Anna Shevchenko[1]
Alexander Loboda[2, 3]
Werner Ens[2]
Burkhart Schraven[4]
Kenneth G. Standing[2]
Andrej Shevchenko[1]

[1]Peptide and Protein Group, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany
[2]Department of Physics and Astronomy, University of Manitoba, Winnipeg, Manitoba, Canada
[3]MDS Sciex, Concord, Canada
[4]Institute for Immunology, University of Heidelberg, Heidelberg, Germany

Archived polyacrylamide gels as a resource for proteome characterization by mass spectrometry

Mass spectrometry was applied to identify protein spots excised from an archived two-dimensional polyacrylamide gel that had been dried and stored for eight years at room temperature. All proteins were successfully identified. Detailed characterization of protein digests by matrix-assisted laser desorption/ionization (MALDI) peptide mapping, nanoelectrospray tandem mass spectrometry and MALDI-quadrupole time-of-flight mass spectrometry revealed no evidence of protein degradation or modifications that could hamper identification of proteins in a sequence database. The experiment with a model protein demonstrated that the pattern of tryptic peptides and the yield of individual peptides were not noticeably changed in the in-gel digest of the archived protein spot compared to the digest of the spot excised from a fresh gel. Thus, the characterization of "archived proteomes" has the potential to advance proteomic research without repeating "wet" biochemistry experiments, that had been perfected in the laboratory years ago.

Keywords: Protein identification / Proteomics / Protein modifications / Mass spectrometry
EL 4285

1 Introduction

For many years molecular biology experiments have yielded femtomole amounts of proteins that were separated by one-dimensional or two-dimensional polyacrylamide gel electrophoresis. However, no analytical technique has been available to utilize this valuable information efficiently. Internal sequencing of proteins by Edman degradation requires about 10 pmol of a protein present on a gel [1]. This technique was (and largely still is) expensive and slow, and, therefore, can only be applied to a small selection of interesting proteins. Furthermore, public sequence databases contained a modest number of protein entries up to a couple of years ago, so the characterization of any novel protein inevitably demanded its *de novo* sequencing and cloning using degenerate oligonucleotide probes and PCR. Polyacrylamide gels containing femtomole amount of proteins per spot or band have not been amenable to rapid and comprehensive characterization. Such analytical gels were typically dried on a plastic or paper support und were kept for years in folders simply as reference images. Changes in protein expression patterns could still be evaluated by visual inspection of images of these high-quality two-dimensional gels, even in the absence of adequate protein sequencing capacities [2, 3].

In recent years, genomic sequencing has led to an unprecedented growth of sequence databases. Complete genomes of important model organisms have become available [4–6] and the human genome is scheduled for completion by the year 2003 [7, 8]. Partial sequences of many more yet uncharacterized genes have become accessible *via* expressed sequence tag (EST) databases [9]. Therefore it has become possible to depart from tedious and laborious *de novo* protein sequencing and cloning of genes [10, 11] by identifying cognate sequences in protein and EST databases [12–15].

During this period mass spectrometric technology has also made significant progress. It is now possible to characterize proteomes of prokaryotic and eukaryotic organisms at the subpicomole level (reviewed in [16, 17]). The entire analytical routine, starting from in-gel digestion of proteins [18] to the acquisition of MALDI spectra and database searching [19, 20], has been automated. Hybrid quadrupole (Qq) time-of-flight mass spectrometers [21, 22] have improved the quality of tandem mass spectrometric data and have enabled automated LC-MS/MS runs and rapid off-line [23–25], and on-line [26] data processing and database searching. Finally, the emerging technique of MALDI-QqTOF-MS has given prospects for combining MALDI mass fingerprinting and tandem mass spectrometric investigation of peptide precursor ions in an efficient manner [27, 29].

Correspondence: Dr. Andrej Shevchenko, Peptide & Protein Group, European Molecular Biology Laboratory (EMBL), Meyerhofstraße 1, 69012 Heidelberg, Germany
E-mail: shevchenko@EMBL-Heidelberg.de
Fax: +49-6221-387 306

Abbreviations: **DHB**, 2,5-dihydroxybenzoic acid; **HCCA**, α-cyano-4-hydroxycinnamic acid; **QqTOF**, quadrupole time-of-flight

0173-0835/01/0606–1194 $17.50+.50/0

The identification of proteins from archived gels has now become an appealing prospect because of the progress in MS and the expansion of sequence databases. Rasmussen *et al.* [30] demonstrated that proteins could be extracted from pooled protein spots excised from a large number of archived Coomassie-stained gels. Although more than 50 proteins were identified *via* Edman degradation of tryptic peptides, only a few peptides per protein were recovered and many of the reported sequences were ambiguous and/or incomplete. More recently it has been demonstrated by us [31] and later by Matsumoto and Komori [32] that proteins from archived gels could also be identified by in-gel trypsinolysis followed by MALDI peptide mass mapping.

However, it has still not been clear how well proteins embedded in a gel matrix survive years of storage. In the latter experiments, a number of intense peptide peaks in a MALDI peptide map of the in-gel digest of an archived protein was not assigned to the protein sequence even at the mass tolerance exceeding 400 ppm [32]. Furthermore, the sequence coverage of peptide maps acquired from in-gel digested archived proteins was significantly lower than from the corresponding spots excised from fresh gels. Taken together, these observations suggest that degradation of archived proteins might occur [30, 32]. However, preserving proteins intact is paramount for deciphering protein mixtures [33], for identifying proteins in an EST database [34] and for the characterization of protein modifications [35] and polymorphism [36]. We therefore set out to characterize the protein modifications, which might occur during archiving and storage of gels. We also sought to determine whether archiving of gels affects the pattern and the yield of peptides recovered after enzymatic cleavage of proteins at the picomole level.

2 Materials and methods

2.1 Materials and reagents

All major chemicals were purchased from Sigma (Sigma Chemicals, St. Louis, MO, USA), and were of analytical grade. HPLC-grade water, methanol and acetonitrile (LabScan, Dublin, Ireland) were used for mass spectrometric analysis and preparation of digests. $H_2{}^{18}O$ (Cambridge Isotope Laboratories, Andover, MA, USA) was purified by microdistillation as described [11]. Human proteins immunoprecipitated by polyclonal antibodies against the CD45 receptor had been isolated as described [37], separated by two-dimensional gel electrophoresis and visualized by Coomassie staining. The gel had been subsequently dried between two sheets of a cellophane film and stored since July, 1992 in a folder at room temperature.

