Identification of In-Gel Digested Proteins by Complementary Peptide Mass Fingerprinting and Tandem Mass Spectrometry Data Obtained on an Electrospray Ionization Quadrupole Time-of-...

ARTICLE *in* ANALYTICAL CHEMISTRY · APRIL 2000

Impact Factor: 5.64 · DOI: 10.1021/ac990937m · Source: PubMed

5 AUTHORS, INCLUDING:

Christoph Borchers
University of Victoria
258 PUBLICATIONS **8,949** CITATIONS
SEE PROFILE

Jochen Peter
Solaton Power GmbH
12 PUBLICATIONS **376** CITATIONS
SEE PROFILE

Mark C Hall
Purdue University
39 PUBLICATIONS **1,581** CITATIONS
SEE PROFILE

Kenneth Tomer
National Institute of Environmental Health S...
354 PUBLICATIONS **11,952** CITATIONS
SEE PROFILE

Anal. Chem. **2000,** *72,* 1163−1168

Identification of In-Gel Digested Proteins by Complementary Peptide Mass Fingerprinting and Tandem Mass Spectrometry Data Obtained on an Electrospray Ionization Quadrupole Time-of-Flight Mass Spectrometer

Christoph Borchers, Jochen F. Peter, Mark C. Hall, Thomas A. Kunkel, and Kenneth B. Tomer*

Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709

The present study reports a procedure developed for the identification of SDS-polyacrylamide gel electrophoretically separated proteins using an electrospray ionization quadrupole time-of-flight mass spectrometer (Q-TOF MS) equipped with pressurized sample introduction. It is based on in-gel digestion of the proteins without previous reduction/alkylation and on the capability of the Q-TOF MS to provide data suitable for peptide mass fingerprinting database searches and for tandem mass spectrometry (MS/MS) database searches (sequence tags). Omitting the reduction/alkylation step reduces sample contamination and sample loss, resulting in increased sensitivity. Omitting this step can leave disulfide-connected peptides in the analyte that can lead to misleading or ambiguous results from the peptide mass fingerprinting database search. This uncertainty, however, is overcome by MS/MS analysis of the peptides. Furthermore, the two complementary MS approaches increase the accuracy of the assignment of the unknown protein. This procedure is thus, highly sensitive, accurate, and rapid. In combination with pressurized nanospray sample introduction, it is suitable for automated sample handling. Here, we apply this approach to identify protein contaminants observed during the purification of the yeast DNA mismatch repair protein Mlh1.

The combination of protein separation by one- or two-dimensional polyacrylamide gel electrophoresis with mass spectrometric analysis of proteins digested enzymatically in-gel, followed by protein database searching is a very efficient tool for protein identification in complex biological systems at the nanogram level.[1] The most commonly used mass spectrometric identification methods rely on the highly accurate mass measurement of peptides obtained by specific enzymatic cleavage of a protein, which is commonly named "peptide mass fingerprinting".[2−4] Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) is the simplest, fastest, and most sensitive MS approach. Furthermore, MALDI-MS is suitable for automated sample handling, which is an important issue for highly demanding high-throughput protein characterization. Peptide mapping-based methodologies, however, allow only for the identification of a protein that is already in the protein database. It assumes that the detected peptides originate from a single, nonmodified polypeptide or from a component of simple protein mixtures,[5] and it is not compatible with the identification of proteins by correlation with expressed sequence tag databases.[6] A complementary mass spectrometric approach based on electrospray tandem mass spectrometric analysis (ESI-MS/MS)[7] has been developed in Mann's group[8] that permits the identification of such proteins. In this approach, the partial amino acid sequence of a single peptide is obtained by MS/MS. The associated masses of this "sequence tag" combined with the mass of the peptide provide a highly specific probe for sequence database searches. In contrast to the use of MALDI-MS, however, sample preparation for electrospray analysis requires a concentration and desalting step, due to the lower sensitivity and lower tolerance to salt and buffer contamination of the ESI technique. This desalting/concentration step is an absolute requirement if, to obtain single polypeptides, in-gel reduction/alkylation of the protein is performed and/or the nanoelectrospray (nano-ESI) is used.[9] In general, the nano-ESI technique is difficult because the sample is usually introduced via a fine, fragile glass needle which requires a small sample volume (∼0.5 μL). Additional drawbacks to this technique include the absolute loss of sample during the desalting/concentration step, the time required for sample handling, and the use of needles, which is not well suited for automation.

