International Journal of Mass Spectrometry 307 (2011) 16–27

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

Investigation of Se-containing proteins in *Bertholletia excelsa* H.B.K. (Brazil nuts) by ICPMS, MALDI-MS and LC–ESI-MS methods

Sarath B. Jayasinghe, Joseph A. Caruso*

Department of Chemistry, University of Cincinnati, Cincinnati, OH 45221, USA

ARTICLE INFO

Article history:
Received 21 October 2010
Received in revised form 7 December 2010
Accepted 8 December 2010
Available online 16 December 2010

Keywords:
Selenoproteome
Isotope pattern
SEC
ICPMS
MALDI-TOF MS
capLC–ESI-MS
Deamidation

ABSTRACT

Owing to recent studies which showed that Se supplementation in the diet can reduce the risk of several forms of cancer, efforts have been directed to identify anticarcinogenic activity in Se compounds from natural sources with the objective of using those as food supplements. As a part of a continuing research project directed at identifying Se-containing species in food stuffs, the distribution of selenium-containing proteins in water-soluble protein fraction of seeds of *Bertholletia excelsa* H.B.K., which typically have a high Se content, was studied under non-denaturing conditions and the effectiveness of protein fractional precipitation by ammonium sulfate was investigated with the objectives to preconcentrate the protein(s) of interest and to reduce the matrix complexity. By SEC–ICPMS, Se-containing proteins (selenoproteome with some additional proteins) were found demonstrating the usefulness of ICPMS as an "on-line assay" in sub-proteomics investigations. Fractions with Se-containing proteins, collected from SEC, were tryptically digested and tryptic digests were analyzed by MALDI-TOF-reflectron MS and capLC–ESI-QTOF-MS. Isotope patterns which are different from the typical isotope patterns of peptides containing C, H, O, N and S were used to identify selenium-containing peptides in the digests. Observed isotope patterns were slightly different from the predicted isotope patterns and partial deamidation of selenium-containing peptides is suggested to explain the modified isotope pattern. To our knowledge, this is the first report of the effect of partial deamidation of selenium-containing peptides on the observed selenium isotope pattern, although the evidence is indirect. According to the SEC–ICPMS study, Se-containing proteins in water-soluble protein fraction of *Bertholletia excelsa* H.B.K. seeds can be divided into two main sub-groups-high molecular weight Se-containing proteins and low molecular weight Se-containing proteins. LC–ESI-MS and MS/MS analysis of tryptic digest of low-molecular weight Se-containing proteins identified 2S albumins, which are rich in methionines; therefore, with a high probability of non-specific selenomethionine incorporation—as the proteins present in that fraction.

1. Introduction

Se, discovered in 1817, was considered a toxic element until it was reported in 1957 to be an essential trace element for mammals [1,2]. Now it is well established that Se is beneficial at lower concentrations, but toxic at higher concentrations, and the range from deficiency, essentiality, and toxicity is narrow. Regular consumption of food, which has more than 1 mg kg^{-1} Se, results in toxicity while food with less than 0.1 mg kg^{-1} Se results in deficiency [3]. Even though the effects of Se were described in terms of the total elemental concentration in earlier studies, now it has been shown that the chemical form of Se, as well as the dose, determine its biological activities as an essential element or a toxicant. In environmental and in biological systems, Se has been detected as inorganic species as well as in organoselenium forms, which range from small molecules, such as methyl selenol (MeSeH) to very complex selenoproteins. Most of these are analogous to their sulfur counterparts. In selenoproteins, Se is incorporated in the amino acid chain in the form of selenocysteine (now recognized as the 21st amino acid) or as selenomethionine.

In 1973, Se was identified as a stoichiometric, covalently bound component of glutathione peroxidase (GPx) [4,5]. Se is present in this enzyme as a selenocysteine residue integrated to the amino acid chain. In the majority of mammalian selenoproteins, selenium occurs in the form of selenocysteine.

Studies carried out in 1969 generated interest in the research on Se for cancer chemoprevention and in 1977 Se was reported to be a potential human cancer protective agent [6,7]. As far as

Abbreviations: SEC, size exclusion chromatography; ICPMS, inductively-coupled plasma mass spectrometry; MALDI-TOF-reflectron MS, matrix-assisted laser desorption ionization time-of-flight reflectron mass spectrometry; CapLC–ESI-MS, capillary liquid chromatography electrospray ionization mass spectrometry.

