRER
1235.5787	1235.5149	PyroGlu 1Met-ox	582	590	1	QREMEEQMR
1235.5787	1235.6497		287	296	1	RPGEKTYTQR
1245.6538	1245.6955		414	425	0	GIVEFASKPAAR
*1252.5765	1252.5414	1Met-ox	582	590	1	QREMEEQMR
*1252.5765	1252.6214		320	330	0	YGEPGEVFINK
*1267.5986	1267.6217		33	44	0	SPPPGMGLNQNR
*1341.6605	1341.6664		667	681	0	FGQGGAGPVGGQGPR
*1572.7620	1572.7263		549	560	1	MEELHNQEMQKR
*1572.7620	1572.7263		548	559	1	RMEELHNQEMQK
*1743.8246	1743.8951		343	358	1	ALAEIAKAELDDTPMR
*1762.7778	1762.7825		480	493	0	FAQHGTFEYEYSQR
*1807.8894	1807.8241	1PO4	664	681	1	TERFGQGGAGPVGGQGPR
*1807.8894	1807.9118		299	314	0	LFVGNLPADITEDEFK
*1963.9795	1964.0129		299	315	1	LFVGNLPADITEDEFKR
*2264.9968	2264.9647	1PO4	519	536	0	LESEMEDAYHEHQANLL
*2370.1633	2370.2253		212	236	0	MPGGPKPGGGPGLSTPGGHPKPPHR
*2403.0925	2403.1356		246	267	0	QHHPPYHQQHHQGPPPGGPGGR
*2428.0911	2428.1203		517	536	1	DKLESEMEDAYHEHQANLLR
*2639.1956	2639.2993		377	399	0	NLSPYVSNELLEEAFSQFGPIER
*3195.3171	3195.4372		631	663	0	FPPLGGGGGIGYEANPGVPPATMSGSMMGSDMR
*3467.1528	3467.3260	2Met-ox 3PO4	631	663	0	FPPLGGGGGIGYEANPGVPPATMSGSMMGSDMR

immobilization [44] on the plate to generate less complicated mass spectra for unambiguous interpretation. In order to further confirm the protein identification procedures, LC/MALDI-based sequencing analysis was performed using MALDI-TOF/TOF MS. Figure 4 shows an example for one of the proteins in Table 1, calgranulin B, where one of its tryptic peptides, NIETIINTFHQYSVK (11–25) has been successfully sequenced. This protein was identified with >50% of sequence coverage through sequencing analysis.

Also, since proteins are sufficiently separated from each other for direct deposition onto the MALDI target plate to be confined into a very small spot size with minimal sample loss expected, several less abundant proteins were identified in this work, such as cyclin H [45] in Table 1. Another protein, MAPK-interacting serine/threonine kinase 1 [46], activated by phosphorylation, is also known to be expressed at low levels. It would be difficult to

detect these proteins using a traditional proteomics approach, where sample loss due to sample-handling procedures is inevitable.