* Corresponding author: National Institute of Environmental Health Sciences, P.O. Box 12233 MD F0-03, Research Triangle Park, NC 27709; (phone) (919) 541-1966; (fax) (919) 541-0220; (e-mail) tomer@niehs.nih.gov.

(1) Jensen, O. N.; Larsen, M. R.; Roepstorff, P. *Proteins Suppl.* **1998**, *2*, 74−89.
(2) Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5011−5015.
(3) Pappin, D. J. C.; Højrup, P.; Bleasby, A. J. *Curr. Biol.* **1993**, *3*, 327−332.
(4) Mann, M.; Højrup, P.; Roepstorff, P. *Biol. Mass Spectrom.* **1993**, *22*, 338−345.
(5) Jensen, O. N.; Podtelejnikow, A. V.; Mann M. *Anal. Chem.* **1997**, *69*, 4741−4750
(6) Mann, M. *Trends Biol. Sci.* **1996**, *21*, 494−495.
(7) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6233−6237.
(8) Mann, M.; Wilm, M. S. *Anal. Chem.* **1994**, *66*, 4390−4399.
(9) Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *66*, 1−8.

10.1021/ac990937m
Published on Web 02/11/2000

On-line coupling of liquid chromatography (LC) with electrospray MS/MS can overcome these disadvantages. Miniaturized LC/MS interface systems have been developed and used for the identification of in-gel digested proteins.[10−12] Furthermore, a fast Fourier transform algorithm has been developed that correlates the experimental data with the data predicted from the protein sequences in databases.[13,14] Recently, automated protein identification by LC/MS/MS and database searching has been demonstrated at the 5 pmol level.[15] The automation of this combined method, however, is limited by the robustness and sample turnaround of the LC system.

More recently, a microfabricated, electrokinetic device coupled with an electrospray ionization quadrupole time-of-flight (ESI-Q-TOF) mass spectrometer has been developed that is suitable for automated protein identification at the 290 fmol/μL protein level.[16] The enhanced resolution and mass accuracy of the Q-TOF mass spectrometer and its capability for MS/MS analysis[17,18] offers the unique combination of peptide mass fingerprinting and sequence tag database searching. The microfabricated device, however, is not commercially available, its construction is complicated, and the use of electrokinetics for sample transportation is always accompanied by potential discrimination effects. Furthermore, the successful use of this system has been demonstrated only for the identification of a standard solution of a myoglobin tryptic digest.

Here, we report the application of nanoflow sample introduction combined with an ESI-Q-TOF mass spectrometer suitable for automated identification of in-gel digested proteins at the subpicomole protein level. The sample introduction system combines a commercially available nanospray interface equipped with a simple pressurized sample infusion and is suitable for automation. In addition, this system does not require a desalting/concentration step following the in-gel digestion procedure. The sensitivity, minimal protein and time consumption, and fidelity of this approach are described for identifying protein contaminants in a preparation of the yeast DNA mismatch repair protein Mlh1.

MATERIALS AND METHOD

Protein Expression and Purification. Mlh1 was introduced into protease-deficient yeast cells on a multicopy plasmid as a translational fusion with the yeast *VMA1* intein and the *Bacillus circulans* chitin binding domain (CBD) obtained from the Impact I purification system (New England Biolabs, Beverly, MA). Overexpression of the Mlh1-intein-CBD fusion from a *GAL1* promoter was induced during early log-phase growth by transferring the cells to rich media containing 2% galactose. Cells were harvested during late log phase and lysed using a Bead Beater (Biospec Products, Bartlesville, OK), and the clarified extract was applied to a 5 mL chitin column (New England Biolabs). Following extensive washing with buffer containing 500 mM NaCl, 100 mM 2-mercaptoethanol was added to the column to induce the intein-catalyzed cleavage reaction. After a 24 h incubation period at 4 °C, free Mlh1 was eluted from the column.

SDS-Polyacrylamide Gel Electrophoresis and Densitometry. For the SDS-polyacrylamide gel electrophoresis, a 1 mm thick 10% Tris−glycine precast gel from Novex, Inc. (San Diego, CA) was used. The sample was boiled for 2 min in the presence of 60 mM SDS and 570 mM 2-mercaptoethanol prior to loading. For the molecular weight markers, the standard broad-range Bio-Rad MW marker (Bio-Rad, Hercules, Ca) was used. The electrophoresis was performed at 125 V. The gel was fixed for 20 min in 10% acetic acid (Mallinckrodt Baker, Paris, KY)/25% 2-propanol (J. T. Baker, Phillipsburg, NJ) and then stained with 0.006% Coomassie Brilliant Blue R-250 (Bio-Rad) in 10% acetic acid for several hours. The gel was destained in 10% acetic acid and stored at 4 °C in this solution prior to in-gel digestion. The protein amount of the gel bands was quantified by densitometric analyses using the image system from Eastman Kodak (New Haven, CT). For calibration of the densitometer, the standard molecular weight marker proteins were used. The loading amount of the molecular weight markers was 1 μg of each standard protein.