* Corresponding author. Tel.: +1 513 556 5858; fax: +1 513 556 0142.
E-mail address: joseph.caruso@uc.edu (J.A. Caruso).

1387-3806/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2010.12.005

human nutrition is concerned, the most recognized role of Se is its ability to serve as an antioxidant and possible cancer-preventive agent [8,9]. Owing to the high prevalence of this disease, searching for naturally occurring agents, which may serve as cancer chemopreventives, has been an important objective of various research efforts. Since recent studies have shown that Se supplementation in the diet can reduce the risk of several forms of cancer [8], efforts have been directed to identify anticarcinogenic activity in Se compounds from natural sources, with the objective of using those as food supplements. The Se content of food is highly dependent on the amount of Se in the soil that varies, not only from country to country, but also from region to region within a country. Moreover, it is also dependent on the ability of plants to take up and accumulate various Se forms. When considering selenoproteins, Se content is dependent on the developmental stage of the plant as well, since different proteins are expressed in different levels at different times. The majority of biological and biomedical effects appear to be mediated by proteins that contain selenium. The amount of selenium in proteins from plants, yeast and yogurt has been shown to be dependent on the extent of nonspecific incorporation of selenomethionine and selenocysteine [10]. It is known that selenomethionine can be nonspecifically inserted into proteins by the usual methionine incorporation process [11,12]. This type of substitution is very common in biological systems, and in humans and other animals, selenomethionine from ingested foods can be incorporated randomly into cellular proteins. It has been shown in studies with bacteria, that several proteins, especially rich in methionine, were radio-labeled because of nonspecific ^{75}Se selenomethionine substitution throughout the polypeptide chain, even when a relatively low level of labeled substrate was provided [13].

The proteins which contain selenium are classified into three groups: (A) the proteins which contain selenium in the form of genetically encoded selenocysteine. These are named selenoproteins. (B) In a second category, selenium is incorporated nonspecifically into proteins in place of methionine by replacing the sulfur and these are termed selenium-containing proteins; therefore, these are mainly found in methionine-rich proteins. (C) The third group comprises the specific proteins in which selenium is only attached to the molecule. There are few examples reported so far for this category [14–16]. Any peptide fragment containing selenocysteine and/or selenomethionine are termed as Se-containing peptides.

According to previous research, Brazil nuts, the seeds of *Bertholletia excelsa* H.B.K., which is a large tropical tree of the *Lecythidacea* family, widespread in Latin American high and middle basins, are rich in lipids (65–70%) and contain 15–17% protein by fresh weight and around 50% protein in its defatted flour [17]. As early as 1892, it was discovered that Brazil nuts have unusually high content of the sulfur-containing amino acids, cysteine and methionine, 8.3% by weight [18,19]. The protein content of the Brazil nut has been fractionated into three size classes of proteins, namely the 11S, 7S and 2S proteins [20]. While the level of methionine in all fractions is higher than that of many other plant seed storage proteins [21], the 2S fraction, albumin, is exceptionally rich in sulfur containing amino acids, about 30% methionine and cysteine, and comprises about 30% of the total protein fraction [22,23]. It is expected that there will be similar chemical or biochemical behaviors for S and Se because of their relationship in the periodic table.

It has been indicated that soil in some parts of Brazil is rich in Se and previous studies have shown that Brazil nuts from trees growing in those areas are rich in Se [24]. Previous research carried out by our group has shown that Se in Brazil nuts is bound to proteins and according to these results, Se bound to protein exists in two different forms, namely firmly bound Se and weakly bound Se. The term weakly bound has been ascribed to selenodisulfides (RS-Se-SR′) or methylselenylsulfides (RS-SeCH_3) [25]. In an attempt to characterize the Se-containing species in a Proteinase-K digest of Brazil nut proteins, a tentative structure was suggested but unconfirmed, based on electrospray-MS data [26]. Another study has shown that Se in the protein fraction of Brazil nut is covalently linked to the two amino acids, selenomethionine and selenocysteine [27].