2.2 Identification of proteins

2.2.1 *In-gel* digestion of proteins

Protein spots were excised from the dry gel and the gel plugs were rehydrated in water for 30 min at room temperature. The gel plugs were then rinsed twice with water, after which the cellophane film was easily detached from the gel slab by using a scalpel. The gel plugs were subsequently cut in *ca.* 1 × 1 mm pieces, transferred to an 0.5 mL Eppendorf test tube, rinsed with water and the proteins were in-gel digested with bovine trypsin (unmodified, sequencing grade; Roche Diagnostics, Mannheim, Germany) as described [11, 38].

2.2.2 Mass spectrometric analysis and database searching

Proteins were identified by MALDI peptide mapping and nanoES-MS/MS combined in a layered approach [39]. Briefly, a 0.3–0.5 μL aliquot of the supernatant was withdrawn from the in-gel digest of the protein band and analyzed by MALDI-MS as described [38, 40]. MALDI-MS measurements were performed on a modified Reflex mass spectrometer (Bruker Daltonics, Bremen, Germany). If further analysis by nanoES-MS/MS was required, the gel pieces were extracted with 5% formic acid and acetonitrile. The combined extracts were dried in a vacuum centrifuge, redissolved in 5% formic acid and the unfractionated digest was analyzed as described [41]. The nanoES-MS/MS measurements were performed on an API III triple quadrupole mass spectrometer (MDS Sciex, Concord, ON, Canada). MS and MS/MS analysis of in-gel digests by MALDI-QqTOF-MS was performed as described [27] on a prototype MALDI-QqTOF mass spectrometer [28] built at the University of Manitoba in collaboration with MDS Sciex. Database searching was performed against a comprehensive nonredundant sequence database using the PaptideSearch V. 3.0 software developed at EMBL. No limitations on protein molecular weights and species of origin were imposed. A second pass searching routine [33] was applied to reveal masses of the modified peptides in the MALDI peptide mass maps. Database searching using tandem mass spectra acquired on the MALDI-QqTOF instrument was performed *via* Internet by the MS-Tag program (http://prospector.ucsf.edu/) at the UCSF Mass Spectrometry Facility server as described [27]. If database searching produced no hit, short stretches of peptide sequence were deduced and assembled into a peptide sequence

tag; then error-tolerant database searching [42] was performed using PaptideSearch V. 3.0. Alternatively, the PredictSequence routine (BioMultiview 1.4; MDS Sciex) was used for partial interpretation of the tandem mass spectra [28].

2.3 Quantification of tryptic peptides by MALDI-MS

Two pmol of BSA were loaded onto separate landes of a one-dimensional polyacrylamide gel. After electrophoresis the gel was stained with Coomassie blue and then cut into two parts. One part was dried on a plastic support using a GelAir Dryer (BioRad, Hercules, CA, USA) and another part was kept in destaining solution (45% water/45% methanol/10% acetic acid v/v/v). Two BSA bands were excised both from the wet gel and from the archived gel and in-gel digested in parallel. Two microliters of an in-solution tryptic digest of BSA in 25 mM ammonium bicarbonate buffer in $H_2{}^{18}O$ with a concentration of 0.14 pmol/μL were mixed with 2 μL aliquots withdrawn from the above experimental digests and analyzed by MALDI-MS. Upon digestion in $H_2{}^{18}O$, BSA peptides incorporated one or two ^{18}O-atoms into their *C*-terminal carboxyl groups [43], and these ^{18}O-labeled peptides were used as internal standards. The relative concentration of the individual digestion products was calculated as a ratio of the intensity of the monoisotopic peak of the unlabeled peptide and the intensity of the monoisotopic peak of the corresponding labeled peptide that incorporated two ^{18}O atoms [44, 45].

3 Results and discussion

3.1 In-gel digestion of proteins from archived gels

No substantial modifications of the conventional in-gel digestion protocol [11, 38] were required to analyze proteins from the archived gel. To avoid crushing of the gel slab during rehydration, we excised spots from the dry gel first and then rehydrated the gel plugs separately. After a brief incubation in water, pieces of cellophane could be easily detached from the gel plug. Importantly pieces of dust, threads and traces of human fingerprints (frequent sources of keratin contamination) remain at the surface of the film during rehydration and can be subsequently removed together with the film after the rehydration step is completed. For gels dried on a paper support, it is often difficult to remove debris of the paper from the reaction mixture, but residual paper filaments do not noticeably affect the recovery of digestion products (data not shown). However, their presence complicates handling of samples because of frequent clogging of the pipette tips.

3.2 Identification of archived proteins

In the present work we address two major issues: (i) whether archived proteins were severely modified or damaged during long storage; (ii) if such modifications occurred, how can they be accounted for so that the certainty of protein identification is not affected? To this end, 14 Coomassie-stained protein spots were excised from the archived gel (Fig. 1) and digested in-gel with trypsin. The digests were analyzed by MALDI-MS, nanoES-MS/MS [46] and by MALDI-QqTOF-MS. A combination of these methods allowed us to identify the proteins and to establish the identity of almost all of the prominent peaks observed in the peptide maps, thus providing a detailed account of the protein modifications detected.