In-Gel Digestion. The in-gel digestion was performed using a procedure that has been previously described.[19] Unless otherwise noted, the various steps of the procedure were performed at room temperature and all incubation steps were performed under shaking. The protein double band of interest in the SDS-gel was excised, cut into small pieces, transferred into an 1.5 mL polyethylene sample vial, and washed with 100 μL of deionizied water. The SDS-gel pieces were destained by adding 50 μL of acetonitrile and incubating for 5 min. The supernatant was discarded and 50 μL of acetonitrile/50 mM NH_4HCO_3 (1:1) was added to the SDS-gel pieces. This step was performed several times until the SDS-gel pieces was completely destained. After drying the gel pieces by lyophilization, the SDS-gel pieces were rehydrated with 100 μL of 50 mM NH_4HCO_3 (pH 8.5). The supernatant was removed and the SDS-gel pieces were incubated with 100 μL of acetonitrile for 5 min. This step was repeated until the SDS-gel pieces were completely white. After removing the supernatant, the SDS-gel pieces were lyophilized, rehydrated, and incubated with 20 μL of ice-cold TPCK-treated trypsin (Worthington Biochemical Corp., Freehold, NJ) solution (10 ng/mL in 50 mM NH_4HCO_3) for 30 min in an ice bath. The enzyme solution was removed, and the SDS-gel pieces were incubated in 50 μL of 25 mM NH_4HCO_3 for 18 h at 37° C. Acetonitrile (50 μL) was added to the solution. After a 10 min incubation, the supernatant was removed and transferred to a 750 μL polyethylene sample vial. The SDS-gel pieces were incubated for 10 min with 30 μL of water, following the addition of 50 μL of acetonitrile. The supernatants were removed, combined with the solvents in the 750 μL polyethylene sample vial, lyophilized, and redissolved for mass spectrometric analysis in 20 μL of 50% acetonitile/0.1% formic acid (J. T. Baker Inc., Phillipsburg, MA).

(10) Hunt, D. F.; Henderson, R. A.; Shabanowitz, J.; et al. *Science* **1992**, *255*, 1261−1263.
(11) Davis, M. T.; Lee, T. D. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 194−201.
(12) Qin, J. H. C.; Zhang, X. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 209−216.
(13) Eng, J. K.; McCormack, A. L.; Yates, J. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976−989.
(14) Dongre, A. R.; Eng, J. K.; Yates, J. R. *Trends Biotechnol* **1997**, *15*, 418−425
(15) Ducret, A.; Van Oostveen, I.; Eng, J.; Yates, J. R.; Aebersold, R. *Protein Sci.* **1998**, *7*, 706−719.
(16) Figeys, D.; Lock, C.; Taylor, L.; Aebersold, R. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1435−1444.
(17) Morris, H. R.; Paxton, T.; Dell, A.; et al. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 889−896.
(18) Shevchenko, A.; Chernuchevich, I.; Ens, W.; et al. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1015−1024.
(19) Wilm, M.; Shevchenko, A.; Houthaeve, T.; Breit, S.; Schweigerer, L.; Foptsis, T.; Mann, M. *Nature* **1996**, *379*, 466−9.