Several studies have been carried out in the last few years to characterize Se-containing peptides by experimental procedures used in proteomics research, namely single or multidimensional chromatography and mass spectrometric techniques, matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) and electrospray ionization mass spectrometry (ESI-MS). In these studies, Se-containing peptides were sequenced by high performance liquid chromatography ESI-MS/MS (LC–ESI-MS/MS). In LC–ESI-MS/MS experiment, ions with intensities which are below the pre-determined intensity value are not subjected to collision-induced dissociation. In such a case, m/z value of the ion of interest must be selected by the operator. When compared with the intensities of other peptides, the intensity of Se-containing peptides are very low (except in Se-enriched samples) and as a result those are not selected for collision-induced dissociation. Therefore, it is necessary to determine the m/z value of Se-containing peptide in order to input that value in setting up the ESI-MS/MS experiment. As Se has a unique isotope distribution, Se-containing peptides can be identified by MALDI spectrum through the characteristic isotope pattern (though modified to some extent) and m/z value of Se-containing peptides can be obtained. In some of the studies carried out to characterize Se-containing peptides, purified selenoproteins have been used [28–30] and in all other studies, a peptide mixture containing selenopeptides has been examined [31–40]. For studies in which a peptide mixture was analyzed, the sample investigated was selenium-enriched yeast (*Saccharomyces cereviseae*) except in a very recent study in which Se-containing peptides from Brazil nuts were investigated [41]. In that study, tryptic digests of the selenium-containing proteins were fractionated by size-exclusion chromatography, and then concentrated fractions were analyzed by reversed phase nanoLC–ICPMS (RP-nanoLC–ICPMS) and RP-nanoLC-nanospray-quadrupole time of flight (QTOF) mass spectrometry under identical elution conditions. Selection of Se-containing peptides for an ESI-MS/MS study has been carried out by examining the LC–ESI spectra for characteristic Se-isotope patterns at elution times when a Se signal is observed in nano-RP-LC–ICPMS. Their justification for this approach is that it is necessary because of the low-level of Se-containing peptides in Brazil nuts [41]. In studies in which purified selenoproteins were used (as in Ref. [28]), conventional or standard proteomics technologies have been used and expected results have been obtained-that is, sequencing Se-containing peptides by LC–ESI-MS/MS methods. Also, in such cases, necessities such as sample enrichment and minimizing or removing signal suppression do not arise. In studies in which a peptide mixture was analyzed, the sample investigated was selenium-enriched yeast (*S. cereviseae*) (except in one study, Ref. [41]). Therefore, in those studies too, expected results have been obtained by following standard proteomics methodologies as Se-containing proteins are expressed in considerable quantities in Se-enriched yeast. That is, such a sample is also not an analytical challenge. Even though Brazil nut is a natural sample, it has a higher content of Se (when compared with the content of Se in normal plants or, in general, other living specimens) owing to the elemental composition of the soil in which Brazil nut plants grow, but less than Se-enriched yeast. Therefore, standard or conventional proteomics methodologies may not generate polypeptides mixtures (by enzymatic digestion) with a considerable amount of Se-containing peptides (when compared with the quantity of other peptides) and free of signal suppression.

In this present study, distribution of Se-containing proteins in the water-soluble protein fraction of Brazil nuts under non-denaturing conditions was studied and the effectiveness of fractional precipitation of proteins by ammonium sulfate was investigated with the objectives to preconcentrate the protein(s) of interest and reduce the matrix complexity. The main objective of this research is to discover the selenium-containing peptides in tryptic digests (with no or minimal separation) of selenium-containing proteins by ICPMS, MALDI-MS-directed LC–ESI-MS/MS methods and, thereby, to evaluate the applicability of this method—ICPMS-MALDI-directed LC–ESI-MS/MS—i.e., without using nano-based methods and custom-made interfaces, in the investigations of Se-containing peptides in non-enriched biological samples.

In order to investigate feasibility of using the proposed experimental procedure with the facilities available to us, bovine glutathione peroxidase was subjected to tryptic digestion and analyzed by MALDI-MS and LC–ESI-MS/MS.