First MALDI peptide maps were acquired and examined in detail. A second pass searching algorithm [33] was employed to account for common protein modifications such as oxidation of methionine residues and acrylamidation of cysteine residues. If intense peptide ions were observed that could not be matched to the sequence of a protein already identified, the recovered pool of tryptic peptides was analyzed by nanoES-MS/MS and MALDI-QqTOF-MS. Proteins from all 14 spots excised from the gel were identified by MALDI peptide mapping (Table 1) using special database searching settings. For example, a peptide mass map was acquired from an in-gel digest of a 38 kDa protein from spot 12 (Fig. 2), and the masses of

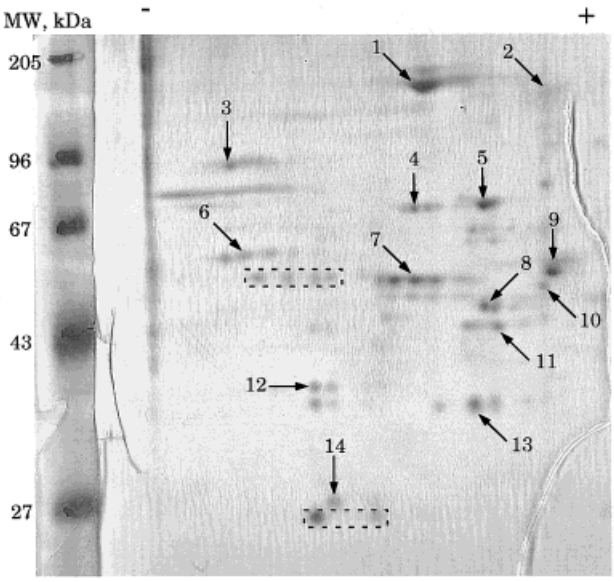

Figure 1. Scanned picture of a Coomassie-stained archived gel. Protein spots labeled with arrows were analyzed by MS. Identification of these proteins is presented in Table 1. Spots originating from antibodies, used in immunoaffinity purification are boxed.

Table 1. Identification of proteins from the archived gel

Spot No.	Protein	Acc. number, submission date	Peptide masses, matched/ unmatched[a)]	Sequence coverage (%)	Met residues[b)] Σ/M(ox)/ M&M(ox)	Cys residues[c)] Σ/C(acr)/ C&C(acr)	Trp residues[d)] Σ/W(ox)/ W&W(ox)
1	SWI/SNF complex 155/170 kDa	Q92923;Q92922 from 02.97	32 / 2[e)]	25 / 12	13 / 4 / 1[e)]	6 / 4 / 1[e)]	4 / 2 / 0[e)]
2	Fibronectin receptor CD29	P05556 from 11.88	9 / 1	12	2 / 2 / 0	2 / 0 / 1	–
3	DEAD box protein 1	Q92499 from 02.97	37 / 0	52	6 / 5 / 1	8 / 2 / 6	3 / 1 / 1
4	Bovine serum albumin	P02769 from 07.86	20 / 0	42	2 / 2 / 0	9 / 8 / 0	–
5	Heat shock 71 kDa protein	P11142 from 07.89	21 / 0	37	7 / 4 / 2	1 / 0 / 0	1 / 1 / 0
6	Hypothetical 55.2 kDa protein	Q9Y3I0 from 11.99	15 / 0	30	6 / 5 / 0	1 / 0 / 1	3 / 1 / 1
7	Ribonucleoprotein H	P31943 from 07.93	19 / 2	45	4 / 2 / 2	3 / 2 / 0	1 / 0 / 0
8	Actin-related protein BAF53A	O96019 from 05.99	16 / 1	40	4 / 1 / 1	1 / 1 / 0	2 / 0 / 0
9	BAF57 protein	O43539 from 06.98	29 / 1[e)]	55	8 / 1 / 0	–	4 / 2 / 1[e)]
	TGF protein[f)]	Q92734 from 02.97		10	2 / 2 / 0	–	
10	Disulfide isomerase	Q99778 from 05.97	17 / 1	50	–	1 / 1 / 0	2 / 0 / 1
11	γ-Actin	P02571 from 07.86	11 / 0	45	4 / 4 / 0	1 / 1 / 0	3 / 1 / 0
12	Ribonucleoprotein	Q9Y4J5 from 11.99	9 / 2	41	5 / 5 / 0	–	–
13	Transducin β-chain 1	P04901 from 08.87	16 / 0[e)]	35	1 / 0 / 1[e)]	6 / 6 / 0[e)]	2 / 0 / 2[e)]
	Transducin β-chain 2	P11016 from 07.89		38			
14	Homebox Prox 1 protein	Q9Y224 from 11.99	21 / 1	65	1 / 0 / 0	2 / 0 / 2	3 / 1 / 1

a) Number of unmatched peaks having the intensity over 25% of the intensity of the most abundant peak in the spectrum
b) Methionine residues in detected peptides. Σ, total number of methionine residues; M(ox), detected in sulfoxide form only; M&M(ox), detected both in native and sulfoxide forms
c) Cysteine residues in detected peptides. Σ, total number of cysteine residues; C(acr), detected in *S*-acrylamide form only; M&M(ox), detected both in *S* – acetamide and *S*-acrylamide forms
d) Tryptophane residues in detected peptides. Σ, total number of tryptophane residues; W(ox), detected only in oxidized forms; W&W(ox), detected both in native und oxidized forms
e) Calculated in total for both identified proteins
f) Protein identified by nanoES-MS/MS

14 peptide peaks were used for searching a database with a mass tolerance better than 100 ppm. The search produced 20 protein hits (including three human proteins of various molecular weights) that matched more than four peptide masses. Thus, no unequivocal identification could be reached.

Database searching was then repeated with the same set of peptide masses but assuming that methionine residues are present in the oxidized form (methionine sulfoxides) and cystein residues are present in acrylamidated form (cysteine *S*-acrylamide). The 37 kDa human Q9Y4J5 ribonucleoprotein then appeared at the top of the list with nine matching peptides covering more than 50% of the protein sequence and only two peptide ions of low abundance were unmatched (Fig. 2). An 80 kDa plant protein that matched only four peptide masses occupied the second position in the list of hits and no human proteins other than Q9Y4J5 appeared in the list. MALDI peptide maps of the archived proteins (Table 1) contain a substantial proportion of methionine – contaning peptides only in the methionine sulfoxide form, cysteine-containing peptides only in the cysteine-*S*-acrylamide form and tryptophane residues only in dioxidized form (see also the discussion below). Thus accounting for these forms in database searching sharply increases the chances of producing a statistically reliable hit.