In addition to helping to confirm protein identifications determined by PMF and sequencing analyses, intact protein MW values were obtained separately by NPS-RP-HPLC/ESI-TOF MS for comparison to theoretical MW values with excellent agreement, as shown in Table 1. Although the LC/MALDI method alone can provide highly confident protein identification due to relatively high sequence coverage, the usefulness of intact MW values can be observed from a close examination of several proteins. For instance, the splicing factor, proline- and glutamine-rich, is annotated with two distinct isoforms A and F (accession nos. P23246-1 and P23246-2, respectively) that slightly differ in amino acid compositions in the region of 663–707 to result in the isoform A being larger than the isoform F by *ca.* 4000 Da (www.expasy.org).
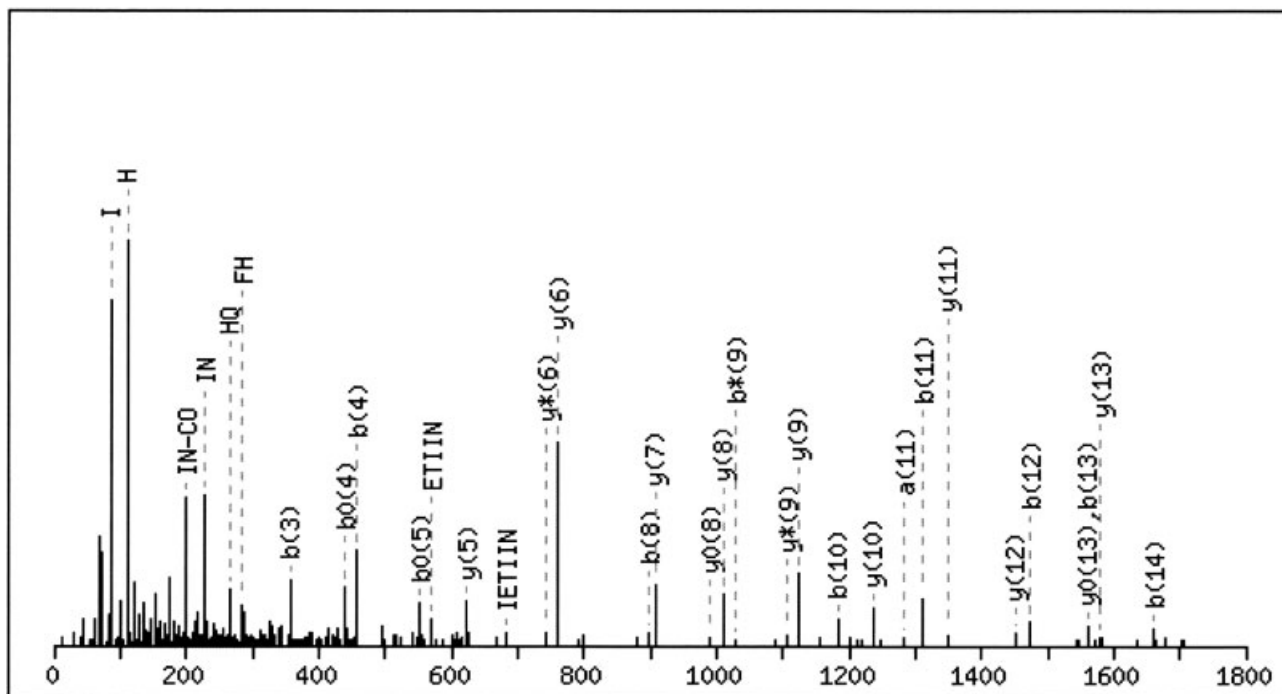
Figure 4. MALDI-TOF/TOF MS spectrum for a tryptic peptide NIETIINTFHQYSVK (11–25) of calgranulin B.

Protein identification by PMF or tandem MS-based sequencing analysis is solely based on partial sequence coverage of proteins, where sequence variations, such as truncation and isoforms, are often undetected. In our experiment, however, comparison between theoretical and experimental intact protein MW allows us to conclude that it is likely that the isoform A of splicing factor, proline- and glutamine-rich, is present in the esophageal tumor sample. Also, significant deviation observed between the experimental and theoretical protein MW values can suggest the presence of possible modifications in certain proteins. An example can be found from ATP synthase beta chain that becomes truncated by the loss of transit peptide sequence upon entry into mitochondria, where this can also be explained by an excellent agreement between its theoretical intact MW value and experimental MW value, as shown in Table 1. A close match between theoretical and experimental MW values of all proteins in Table 1 further confirms the protein identification obtained by PMF analysis, where slight deviations can suggest the presence of potential PTM [24].

4 Concluding remarks

We have developed a novel method that automates the integration of monolithic capillary HPLC separation of intact proteins and on-plate enzymatic digestion for MALDI-based MS analysis. The method is shown to be highly versatile and robust and is suitable for the analysis of complex proteomes, as successfully demonstrated for the analysis of esophageal tissue proteins. It is important to emphasize that the current study analyzed the proteome at the protein level, unlike the majority of previous LC/MALDI work performed by shotgun proteomics that may result in ambiguous identifications due to the complexity of mixtures upon digestion of a large number of proteins. Based on the measured intact protein MW values and substantially higher protein sequence coverage from the protein-based LC/MALDI, compared to shotgun-derived LC/MALDI, it is likely that the developed method can help analyze PTM or other sequence variations. Although the analysis of complex proteomes is often performed by ESI-based MS, integration of protein separation directly to MALDI is an effective means to solve compatibility issues with various LC solvent compositions and contaminants to which the ESI process is substantially less tolerant.