2. Experimental

2.1. Materials and reagents

Tris base, dithioerythritol (DTT), iodoacetamide (IAA), phenylmethylsulfonyl fluoride (PMSF), polyvinylpyrrolidone (PVP) were obtained from Sigma Aldrich (St. Louis, MO, USA). Trypsin was obtained from Promeaga (Madison, WI, USA). Water was doubly deionized (18.2 MΩ cm), prepared by passing water through a Nanopure treatment system (Barnstead, Boston, MA, USA). All reagents used were analytical grades and used without further purification. Centrifuge filters for ultrafiltration with a cut-off of 5 kDa were purchased from Millipore (Bedford, MA, USA).

2.2. Instrumentation

Chromatographic separations were performed using an Agilent 1100 liquid chromatographic system (Agilent Technologies, Santa Clara, CA, USA) equipped with an LC binary pump, an auto sampler, vacuum degasser, a thermostatic column compartment and a diode array detector. The chromatographic columns used were a Superdex^{TM} 200 (10 mm × 300 mm; 13 μm particle size), Superdex^{TM} peptide (10 mm × 300 mm; 13 μm particle size) and HiTrap^{TM} Desalting (16 mm × 25 mm; G-25 Superfine) columns (Amersham Pharmacia Biotech AB, Uppsala, Sweden) for protein and peptide SEC and, desalting, respectively. SEC column mass ranges were calibrated using protein and peptide standards according to Amersham Pharmacia Biotech instructions.

An Agilent 7500ce ICPMS (Agilent Technologies, Santa Clara, CA, USA) with a micromist nebulizer and a Peltier cooled spray chamber (2 °C) and a shield torch system was used for selenium specific detection. The outlet of the UV detector was connected online to the liquid sample inlet of the ICPMS nebulizer using a 300 mm long (i.d. 0.25 mm) polyetheretherketone (PEEK) tubing.

MALDI-MS experiments were done on a Bruker Daltonics Reflex IV MALDI mass spectrometer (Billerica, MA, USA), equipped with a nitrogen laser. The instrument was calibrated externally using two standards [ACTH fragment (18–39) m/z 2465.1 $(M+H)^+$and bradykinin fragment (1–5) m/z 573.3 $(M+H)^+$] (resolution: 1000; mass accuracy: within ±0.5 Da for peptides). The data were acquired as positive ions in the reflectron mode.

LC–ESI-MS experiments were carried out using a Waters Cap-LC system coupled to Q-Tof-2 mass spectrometer (Micromass, Manchester, UK) (resolution: 4000; mass accuracy: within 10–20 ppm).

2.3. Sample preparation

Seeds of *Bertholletia excelsa* H.B.K. were purchased from "Nuts" www.nuts.com. Nuts with shells were dipped in liquid nitrogen for 15 min and after that, shells were removed and the seeds were ground. The ground material was defatted with *n*-hexane by stirring at room temperature (12–14 h; two defating cycles per sample). The defatted powder was used for the protein extraction. Four grams of defatted Brazil nut powder were extracted in 100 ml of 30 mM Tris–HCl pH 7.5 containing 100 mg of polyvinylpyrrolidone (PVP), 60 mg of ethylene diaminetetracetic acid (EDTA), 16 mg of dithiothreitol (DTT) and 18 mg of phenylmethylsulfonyl fluoride (PMSF) for 12 h under nitrogen at room temperature. The extract was centrifuged at 5000 rpm for 30 min at 25 °C (Sorvill SS-34 rotor) [2987 × g]. The supernatant was separated and made up to 100 ml with 30 mM Tris–HCl pH 7.5. The supernatant was used for ammonium sulfate precipitation.

Traditionally, ammonium sulfate concentration is expressed as % of saturation. The amount of ammonium sulfate necessary to reach a certain % is calculated by the formula

$$g = \frac{533\,(S_2 - S_1)}{100 - 0.3\,S_2}$$

g is the amount of ammonium sulfate (in grams) to be added to 1 l of the solution at 20 °C to increase the ammonium sulfate concentration from S_1% saturation to S_2% saturation [42].

2.3.1. Bulk precipitation of proteins by ammonium sulfate

5 ml of extract was cooled in an ice-water bath and the amount of ammonium sulfate necessary to reach 90% of saturation was added while stirring. After completing the addition, the solution was stirred for another 30 min in the ice-water bath. After that, the solution was centrifuged at 5000 rpm for 30 min at 5 °C (Sorvill SS-34 rotor) [2987 × g]. The supernatant was separated from the precipitate and the precipitate was dissolved in 2 ml of 30 mM Tris–HCl (pH 7.5).