In several cases (spots 9, 13 and 14) the MALDI peptide maps contained intense signals that could not be assigned to the sequences of the identified proteins by second pass searching. In these cases, tryptic peptides were extracted from a gel matrix and half of the extract was analyzed by nanoES-MS/MS. In the case of spot 9, nanoES-MS/MS sequencing identified another component – the Q92734 TGF protein, comigrating with the main protein component BAF57 (Table 1). However, only three peptides originating from the TGF protein were found retrospectively in the MALDI spectrum of the spot 9 digest (data not shown), which may explain why MALDI-

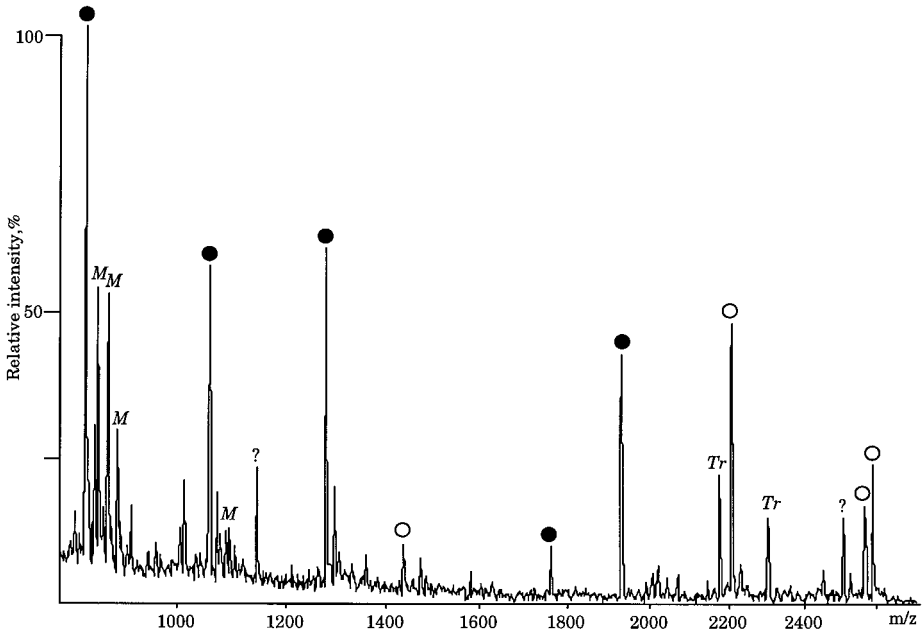

Figure 2. Identification of the protein in spot 12 by MALDI-MS. Five peptides (peaks labeled with filled circles) were matched to the sequence of the Q9Y4J5 protein when the mass of native methionine was used in database searching. However, additional four peptides were matched (peaks labeled with unfilled circles) when methionine masses were set to the mass of methionine sulfoxide thus producing a very strong hit. Trypsin autolysis products are designated with Tr. M, matrix peaks. Two peptide peaks (labeled with a question mark) were not matched.

MS failed to identify it. The reason for such exceptionally low sequence coverage remains unclear.

In general, the two ionization methods applied – MALDI-MS and nanoES-MS/MS do not generate identical patterns of peptide ions when the same digest is analyzed. As demonstrated above for spot 9, this presents only a minor problem for identification of proteins in mixtures. However, it may become a major limitation if the identity of particular masses in the MALDI spectrum is sought, so such a combination of methods does not readily lend itself to accurate accounting of protein modifications.

3.3 Comprehensive characterization of peptide mass maps by MALDI-QqTOF-MS

It can, therefore, be a significant advantage if any particular peak observed in a MALDI peptide map of the digest can be characterized independently by a low energy CID tandem mass spectrum. Therefore, we applied MALDI-QqTOF-MS to analyze the digests in which conventional MALDI-MS detected intense peaks that were not assigned to the identified protein, and nanoES-MS/MS analysis did not reveal the presence of another comigrating protein.

The analysis of spot 13 is discussed as a case study. Database searching using the MALDI peptide map (Fig. 3A) hit two human proteins both having a molecular mass close to the expected 40 kDa: P04901 transducin beta-chain 1 (gene name GNB1) and P11016 transducin beta-chain 2 (gene name GNB2), which produced partly overlapping peptide mass maps. When second pass searching was separately applied to GNB1 and GNB2, almost all the intense peptide peaks detected in the MALDI spectrum could be assigned to either of these proteins (Fig. 3A and Table 2), except for two intense peptide peaks having *m/z* 1728.78 and 1770.81 (labeled as T_{17} and T_{18} in Fig. 3A). NanoES-MS/MS sequencing confirmed that spot 13 is a mixture of GNB1 and GNB2 (Fig. 3B, Table 2). Sequences of two peptide prescursors T_7 and T_8 (Table 2) differed from the sequence of the corresponding peptide in a database (T_6) by a single amino acid residue. However, the observed profile of peptide peaks looked strikingly different from the MALDI map, and no apparent candidates for T_{17} and T_{18} were revealed.

Therefore, the second half of the digest was analyzed by MALDI-QqTOF-MS (Fig. 3C). MALDI-QqTOF-MS required a “cold” matrix like 2,5-dihydroxybenzoic acid (DHB) [28] for best results, although conventional MALDI had demonstrated best resolution and sensitivity when α-cyano-4-hydroxycinnamic acid (HCCA) doped with nitrocellulose was applied [38, 47]. Thus, it is not surprising that the relative intensities of peptide ions in the spectra obtained by conventional MALDI and MALDI-QqTOF-MS were somewhat different, although in general similar profiles of peptide ions were observed (Fig. 4A and C). The unidentified peptide ions (*m/z* 1728.854 and *m/z* 1770.874) were also detected as fairly intense peaks. Upon partial manual interpretation and error-tolerant