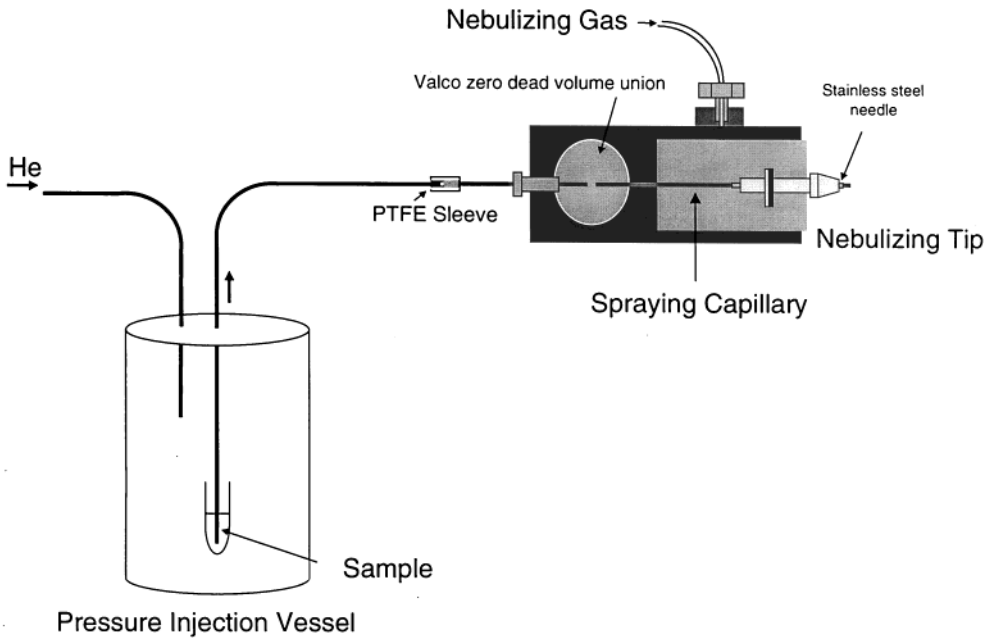

Figure 1. Scheme of the nanoflow sample introduction system coupled to an electrospray ionization mass spectrometer. The sample introduction system combines a commercially available nanospray interface equipped with simple pressurized sample infusion system.

Mass Spectrometry. All experiments were performed on a Micromass Q-TOF mass spectrometer (Micromass, Altrincham, U.K.), equipped with a nebulized nanoelectrospray z-spray source. The capillary voltage used was 2800 V, and the cone voltage was 22 V. In the nebulized nanospray source, nitrogen gas is introduced coaxially as the nebulizing gas and the sample is introduced by flow injection analysis (Figure 1). The injections were performed by placing the end of a 50 μm i.d. fused-silica into the sample vial that is inside a stainless steel pressure vessel.[20] Samples were forced into the capillary by pressurizing with helium. The helium pressure was adjusted to give a flow rate of 150−200 nL/min (nanoflow mode). The end of the capillary tubing was connected to the entrance tubing of the nebulized nanospray interface with a poly(tetrafluoroethylene) (PTFE) sleeve.

Database Searching. The peptide mass data and the tandem-MS data were searched using the program MS-Fit and MS-Tag, respectively, against a nonredundant database (NCBlnr) located at the University of California San Francisco (Internet address: http://prospector.ucsf.edu).

RESULTS AND DISCUSSION

The combination of a nanoflow sample introduction system and an ESI-Q-TOF mass spectrometer has been successfully applied to the analysis of peptides and to the identification of modification sites in proteins by using a variety of mass spectrometric approaches, including precursor ion scanning.[21] In this study, we investigated the analytical criteria of this combined system for identifying in-gel digested proteins at low-picomole/subpicomole protein levels. The sample introduction system combines a commercially available nanospray interface equipped with a simple, pressurized sample infusion system that uses helium for applying pressure to a pressure injection vessel containing the sample (Figure 1). The flow rate can be easily adjusted by pressure regulation of the helium gas tank and was approximately 150−200 nL/min for the experiments described. A stable electrospray was obtained for several hours using water or ammonium bicarbonate buffer (5−50 mM) containing 0.1% formic acid and 10−90% methanol or acetonitrile as organic solvents (data not shown).

In this report, we demonstrate the capability of this approach for the identification of in-gel digested proteins, by applying this approach to the identification of minor components observed in the SDS-gel electrophoretic analyses of the yeast DNA mismatch repair protein Mlh1 after its overexpression in yeast cells and purification by affinity chromatography. The Mlh1 was overexpressed and purified to study its structure and biochemical properties and to obtain more details about its role in DNA mismatch repair. The SDS−PAGE separation shows a broad, intense band with an apparent molecular weight around 80 000, corresponding to Mlh1, and additional less intense protein bands at lower molecular weights (Figure 2). The protein bands in the molecular weight range of 40 000−50 000 were identified as fragments of Mlh1 by Western blot analysis (data not shown). The double band with an apparent MW of 70 000, however, did not react with the anti-Mlh1 antibody, suggesting that it may contain protein unrelated to Mlh1. These protein impurities will be present in the Mlh1 preparation being studied. It is therefore important to know what these impurities are in order to avoid confusing their biochemical and biological functions with those of authentic Mlh1. The double protein band, therefore, was excised, in-gel digested, and analyzed by MS and MS/MS using the previously described nanoflow sample introduction combined with an ESI-Q-TOF mass spectrometer. The double protein band was excised as a single band because the single bands could not be excised separately.