2.3.2. Fractional precipitation of proteins by ammonium sulfate

5 ml of extract was cooled in an ice-water bath and the concentration of ammonium sulfate was increased from 0% to 90% in 10% increments while stirring. After completing the addition of ammonium sulfate necessary to reach certain saturation, the solution was stirred for another 30 min in the ice-water bath. If cloudiness or formation of a precipitate was not observed, the ammonium sulfate concentration was increased in 10% saturation steps. If formation of precipitate was observed, then the solution was centrifuged at 5000 rpm for 30 min at 5 °C (Sorvill SS-34 rotor) [2987 × g]. The supernatant was separated from the precipitate and the precipitate was dissolved in 1 ml of 30 mM Tris–HCl (pH 7.5). The amount of ammonium sulfate necessary to reach the next 10% increment was calculated, added to the supernatant and the same procedure was followed.

2.3.3. Separation of Se-containing protein fraction by ammonium sulfate precipitation and size exclusion chromatography

50 ml of extract was cooled in an ice-water bath and the concentration of ammonium sulfate was increased from 0% to 50% while stirring. The precipitate of 50% of saturation was discarded and the ammonium sulfate concentration of the supernatant of 50% precipitation was increased to 80% of saturation of ammonium sulfate. The precipitate of this step was separated after centrifugation and dissolved in 5 ml of Tris–HCl buffer pH 7.5 immediately after centrifugation. 3 ml of that solution was subjected to desalting using a HiTrap^{TM} desalting column and the fraction containing selenium was concentrated by ultrafiltration (Spin column [Millipore, Bedford, MA, USA] 5 kDa cut-off)—final volume ∼1 ml. 500 μl of the

concentrated fraction was fractionated by SEC [Column: Superdex 200, eluent: 30 mM Tris–HCl pH 7.5, flow rate: 0.7 ml/min: sample volume: 100 μl]. Fractions eluted from 15 min to 18 min [fraction A], from 21 min to 23.5 min [fraction B], from 23.5 min to 25.5 min [fraction C] and from 25.5 min to 27.5 min [fraction D] were collected and concentrated by ultrafiltration to final volume ~250 μl.

2.3.4. Separation of tryptic digests of fraction D by size exclusion chromatography

100 μl of the tryptic digest was fractionated by SEC [Column: Superdex™ peptide, eluents: 30 mM Tris–HCl pH 7.5 and 30 mM NH_4Ac pH 6.8, flow rate: 0.6 ml/min: sample volume: 100 μl]. Fractions which eluted from 18 min to 21 min [fraction E], from 23 min to 26 min [fraction F], from 26 min to 28 min [fraction G], from 28 min to 30 min [fraction H] and from 31 min to 33 min [fraction I] when eluent was 30 mM NH_4Ac, were collected and concentrated by vacuum centrifugation at 37 °C.

2.4. In-solution tryptic digestion

80 μl of 8 M urea, 0.4 M ammonium bicarbonate (pH 7.5–8.8) solution was added to 60 μl of concentrated protein fraction A and followed by the addition of 20 μl of 45 mM DTT solution. The reaction mixture was incubated at 50 °C for 15 min and then cooled to room temperature. Then, 20 μl of 100 mM iodoacetamide (IAA) was added to the reaction mixture and left it in dark at room temperature for 15 min. After that, 140 μl of water was added to dilute the digestion buffer to 2 M urea and 0.1 M ammonium bicarbonate. Then, 20 μl of trypsin (0.1 μg/μl) was added to the reaction mixture and incubated at 37 °C overnight [43].

2.5. MALDI analysis

10 μl of the tryptic digest was desalted by C-18 Zip-Tip and mixed with α-cyano-4-hydroxycinnamic acid (α-CHCA) matrix (in 50% acetonitrile [ACN]/0.1% trifluroacetic acid [TFA]) in a ratio of 1:1 and analyzed by MALDI-MS.