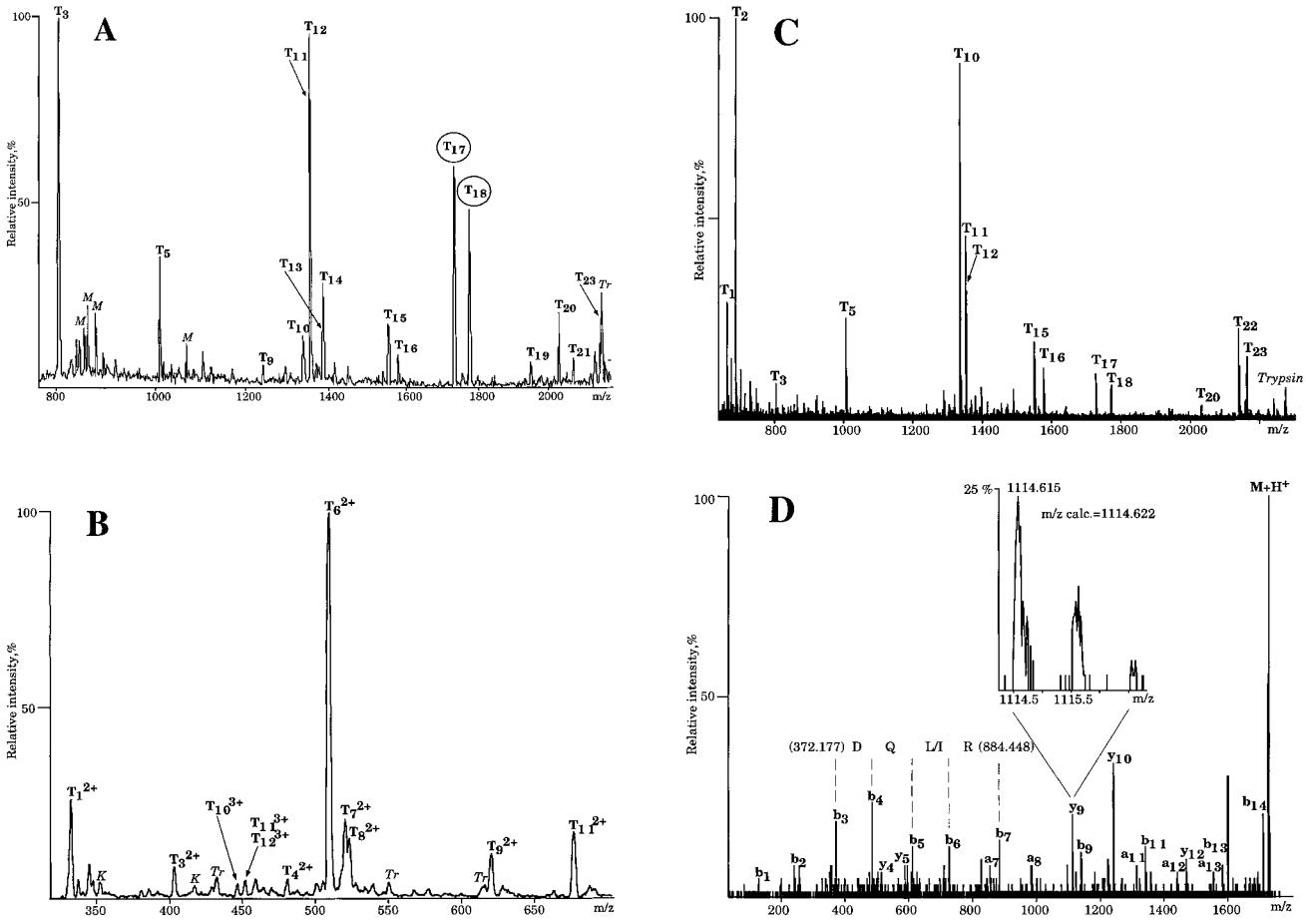

Figure 3. (A) MALDI peptide map acquired from the digest of spot 13. Peptide ions labeled in the panel were matched to the sequences of GNB1 and/or GNB2 (Table 2). Peaks designated with T_{17} and T_{18} (circled) were not matched. (B) nanoES spectrum of the same digest. All peaks labeled in the spectrum were assigned (Table 2). (C) MALDI-QqTOF spectrum of the digest. Peaks of the peptides T_{17} and T_{18} were observed in the spectrum and their MS/MS spectra were subsequently acquired. (D) The tandem mass spectrum acquired from the precursor ion with *m/z* 1728.853. A stretch of the peptide sequence was deduced from the series of b-ions and assembled into a sequence tag together with masses of the corresponding fragments. Upon error tolerant searching the identity of T_{17} was established (Table 2).

database searching, tandem mass spectra acquired from T_{17} and T_{18} (Fig. 4D) revealed that these peptides originated *via* *N*-terminal processing of GNB1 and GNB2, respectively. The *N*-terminal methionine residues had been removed and the resulting peptides were acetylated (Table 2). Thus, despite of the complexity of a mixture of two highly homologous *N*-terminally processed proteins, almost all prominent peptide ions in the peptide maps were finally assigned to the protein sequences.

Similarly, the combination of MALDI-MS, nanoES-MS/MS and MALDI-QqTOF-MS allowed us to obtain representative peptide mass maps with 40% (on average) sequence coverage and to establish the identity of the vast majority of the prominent peptide peaks (Table 1). Less than 3% of the total number of peptide peaks observed in the digests of 14 proteins remained unassigned because of insufficient quality of the acquired MS/MS data.

3.4 Observed protein modifications

Three major types of protein modification were detected in the analysis of the 14 protein spots: oxidation of methionine residues, acrylamidation of cysteine residues and double oxidation of tryptophane residues (Table 1). We note here that such modifications are frequently encountered in sequencing of proteins from conventional

Table 2. Identification of peptides in the digest of spot 13 by MALDI-MS, nanoES-MS/MS and MALDI-QqTOF-MS[a)]