(20) Deterding, L. J.; Moseley, M. A.; Tomer, K. B.; Jorgenson, J. W. *Anal. Chem.* **1989**, *61*, 2504.
(21) Borchers, C.; Parker, C. E.; Deterding, L. J.; Tomer, K. B. *J. Chromatogr., A* **1999,** *854*, 119−130.

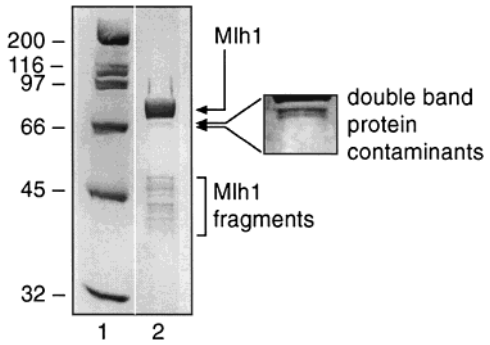

Figure 2. SDS-polyacrylamide gel electrophoresis of Mlh1 after its expression and purification by affinity chromatography. Lane 1: standard molecular weight markers. Lane 2: The SDS-gel shows a broad band at an apparent MW of 78 000 corresponding to Mlh1, minor bands at MW range 40 000−50000 that were identified as Mlh1 fragments by Western blot analysis, and a less intense double band at 70 kDa. The inset shows a magnified exposure of the double band which was excised as a single gel piece because the single bands could not be excised separately. The double band most likely corresponds to protein contaminants not structurally related to Mlh1 because the double band was not detectable by Western blot analysis.

To determine the sensitivity of this analytical approach, the proteins in the double band were first analyzed using densitometry and were found to be present at the 65 (0.9 pmol) and 95 ng (1.3 pmol) levels for the lower and higher molecular weight bands, respectively. The digestion of the two proteins in the gel piece was done using a modification of the in-gel procedure of Wilm et al.[19] Of particular importance, our modification of Wilm's procedure eliminated the alkylation and reduction step; instead, the gel pieces were vigorously washed with 50 mM NH_4HCO_3 (pH 8.5). This procedural modification is advantageous because the sample does not require cleanup prior to analysis, thereby reducing sample loss. After in-gel digestion, the sample was dissolved in 20 μL of 50% acetonitrile/0.1% formic acid for analysis by ESI-Q-TOF, corresponding to a concentration of 8 ng/μL (110 fmol/μL) protein, assuming that the proteins were completely digested and that complete recovery of the in-gel digested proteins was achieved.

To obtain a mass spectrum (Figure 3) of the in-gel digested double-banded protein, approximately 1.5−2.0 μL of the sample was nanosprayed for 10 min. This corresponds to a consumption of 165−220 fmol of protein assuming 100% recovery following in-gel digestion. Due to the high resolution of the ESI-Q-TOF mass spectrometer, the charge state of the ions could be determined from the mass spectrum (Figure 3), allowing for the unambiguous determination of the monoisotopic molecular weight of the peptides. The resolution obtained via ESI-Q-TOF analysis is shown in the insets in Figure 3 and demonstrates that high isotopic resolution was achieved not only for the relatively abundant peaks with an ion distribution characteristic of peptides (see triply charged monoisotopic ion at m/z 596.72) but could also be obtained for lower abundance peaks and peaks in the higher m/z range (see triply charged ion at m/z 990.15, which corresponds to a peptide containing a ^{13}C-atom). Overall, 38 peaks with an ion distribution characteristic of peptides (i.e. multiply charged ions) were found in the mass spectrum of the digest. There are still a few ions, however, to which a charge state could not be unequivocally assigned because of overlapping ion signals. The mass accuracy of the mass spectrometer in this experiment was determined from the doubly charged autoproteolytic tryptic peptides T3 (aa 50−69) and T8 (aa 126−136) of trypsin (doubly charged ions of m/z 1082.06 and of m/z 577.35, respectively) and was calculated to be 32 and 65 ppm for these peptides, respectively.