2.6. LC–ESI-MS analysis

10 μl of the tryptic digest (acidified with 1% formic acid [FA]) was loaded on a 10 cm × 75 μm C-18 capillary column (made in-house) using a Waters Cap-LC system. The peptides were separated and eluted from the column using a Cap-LC gradient of solvent A, which was 95% water, 5% acetonitrile (ACN), 0.1% formic acid (FA) and solvent B, which was 95% ACN, 5% water and 0.1% FA. A linear gradient was run starting at 15% solvent B, 85% solvent A to 80% solvent B, 20% solvent A in 70 min. The flow rate was set to 7 μl per minute, but split 50:50 before the column. The Cap-LC system was directly connected to the Waters Q-Tof-2 mass spectrometer, in which the eluted peptides were directly detected (LC–MS), fragmented (LC–MS/MS) and data collected accordingly. The collected data were processed using Mass Lynx 4.0 software.

3. Results and discussion

3.1. Importance of elemental and molecular mass spectrometries in selenoproteome research

By appropriate selection of mobile phases and other parameters, ICPMS in normal configuration can be used for the detection of Se-containing proteins in size-exclusion chromatography and ion-exchange chromatography. Wide dynamic range, high sensitivity, low detection limits and simultaneous determination of multiple elements (even different isotopes of the same elements) are some of the advantages which ICPMS possesses over traditional bio-assay based detection methods, at least for the screening stage. This use of ICPMS in conjunction with molecular MS methods as discussed below encompasses what is more newly defined as a metallomics approach to further unraveling the metalloproteome, and in this case, the selenoproteome.

With the advent of soft ionization techniques, MALDI and ESI mass spectrometries have been indispensable tools in proteomics research. Since pure proteins or simple protein mixtures were not used in this study, protein identification is via peptide sequencing by MS/MS. The amino acid sequence of a peptide is more specific than its mass for protein identification by database searching. When compared with MALDI-MS, ESI-MS has a special advantage in analyzing complex mixtures, since until recently only it could be readily coupled with chromatographic methods, thereby reducing sample complexity at the time of ionization and minimizing signal suppression. In routine ESI-MS/MS experiments, the ions for which intensity exceeds a pre-determined value are subjected to collision-induced dissociation. But, if the intensity of the ion under study is below that level, then the ion of interest should be selected by the operator. Frequently, this is the case in analyzing low-abundance peptides in LC–ESI-MS/MS studies. In such ESI-MS/MS experiments, the ion subjected to collision-induced dissociation is selected on the basis of its m/z value and for that, the m/z value of the ion must be known.

Selenium has a unique isotope distribution of ^{74}Se (0.89%), ^{76}Se (9.3%), ^{77}Se (7.63%), ^{78}Se (23.77%), ^{80}Se (49.61%) and ^{82}Se (8.73%) and this provides a "finger print" of the presence of Se in a peptide. The importance of this characteristic pattern is that using it, ions containing Se can be identified. In a typical LC–ESI-MS experiment, a large number of mass spectra are generated and therefore, examining all spectra for finding the mass spectrum with the Se isotopic pattern is not a practical option. Moreover, in ESI, multiply charged ions are produced and this will also contribute to the complexity of the observed mass spectrum, since signal overlap is possible. Therefore, in order to identify the peptide signal(s) with the characteristic Se-isotope pattern and, thereby, to obtain the m/z value of the Se-containing peptides, MALDI-MS experiments are employed. This is the basis of the MALDI-MS-directed LC–ESI-MS/MS analysis for selenopeptide mapping. Even though the characteristic isotope pattern for a single Se atom is observed for small compounds containing Se, this pattern is modified in large molecules such as peptides. This is because of the contributions of the isotope pattern of other atoms such as ^{13}C (especially as they become more numerous), the number of selenium atoms and in some cases, the post translational modifications (S.B. Jayasinghe and J.A. Caruso Manuscript in preparation). Therefore, in selecting Se-containing peptides, appropriate attention should be given, since selecting the selenium-containing peptide is the key to the success of the experiment and moreover, it reveals important information about the proteins under study.