Peptide	Residues	Sequence	MALDI	MALDI-QqTOF	ES-MS	GNB 1 or 2
T_1	252–256	LFDLR		●	●	GNB1/2
T_2	210–214	LWDVR		●		GNB1/2
T_3	90–96	VHAIPLR	●	●	●	GNB1/2
T_4	179–197	(M)SLSLAPDTR			●	GNB1
T_5	305–314	AGVLAGHDNR	●	●		GNB1/2
T_6	69–78	LLVSASQDGK			●	GNB1/2
T_7	69–78	LLVSASQ**H**GK			●	— —
T_8	69–78	LLV**D**ASQDGK			●	— —
T_9	198–209	LFVSGAC(acr)DASAK	●		●	GNB1
T_{10}	58–68	IYAMHWGTDSR	●	●	●	GNB1/2
T_{11}	58–68	IYAM(ox)HWGTDSR	●	●	●	GNB1/2
T_{12}	79–89	LIIWDSYTTNK	●	●	●	GNB1/2
T_{13}	58–68	IYAM(ox)HW(ox^2)GTDSR	●			GNB1/2
T_{14}	79–89	LIIW(ox^2)DSYTTNK	●			GNB1/2
T_{15}	138–150	ELAGHTGYLSC(acr)C(acr)R	●	●		GNB1
T_{16}	138–150	ELPGHTGYLSC(acr)C(acr)R	●	●		GNB2
T_{17}	2–15	(*N*-acetvl)SELDQLRQEAEQLK	●	●		GNB1
T_{18}	2–15	(*N*-acetyl)SELEQLRQEAEQLK	●	●		GNB2
T_{19}	24–42	AC(acr)GDSTLTQITAGLDPVGR	●			GNB2
T_{20}	24–42	AC(acr)ADATLSQITNNIDPVGR	●	●		GNB1
T_{21}	23–42	KAC(acr)GDSTLTQITAGLDPVGR	●			GNB2
T_{22}	79–96	LIIWDSYTTNKVHAIPLR		●		GNB1/2
T_{23}	284–301	KAC(acr)ADATLSQITNNIDPVGR	●	●		GNB2

a) Detected peptides are designated with filled circles.
C(acr), cysteine- S-acrylamide; M(ox), methionine sulfoxide; W(ox^2), dioxidized tryptophane

polyacrylamide gels [38, 48–50], so they are not caused specifically by gel archiving and storage. Nevertheless, peptides containing a monooxidized methionine residue or a dioxidized tryptophane residue were observed as abundant ions even though unmodified forms of such peptides were barely detected, contrary to the patterns usually observed in peptide mass maps of proteins from conventional gels [50].

A MALDI peptide map of BSA obtained from a freshly prepared gel and from spot 4 from the archived gel allowed an interesting comparison (Table 3). Both spectra were acquired on the same instrument and the samples were prepared using similar in-gel digestion and probe preparation recipes. In both spectra only a few minor peaks were not matched to the BSA sequence (data not shown). An almost identical pool of the peptides, without methionine and cysteine residues, was observed. However, many more peptides containing *S*-acrylamidated cysteine residues were detected in the digest of archived BSA.

Although a Coomassie-stained gel was used in the present work, sequencing of proteins from archived silver-stained gels may also be possible provided that the staining protocol did not employ chemicals, which might modify proteins covalently [38, 51]. We concluded that archiving and long time storage of polyacrylamide gels did not affect peptide map profiles and did not hamper subsequent protein identification. We also did not find any credible evidence that uncommon protein modification or protein decomposition occurred.

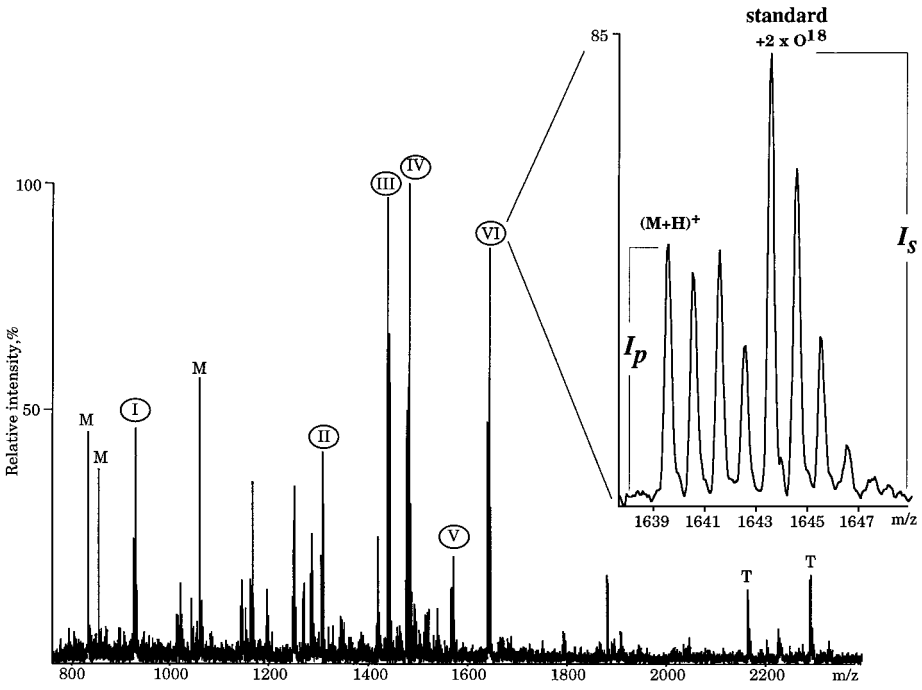

Figure 4. The MALDI peptide map of the in-gel tryptic digest of 2 pmol of BSA spiked with the BSA digest in $H_2{}^{18}O$. Numbers designate peptide peaks used for quantification. M, peaks originating from the HCCA matrix; T, trypsin autolysis products. A blowup of the isotopic cluster of the peptide peak with *m/z* 1639.93 is shown in the inset. The monoisotopic peak of the unlabeled peptide is designated as $(M+H)^+$. The peak of the isotopicaly labeled peptide which incorporated two ^{18}O atoms was used as an internal standard. The relative concentration of the unlabeled peptide (R_c) was calculated as $R_c = I_p/I_s$.