Automated coupling of LC-based protein separation for direct on-plate enzymatic digestion is an attractive means to realize high-throughput proteomics, where the process can be fully automated and enhance reproducibility. Another distinct advantage includes accurate control of the spot surface area upon deposition that allows for the analysis of low-abundance proteins by enrichment of

analytes [47]. This is essential in proteomics studies, where a wide dynamic range of protein concentration presents a great challenge. Considering that the droplet size is dependent on the surface tension and viscosity of the solution [48], further development is ongoing to diminish the spot diameter. The method has potential to become even more powerful when protein identification through automated software analysis can be combined.

5 References

[1] Han, K. K., Martinage, A., *Int. J. Biochem.* 1992, *24*, 19–28.

[2] Cagney, G., Amiri, S., Premawaradena, T., Lindo, M., Emili, A., *Proteome Sci.* 2003, *1*, 5.

[3] Quadroni, M., James, P., *Electrophoresis* 1999, *20*, 664–677.

[4] O'Farrell, P. H., *J. Biol. Chem.* 1975, *250*, 4007–4021.

[5] Weeks, M. E., Sinclair, J., Jacob, R. J., Saxton, M. J. *et al.*, *Proteomics* 2005, *5*, 1669–1685.

[6] Malmstrom, J., Larsen, K., Malmstrom, L., Tufvesson, E. *et al.*, *Electrophoresis* 2003, *24*, 3806–3814.

[7] Klene, T. G., Andreasen, c. M., Kjeldal, H. O., Leonardsen, L. R. *et al.*, *Anal. Chem.* 2004, *76*, 3576–3583.

[8] Foret, F., Preisler, J., *Proteomics* 2002, *2*, 360–372.

[9] Zhong, H., Marcus, S. L., Li, L., *J. Am. Soc. Mass Spectrom.* 2005, *16*, 471–481.

[10] Brombacher, S., Owen, S. J., Volmer, D. A., *Anal. Bioanal. Chem.* 2003, *376*, 773–779.

[11] Keil, O., LeRiche, T., Deppe, H., Volmer, D. A., *Rapid Commun. Mass Spectrom.* 2002, *16*, 814–820.

[12] Tegeler, T. J., Merchref, Y., Boraas, K., Reilly, J. P., Novotny, M. V., *Anal. Chem.* 2004, *76*, 6698–6706.

[13] Zhen, Y., Xu, N., Richardson, B., Becklin, R. *et al.*, *J. Am. Soc. Mass Spectrom.* 2004, *15*, 803–822.

[14] Griffin, T. J., Gygi, S. G., Rist, B., Aebersold, R. *et al.*, *Anal. Chem.* 2001, *73*, 978–986.

[15] Lochnit, G., Geyer, R., *Biomed. Chromatogr.* 2004, *18*, 841–888.

[16] Canelle, L., Pionneau, C., Marie, A., Bousquet, J. *et al.*, *Rapid Commun. Mass Spectrom.* 2004, *18*, 2785–2794.

[17] Young, J. B., Li, L., *J. Am. Soc. Mass Spectrom.* 2006, *17*, 325–334.

[18] Mirgorodoskaya, E., Braeuer, C., Fucini, P., Lehrach, H., Gobom, J., *Proteomics* 2005, *5*, 399–408.

[19] Hattan, S. J., Marchese, J., Khainovski, N., Martin, S., Juhasz, P., *J. Proteome Res.* 2005, *4*, 1931–1941.

[20] Chen, V. C., Cheng, K., Ens, W., Standing, K. G. *et al.*, *Anal. Chem.* 2004, *76*, 1189–1196.