3.2. Importance of fractionating the crude protein mixture

An important aspect in proteomics studies is the dynamic range of the analysis. For example, a human cell type may express up to 20,000 proteins at any given time with a dynamic range of up to five orders of magnitude. But, no existing analytical methodologies are capable of resolving and detecting such protein mixtures with such a wide dynamic range. When selenium-containing proteins are involved, this situation is more complicated, since selenomethionine is randomly incorporated, thereby making such proteins a subset of their sulfur analogs. Therefore, it is essential to pre-fractionate the crude protein mixture and enrich the protein(s) of interest. In general, chromatographic methods are used to fraction-

ate mixtures in separation science. But, in protein separation, crude protein extract cannot be directly applied to chromatographic systems owing to reasons such as low protein concentration and contaminants. So first, proteins are concentrated into a small volume. If the fractionation also can be achieved while concentrating, such an experimental procedure will minimize the number of steps in a purification procedure, and thereby, minimize the sample loss. There are several reagents for the precipitation of proteins such as ammonium sulfate, trichloroacetic acid (TCA), acetone and chloroform/methanol. But reagents such as TCA, acetone and chloroform/methanol cause protein denaturation. Therefore, ammonium sulfate was selected as the precipitating agent in order to concentrate the protein fraction, since one of the objectives of the present study is to investigate the distribution of Se-containing proteins in the non-denatured state, in the water-soluble protein fraction.

3.3. MALDI-TOF reflectron and cap-LC–ESI-QTOF-MS and MS/MS studies of glutathione peroxidase

In glutathione peroxidase, Se is present in the form of selenocysteine at the 45 position of the amino acid chain. The reduced and carboxymethylated protein was subjected to in-solution tryptic digestion and the tryptic digest was analyzed by MALDI-MS. As expected, there was a signal at m/z 1793 with the selenium isotope pattern (Supplementary information-Fig. 1) which is identical with the theoretically predicted Se isotope pattern (Supplementary information-Fig. 2). In LC–ESI-MS, a doubly charged ion at m/z 897 was observed with the characteristic isotope pattern, since in ESI-MS, the prominent ions are doubly and multiply charged ions (Supplementary information-Fig. 3).

Peptide containing selenocysteine—VLLIENVASL**C**GTTVR

C = selenocysteine [reduced and alkylated with iodoacetamide]

Molecular formula of the peptide (reduced and alkylated): $C_{71}H_{133}N_{21}O_{24}Se$

3.4. Distribution of Se-containing proteins in water-soluble protein extract

By size-exclusion chromatography (SEC) on-line coupled with a UV detector and ICPMS detection, molecular weight distributions of Se-containing proteins, which are soluble in 30 mM Tris–HCl (pH 7.5), were studied. Even though the SEC is theoretically based on the hydrodynamic volume of analytes, most column materials used for SEC are slightly anionic and hydrophobic by nature. In size-exclusion chromatography, a salt, in most cases 100 mM NaCl, is used to prevent these types of interactions. But, if the chromatography is coupled to ICPMS, NaCl cannot be used since the nebulizer and cones clog because NaCl is not volatile at sample introduction temperatures. As a result, the estimated molecular weight values may be slightly different from those obtained from ideal size-exclusion chromatographic separation conditions.

In this study, water soluble proteins of Brazil nuts were extracted from defatted Brazil nut powder as mentioned in the methodology section and attention was directed only to the qualitative aspects. The approach was, first to develop a method and then at a later stage, to carry out a quantitative study while fine-tuning the method based on the observation at this stage. Anyway, there is a possibility of some loss of proteins and deamidation (to be explained in detail later) owing to carrying out defating and extraction at room temperature.

The extract was centrifuged and the supernatant was used for ammonium sulfate precipitation (90% of ammonium sulfate saturation) in order to concentrate proteins and remove interfering compounds. The concentrated protein fraction which was in high

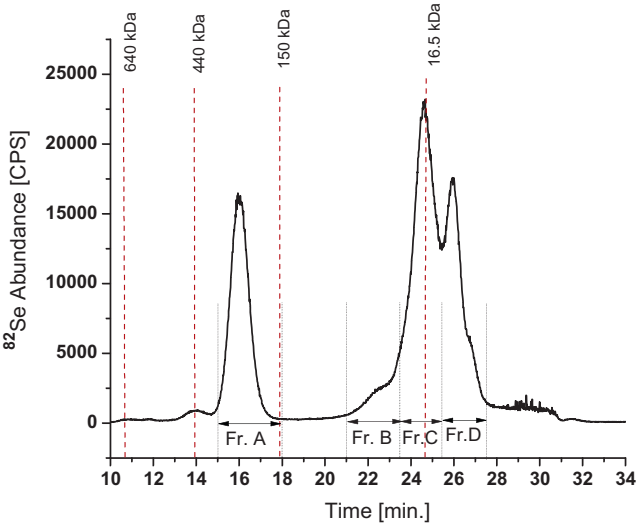

Fig. 1. Distribution of Se-containing proteins in water-soluble protein extract [SEC–ICPMS-^{82}Se].