Table 3. Peptides detected by MALDI-MS in the digests of BSA from the conventional gel and from the archived gel (spot 4)[a)]

m/z	Residues	Sequence	Conventional gel	Archived gel
847.50	242–248	LSQKFPK	•	
927.49	161–167	YLYEIAR	•	•
1083.59	161–168	YLYEIARR	•	•
1142.71	548–557	KQTALVELLK	•	
1163.63	66-75	LVNELTEFAK	•	•
1193.60	25–34	DTHKSEIAHR	•	•
1249.62	35–44	FKDLGEEHFK	•	•
1283.71	361–371	HPEYAVSVLLR	•	
1305.71	402–412	HLVDEPQNLIK	•	•
1415.68	569–580	TVM(ox)ENFVAFVDK		•
1433.70	89–100	SLHTLFGDELC(acr)K		•
1439.81	360–371	RHPEYAVSVLLR	•	•
1479.79	421–433	LGEYGFQNALIVR	•	•
1553.83	483–495	LC(acr)VLHEKTPVSEK		•
1567.74	347–359	DAFLGSFLYEYSR	•	•
1639.93	437–451	KVPQVSTPTLVEVSR	•	•
1754.84	469–482	M(ox)PC(acr)TEDYLSLILNR		•
1809.84	387–401	DDPHAC(acr)YSTVFDKLK		•
1842.86	372–386	LAKEYEATLEEC(acr)C(acr)AK		•
1894.93	508–523	RPC(acr)FSALTPDETYVPK		•
2034.06	588-607	EACFAVEGPKLVVSTQTALA		•
2045.03	168–183	RHPYFYAPELLYYANK		•
2555.18	118–138	QEPERNEC(acr) FLSHKDDSPDLPK		•

a) Detected peptides are designated with filled circles. Abbreviations are the same as in Table 2.

Table 4. Quantification of the recovery of BSA tryptic peptides from archived and from fresh gels

Peak No.[a)]	*m/z*	Sequence	R_c wet / R_c archived (%)[b)]
I	927.49	YLYEIAR	103
II	1305.71	HLVDEPQNLIK	104
III	1439.81	RHPEYAVSVLLR	100
IV	1479.79	LGEYGFQNALIVR	104
V	1567.74	DAFLGSFLYEYSR	83
VI	1639.93	KVPQVSTPTLVEVSR	100

a) As presented in Fig. 4

b) The ratio of the relative concentration (R_c) of the peptide in the in-gel digest of the band excised from the wet gel to the R_c of the same peptide in the digest of the band excised from the archived gel. To determine R_c, five MALDI spectra were acquired from each of the samples and the results were averaged.

3.5 Quantitative evaluation of the yield of tryptic peptides

We further tested whether the yield of tryptic peptides may be affected by archiving of gels. The experiments presented above suggested that proteins embedded in a matrix of polyacrylamide gel largely remained intact after eight years of storage. Therefore, we reasoned that in order to determine whether archiving affects the recovery of peptides we might use dried gels that had not been stored for a long period of time. To this end we applied MALDI-MS to compare the yield of tryptic peptides of BSA recovered after in-gel digestion of Coomassie-stained bands, excised from a wet gel and from a gel which had been dried on a cellophane support 48 h prior to digestion. As internal standards for the quantification of peptides we used ^{18}O-labeled peptides of BSA. Comparison of the relative yields of six tryptic peptides (Fig. 4 and Table 4) demonstrated that the recovery of peptides was not affected by gel archiving.

4 Concluding remarks

This work has demonstrated that proteins from archived polyacrylamide gels can be successfully identified by MS, even if present in mixtures or if only a small number of peptides has been recovered. Almost any peak in the MALDI peptide maps of these proteins could be matched to a corresponding protein sequence and the modifications observed are the same as those commonly encountered in sequencing of gel separated proteins. Archiving and storage did not noticeably affect the pattern of peptide mass maps or the recovery of individual peptides. We concluded that archived gels are fully amenable for the low level protein characterization and accurate accounting of *in vivo* protein modifications and of protein polymorphism is also possible.

Most of the sequences of proteins identified in the course of the project had been only recently submitted to a database (Table 1), so *de novo* sequencing and cloning would have been required to characterize 9 out of 14 spots when the proteins were isolated eight years ago. However, the availability of full length sequences enabled us to identify the proteins by MALDI-MS, a technique that allows very high throughput and completely automated protein identification. Archived gels may now present a valuable resource for research in molecular biology and medicine, since the characterization of proteins and (or) protein modifications can now be achieved rapidly and at low cost.

The work at Manitoba was supported by grants from the Natural Sciences and Engineering Research Council of Canada, from MDS Sciex, and from the US National Institute of Health (GM59240). We are grateful to the other members of the TOF Laboratory of the University of Manitoba and to members of the Peptide and Protein Group in EMBL for technical support and useful discussions and to Dr. W. Zachariae (MPI MCBG, Dresden) for critical reading of the mansucript.

Received July 30, 2000

5 References

[1] Erdjument-Bromage, H., Lui, M., Sabatini, D. M., Snyder, S. H., Tempst, P., *Prot. Sci.* 1995, *3,* 2435–2446.

[2] Boucherie, H., Dujardin, G., Kermorgant, M., Monribot, C., Slonimski, P., Perrot, M., *Yeast* 1995, *11,* 601–613.

[3] Celis, J. E. (Ed.), *Electrophoresis* 1994, *15,* 1349–1490.

[4] Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., Oliver, S. B., *Science* 1996, *274,* 546–567.

[5] The *C. elegans* Sequencing Consortium, *Science* 1998, *282,* 2012–2018.

[6] Blattner, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., Shao, Y., *Science* 1997, *277,* 1453–1474.

[7] Waterston, R., Sulston, J. E., *Science* 1998, *282,* 53–54.

[8] Collins, F. S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R., Walters, L., *Science* 1998, *282,* 682–689.

[9] Boguski, M. S., Lowe, T. M. J., Tolstoshev, C. M., *Nature Genet.* 1993, *4,* 332–333.

[10] Shevchenko, A., Wilm, M., Mann, M., *J. Prot. Chem.* 1997, *16,* 481–490.

[11] Shevchenko, A., Chernushevich, I., Wilm, M., Mann, M., in: Chapman, J. R. (Ed.), *Mass Spectrometry of Proteins and Peptides,* Humana Press, Totowa, NJ 2000, Vol. 146, pp. 1–16.

[12] Yates, J. R., Eng, J. K., McCormack, A. L., *Anal. Chem.* 1995, *67,* 3202–3210.

[13] Mann, M., *Trends Biochem. Sci.* 1996, *21,* 494–495.

[14] Pandey, A., Lewitter, F., *Trends Biochem. Sci.* 1999, *24,* 276–280.

[15] Neubauer, G., Wilm, M., *Curr. Opin. Mol. Ther.* 1999, *1,* 695–701.

[16] Lahm, H. W., Langen, H., *Electrophoresis* 2000, *21,* 2105–2114.