[21] Wang, Y., Rudnick, P. A., Evans, E. L., Li, J. *et al.*, *Anal. Chem.* 2005, *77*, 6549–6556.

[22] Veenstra, T. D., Conrads, T. P., Issaq, H., *Electrophoresis* 2004, *25*, 1278–1279.

[23] Slysz, G. W., Schriemer, D. C., *Anal. Chem.* 2005, *77*, 1572–1579.

[24] Kreunin, P., Urquidi, V., Lubman, D. M., Goodison, S., *Proteomics* 2004, *4*, 2754–2765.

[25] Zhu, K., Miller, F. R., Barder, T. J., Lubman, D. M., *J. Mass Spectrom.* 2004, *39*, 770–780.

[26] Zheng, S., O'Neil, K. A., Barder, T. J., Lubman, D. M., *BioTechniques* 2003, *35*, 1202–1212.

[27] Zhou, F., Johnston, M. V., *Electrophoresis* 2005, *26*, 1383–1388.

[28] Harris, W. A., Reilly, J. P., *Anal. Chem.* 2002, *74*, 4410–4416.

[29] Zheng, S., Yoo, C., Miller, F. R., Huber, C. G., Lubman, D. M., *Anal. Chem.* 2006, *78*, 5198–5204.

[30] Kato, M., Sakai-Kato, K., Jin, H., Kubota, K. *et al.*, *Anal. Chem.* 2004, *76*, 1896–1902.

[31] Zou, H., Huang, X., Ye, M., Luo, Q., *J. Chromatogr. A* 2002, *954*, 5–32.

[32] Hemstrom, P., Nordborg, A., Irgum, K., Svec, F., Fréchet, J. M. J., *J. Sep. Sci.* 2006, *29*, 25–32.

[33] Premstaller, A., Oberacher, H., Walcher, W., Timperio, A. M. *et al.*, *Anal. Chem.* 2001, *73*, 2390–2396.

[34] Walcher, W., Toll, H., Ingendoh, A., Huber, C. G., *J. Chromatogr. A* 2004, *1053*, 107–117.

[35] Lopez, M. F., *Electrophoresis* 2000, *21*, 1082–1093.

[36] Ericson, C., Phung, Q. T., Horn, D. M., Peters, E. C. *et al.*, *Anal. Chem.* 2003, *75*, 2309–2315.

[37] Shaheen, N. J., *Gastroenterology* 2005, *128*, 1554–1566.

[38] Devesa, S. S., Blot, W. J., Fraumeni, J. F., *Am. Cancer Soc.* 1998, *83*, 2049–2053.

[39] Onnerfjord, P., Nilsson, J., Wallman, L., Laurell, T., Marko-Varga, G., *Anal. Chem.* 1998, *70*, 4755–4760.

[40] Zhang, B., McDonald, C., Li, L., *Anal. Chem.* 2004, *76*, 992–1001.

[41] Sluyterman, L. A. A., Elgersma, O., *J. Chromatogr.* 1978, *150*, 17–30.

[42] Zhu, K., Zhao, J., Lubman, D. M., Miller, F. R., Barder, T. J., *Anal. Chem.* 2005, *77*, 2745–2755.

[43] Oberacher, H., Premstaller, A., Huber, C. G., *J. Chromatogr. A* 2004, *1030*, 201–208.

[44] Gabel, D., Kasche, V., *Biochem. Biophys. Res. Commun.* 1972, *48*, 1011–1018.

[45] Lolli, G., Johnson, L. N., *Cell Cycle* 2005, *4*, 572–577.

[46] Parra, J. L., Buxade, M., Proud, C. G., *J. Biol. Chem.* 2005, *280*, 37623–37633.

[47] Miliotis, T., Kjellstrom, S., Nillson, J., Laurell, T. *et al.*, *J. Mass Spectrom.* 2000, *35*, 369–377.

[48] Daniel, J. M., Laiko, V. V., Doroshenko, V. M., Zenobi, R., *Anal. Bioanal. Chem.* 2005, *383*, 895–902.