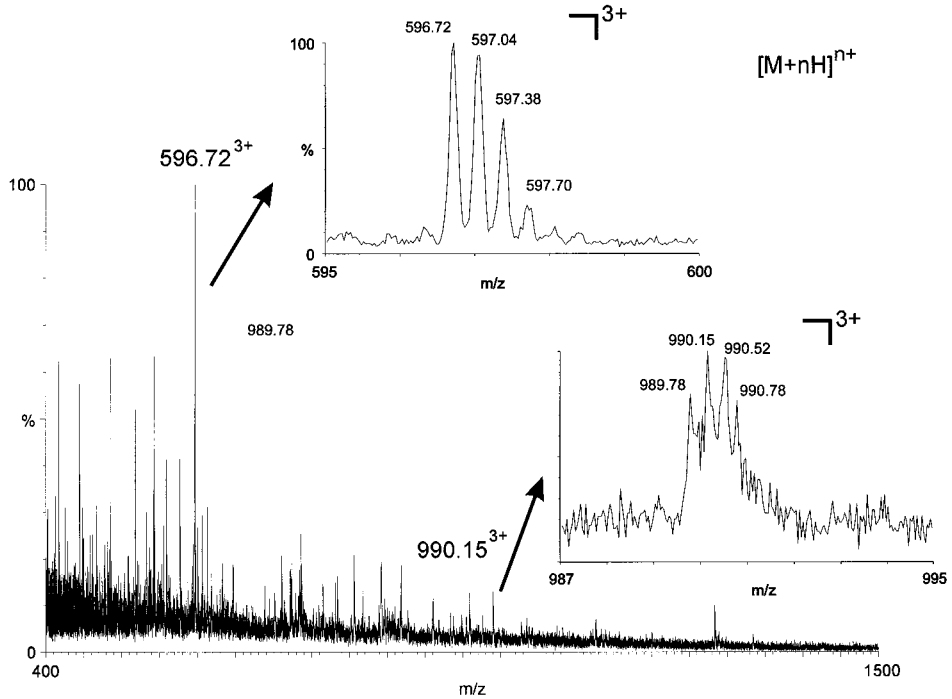

Figure 3. Electrospray mass spectrometric analysis of the in-gel digested double band at 70 kDa in the SDS−PAGE gel of Mlh1 after its expression and purification. The MS spectrum was acquired for 10 min by nanospraying approximately 1.5−2.0 μL of sample, which corresponds to a consumption of 165−220 fmol of protein, assuming 100% protein recovery of the entire in-gel digestion procedure. The insets demonstrate the resolution of the triply charged ion signals at *m*/*z* 596.72 and at *m*/*z* 990.15.

Table 1. Results from the Sequence Database Searching Using the Program MS-Tag (University of California San Francisco)[a]

rank	database entry no.	protein name	protein MW	sequence
1	L22015	Ssa1	69 767.7	(R)^{170}IINEPTAAAIAYGLDKK^{186}(G)
1	X12926	Ssa1	69 657.6	(R)^{170}IINEPTAAAIAYGLDKK^{186}(G)
1	M17583	Ssa1-fragment	37 445.3	(R)^{170}IINEPTAAAIAYGLDKK^{186}(G)
1	Z73129	Ssa2	69 470.3	(R)^{170}IINEPTAAAIAYGLDKK^{186}(G)
1	Z35836	Ssa3	70 547.1	(R)^{170}IINEPTAAAIAYGLDKK^{186}(G)
1	Z26879	Ssa3	68 545.0	(R)^{170}IINEPTAAAIAYGLDKK^{186}(G)
1	U18839	Ssa4	69 651.4	(R)^{170}IINEPTAAAIAYGLDKK^{186}(S)

[a] For the database searching of proteins from the species *Saccharomyces cerevisiae*, the monoisotopic masses of the parent ion (*m*/*z* 1788.31) and the monoisotopic masses of the resultant fragment ions (*m*/*z* 1318.82, 1447.79, 1561.88, and 1674.97) from the MS/MS analysis (Figure 4b) were entered. All entered masses correspond to monoprotonated ion signals.