[17] Pandey, A., Mann, M., *Nature* 2000, *405,* 837–846.

[18] Houthaeve, T., Gausepohl, H., Mann, M., Ashman, K., *FEBS Lett.* 1995, *376,* 91–94.

[19] Jensen, O. N., Mortensen, P., Vorm, O., Mann, M., *Anal. Chem.* 1997, *69,* 1706–1714.

[20] Berndt, P., Hobohm, U., Langen, H., *Electrophoresis* 1999, *20,* 3521–3526.

[21] Morris, H. R., Paxton, T., Dell, A., Langhorn, J., Berg, M., Bordoli, R. S., Hoyes, J., Bateman, R. H., *Rapid Commun. Mass Spectrom.* 1996, *10,* 889–896.

[22] Shevchenko, A., Chernushevich, I., Ens, W., Standing, K. G., Thomson, B., Wilm, M., Mann, M., *Rapid Commun. Mass Spectrom.* 1997, *11,* 1015–1024.

[23] Ducret, A., Van Oostveen, I., Eng, J. K., Yates, J. R., III, Aebersold, R., *Prot. Sci.* 1998, *7,* 706–719.

[24] Moseley, M. A., in: Burlingame, M. A., Carr, S. A., Baldwin, M. A. (Eds.), *Mass Spectrometry in Biology and Medicine,* Humana Press, Totowa, NJ 1999, pp. 179–198.

[25] Perkins, D. N., Pappin, D. J., Creasy, D. M., Cottrell, J. S., *Electrophoresis* 1999, *20,* 3551–3567.

[26] Rappsilber, J., Andersen, J. S., Mann, M., Budin, K., Pedersen, C., Mortensen, P., *Proc. 48th ASMS Conf. Mass Spectrom. Allied Topics,* Long Beach, CA 2000, pp. 165–166.

[27] Shevchenko, A., Loboda, A., Shevchenko, A., Ens, W., Standing, K. G., *Anal. Chem.* 2000, *72,* 2132–2141.

[28] Loboda, A. V., Krutchinsky, A. N., Bromirski, M., Ens, W., Standing, K. G., *Rapid Commun. Mass Spectrom.* 2000, *14,* 1047–1057.

[29] Krutchinsky, A. N., Zhang, W., Chait, B. T., *J. Am. Soc. Mass Spectrom.* 2000, *11,* 493–504.

[30] Rasmussen, H. H., Van Damme, J., Puype, M., Gesser, B., Celis, J. E., Vandekerckhove, J., *Electrophoresis* 1991, *12,* 873–882.

[31] Shevchenko, A., Shevchenko, A., Schraven, B., Mann, M., *Proc. 46th ASMS Conf. Mass Spectrom. Allied Topics,* Orlando, FL 1998, p. 238.

[32] Matsumoto, H., Komori, N., *Anal. Biochem.* 1999, *270,* 176–179.

[33] Jensen, O. N., Podtelejnikov, A. V., Mann, M., *Anal. Chem.* 1997, *69,* 4741–4750.

[34] Neubauer, G., King, A., Rappsilber, J., Calvio, C., Watson, M., Ajuh, P., Sleeman, J., Lamond, A., Mann, M., *Nature Genet.* 1998, *20,* 46–50.

[35] Zhang, X., Herring, C. J., Romano, P. R., Szczepanowska, J., Brzeska, H., Hinnebusch, A. G., Qin, J., *Anal. Chem.* 1998, *70,* 2050–2059.

[36] Gatlin, C. L., Eng, J. K., Cross, S. T., Detter, J. C., Yates, J. R., *Anal Chem.* 2000, *72,* 757–763.

[37] Schraven, B., Schoenhaut, D., Bruyns, E., Koretzky, G., Eckerskorn, C., Wallich, R., Kirchgessner, H., Sakorafas, P., Labkovsky, B., Ratnofsky, S., Meuer, S., *J. Biol. Chem.* 1994, *269,* 29102–29111.

[38] Shevchenko, A., Wilm, M., Vorm, O., Mann, M., *Anal. Chem.* 1996, *68,* 850–858.

[39] Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko, A., Boucherie, H., Mann, M., *Proc. Natl. Acad. Sci. USA* 1996, *93,* 14440–14445.

[40] Jensen, O. N., Podtelejnikov, A., Mann, M., *Rapid Commun. Mass Spectrom.* 1996, *10,* 1371–1378.

[41] Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., Mann, M., *Nature* 1996, *379,* 466–469.

[42] Mann, M., Wilm, M., *Anal. Chem.* 1994, *66,* 4390–4399.

[43] Schnölzer, M., Jedrzejewski, P., Lehmann, W. D., *Electrophoresis* 1996, *17,* 945–953.

[44] Shevchenko, A., Wilm, M., Shevchenko, A., *Proc. 48th ASMS Conf. Mass Spectrom. Allied Topics,* Long Beach, CA 2000, pp. 859–860.

[45] Mirgorodskaya, O. A., Kozmin, Y. P., Titov, M. I., Korner, R., Sonksen, C. P., Roepstorff, P., *Rapid Commun. Mass Spectrom.* 2000, *14,* 1226–1232.

[46] Wilm, M., Mann, M., *Anal. Chem.* 1996, *66,* 1–8.

[47] Vorm, O., Roepstorff, P., Mann, M., *Anal. Chem.* 1994, *66,* 3281–3287.

[48] Clauser, K. R., Hall, S. C., Smith, D. M., Webb, J. W., Andrews, L. E., Tran, H. M., Epstein, L. B., Burlingame, A. L., *Proc. Natl. Acad. Sci. USA* 1995, *92,* 5072–5076.

[49] Swiderek, K. M., Davis, M. T., Lee, T. D., *Electrophoresis* 1998, *19,* 989–997.

[50] Thiede, B., Lamer, S., Mattow, J., Siejak, F., Dimmler, C., Rudel, T., Jungblut, P. R., *Rapid. Commun. Mass Spectrom.* 2000, *14,* 496–502.

[51] Moertz, E., Krogh, T. N., Vorum, H., Görg, A., *Proc. 48th ASMS Conf. Mass Spectrom Allied Topics,* Long Beach, CA 2000, pp. 1115–1116.