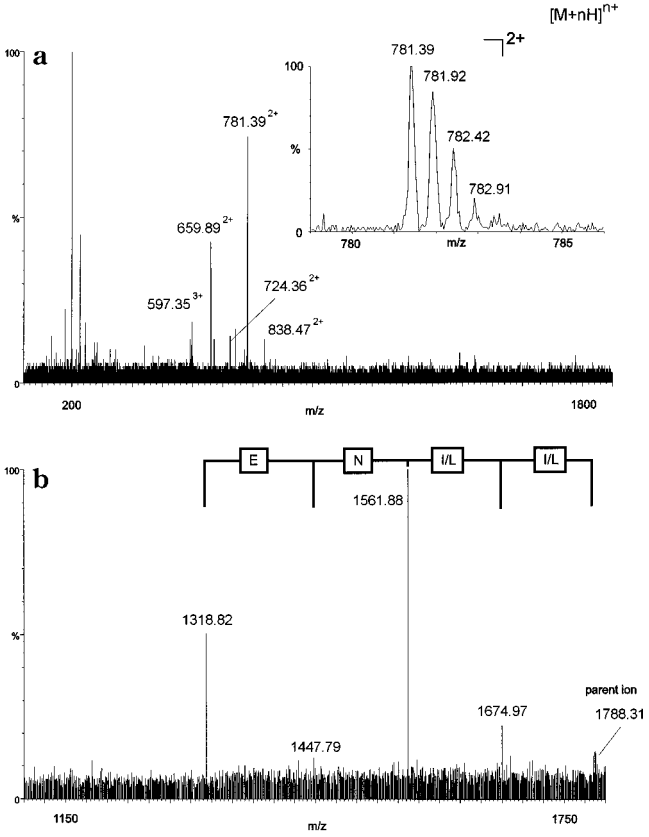

Figure 4. Electrospray tandem mass spectrometric analysis of the triply charged ion signal at *m*/*z* 596.72 from the MS analysis of the in-gel digested double band. The MS spectrum was acquired for 5 min by nanospraying approximately 0.75−1.0 μL of sample, corresponding to a consumption of 83−110 fmol of protein, assuming 100% protein recovery of the entire in-gel digested protein. (a) Aquired MS/MS spectrum showing singly and doubly charged fragment ions. The inset demonstrates the isotopic resolution of the doubly charged ion signal at *m*/*z* 781.39. (b) Conversion of the multiply charged ion signal into a singly charged monoistopic MS/MS spectrum. The differences between the fragment ions correspond to the amino acid sequence I/L-I/L-N-E.

To obtain sequence information, an MS/MS analysis was performed on the most abundant ion (596.72 *m*/*z*) in the mass spectrum. The MS/MS spectrum, which was obtained by using 0.75−1.0 μL of the sample (protein consumption 83−110 fmol) shows singly and doubly charged fragment ions (Figure 4a). As in the MS approach, high resolution also was obtained in the MS/MS-mode. The isotopic resolution allows for the conversion of the multiply charged ion spectrum into a monoisotopic spectrum in order to more easily determine the protein sequence (Figure 4b). The monoprotonated masses of these monoisotopic fragment ions (*m*/*z* 1318.82, 1447.79, 1561.88, and 1674.97) and the monoprotonated mass of the monoisotopic parent ion (*m*/*z* 1788.31) were entered into the UCSF online protein database and a search was done for proteins from the species *Saccharomyces cerevisiae*. The protein database search yielded a peptide whose sequence exactly matched the entered parent ion and the entered fragment ion masses and corresponded to the y-ion series, y_{13}−y_{16}, of the identified peptide using the nomenclature of Roepstorff and Fohlman (Table 1).[22] The identified peptide was found in seven database entries and corresponded to four very closely related proteins in the HSP70 heat shock protein family. Entries L22015 and L12926 are separate entries for the full length Ssa1 protein, which differ by three amino acids, and entry M17583 is a fragment of Ssa1. Entries Z35836 and Z26879 are separate entries for Ssa3 protein, with the latter lacking the 21 C-terminal amino acids present in the former. Of the seven prospective proteins, the Ssa1 fragment M17583 was excluded because its 37 445 Da mass did not correspond to the 70 000 Da mass range established by SDS-gel analysis.

To distinguish among the four Ssa possibilities, the peptide mass fingerprinting methodology was applied. The high resolution and mass accuracy of the ESI-Q-TOF MS allows the accurate determination of the molecular weights of 33 peptides (the autoprotolytic peptides of trypsin are excluded). These peptides were identified manually but can be identified automatically with recent software development. The results from the database search are summarized in Table 2 and illustrate that the six proteins with a MW around 70 000 and which contain the peptide sequenced by MS/MS (Table 1) are among the eight highest ranking proteins. The top three ranking proteins, X12926 (Ssa1), L22015 (Ssa1), and Z73129 (Ssa2), have high MOWSE scores[3] of 3.11e+006 to 2.64e+005, which are approximately 10 000−100 000 higher than the fourth ranking protein, thereby demonstrating the high accuracy of the protein assignment. The peptide coverage of these three proteins ranged from 44 to 33%; however, cysteine-containing peptides were not detected. As mentioned above, the two Ssa1 entries differ by three amino acids, corresponding to a sequence identity of 99.5%. Ssa2 displays 97.7% sequence identity with X12926 (Ssa1). Despite the degree of identity among the three sequences, two ion signals were found in the MS analysis

(22) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601.

Table 2. Results from the Peptide Mapping-Based Database Searches, Using the Program MS-Fit (University of California San Francisco)[a]

rank	MOWSE score	masses matched, no. (%)	protein MW	peptide coverage (%)	database entry no.	protein name
1	3.11e+006	27/33 (81)	69657.6	44	X12926	Ssa1
2	7.55e+005	25/33 (75)	69767.7	39	L22015	Ssa1
3	2.64e+005	21/33 (63)	69470.3	33	Z73129	Ssa2
4	117	10/33 (30)	68545.0	17	Z26879	Ssa3
5	110	11/33 (30)	69070.1	26	Z47815	Cik1
6	97.2	11/33 (33)	65572.5	26	Z72774	Rfg2
7	93.2	10/33 (30)	70547.1	16	Z35836	Ssa3
8	83.2	12/33 (36)	69651.4	27	U18839	Ssa4

[a] The monoisotopic masses of the 33 monoprotononated ion signals that resulted from the MS analysis (Figure 3) of the in-gel digested double-banded proteins were entered into the program. Database searching was performed for proteins in the MW range of 65 000−75 000 from the species *S. cerevisiae*.

that only correspond to peptides of X12926 (Ssa1). The detection of these unique peptides demonstrates that Ssa1 is a contaminant of the Mlh1 preparation and that the X12926 database entry for Ssa1 corresponds better to this contaminant than the L22015 entry. Ssa2 or Ssa3 may also be a contaminant of the Mlh1 preparation but no ions unique to either Ssa2 or Ssa3 were observed, and therefore, no further MS/MS data were acquired. The predicted molecular masses of Ssa1 and Ssa2 are within 300 Da of each other (Ssa3 differs by 900 Da), which could explain the presence of the double band seen by SDS−PAGE.

CONCLUSION

The reported combination of nanoflow pressurized sample introduction with an ESI-Q-TOF mass spectrometer for analysis of proteins enzymatically in-gel digested followed by protein database searching is a powerful technique for the identification of proteins, because of the distinct advantages it offers over traditional MS techniques in proteome analysis. This technique increases the efficiency of sample analysis, requiring only 15 min to obtain MS and MS/MS data. This time of analysis corresponds to a sample volume of ~3 μL with a total protein consumption of ~250−300 fmol, assuming that the 2 pmol of the proteins (double band) applied to the gel was completely recovered. This corresponds to a single protein concentration of ~50 fmol/μL and is comparable to the sensitivity achieved by an ESI-Q-TOF MS using the microfabricated device (290 fmol/μL myglobin in solution). The sensitivity is also comparable with the nanoelectrospray vial technique (160 fmol/μL in-gel digest BSA), although, in that study only 80 fmol of BSA was applied to the gel. The design of the sample introduction device allows a reduction of the sample volume to 4 μL, suggesting that protein identification can be achieved at the 200 fmol level applied to the gel. The MS and MS/MS spectra obtained showed signals with isotopic resolution and mass accuracy of better than 100 ppm over the entire m/z range (up to m/z 1500), for both singly and multiply charged ions (including quadruply charged ions). On the basis of the high resolution and mass accuracy of the Q-TOF-MS, the MS and MS/MS analyses provide data suitable for peptide mass fingerprinting database searches and sequence tag database searches. These two complementary MS approaches increase the accuracy of the identification of the unknown protein. Peptide coverages of up to 44% were achieved, but cysteine-containing peptides were not detected. The absence of the observation of cysteine-containing peptides may be due to omission of the alkylation step. This is one potential drawback to our approach.

The advantages of this sample introduction system are as follows: (i) commercial availability of the nanospray interface, (ii) simple construction of the pressurized sample infusion system (iii), no discrimination effects between analytes with pressurized sample infusion, and (iv) the possibility of automation with this interface. The automation could be performed with a device that might resemble the pressurized sample introduction systems already used by capillary electrophoresis instruments.

This technique also shows great promise for the identification of proteins in protein mixtures. For this purpose, MS sample analysis could be performed in combination with an intelligent computer software capable of database searching selected ions from the MS spectrum.[23] This combinatorial approach would enable investigators to confirm the assignment of peptide masses through an automated computer interface that would select ions from the MS spectrum that matched peptides in the protein database and then perform MS/MS analysis upon these ions for sequence confirmation. Unknown ions would also be analyzed automatically by MS/MS to obtain sequencing information that would enable the complete identification of proteins with high accuracy.

ACKNOWLEDGMENT

The authors thank Robert N. Wine, LT/EPT/NIEHS, for performing the densitometric analysis.

Received for review August 17, 1999. Accepted December 20, 1999.

AC990937M

(23) Yates, J. R. *Electrophoresis* **1998**, *19*, 893−900.