salt concentration was desalted and desalting was monitored by UV detection at 280 nm. Two broad peaks (bands) were observed. Both peaks were collected and analyzed by SEC–ICPMS. The first band contained all Se peaks. Therefore, there is no loss of Se-containing proteins qualitatively as a result of desalting. As the eluting solvent, 30 mM Tris–HCl at pH 7.5 was selected, since at that pH most known Brazil nut proteins are soluble. According to the SEC–ICPMS profile, selenium-containing proteins in water-soluble extract are present in the high molecular weight region (between 440 kDa and 150 kDa) as well as in the low molecular weight region (~16.5 kDa) (Fig. 1).

Fractional precipitation results (Fig. 2) show that no Se-containing proteins are precipitated up to 60% of saturation ammonium sulfate and high-molecular weight Se-containing proteins can be precipitated at 60% saturation of ammonium sulfate with a small amount of low molecular weight Se-containing proteins co-precipitating, while most of low-molecular weight Se-containing proteins precipitate in 60–80% saturation range. According to these results, it is evident that a certain amount of proteins, which do not contain selenium, can be removed by increasing the concentration of ammonium sulfate in the crude extract up to

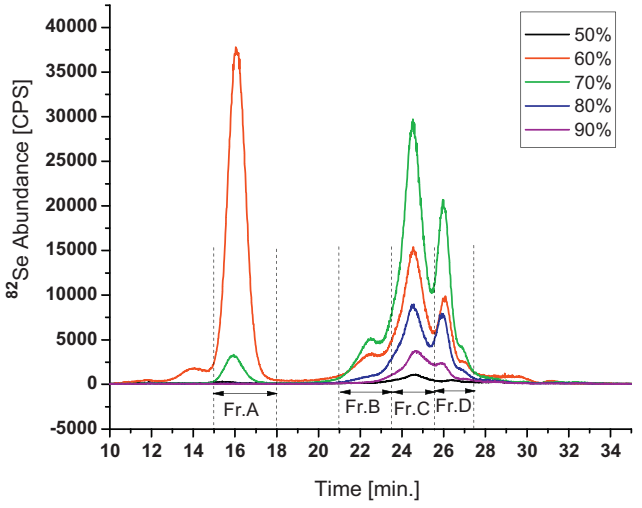

Fig. 2. Fractional precipitation of selenium-containing proteins by ammonium sulfate [SEC–ICPMS-Profiles].

50% of saturation and removing the precipitate after centrifugation. After that, by increasing the concentration of ammonium sulfate in the supernatant of 50% precipitation, up to 80% of saturation, most of the Se-containing proteins can be precipitated. Moreover, from these results, it can be concluded that selenium-containing proteins in the high molecular weight region are more hydrophobic when compared with selenium-containing proteins in the low molecular weight region (Fig. 2). Based on the SEC–ICPMS profile of crude protein mixture, which was obtained by 50–80% precipitation with ammonium sulfate, four different selenium-containing protein fractions were identified—fraction A [from 15 min to 18 min], fraction B [from 21 min to 23.5 min], fraction C [from 23.5 min to 25.5 min] and fraction D [from 25.5 min to 27.5 min]. Ultrafiltration was selected to remove excess buffer solutions from the collected fractions from SEC and desalting columns and thereby concentrate the collected fractions as removal of excess buffer and concentration can be achieved in 45–60 min by that method. When these fractions were collected and re-chromatographed on the same column under identical conditions, all fractions except fraction D showed elution times identical to their corresponding values in the chromatogram of the raw extract, indicating stability of the selenocompounds with respect to the size-exclusion chromatographic resolution. Fraction D showed an additional peak (shown by an arrow in Supplementary information-Fig. 4) when re-chromatographed, which may be due to either the decomposition of fraction D or the presence of selenoproteins, which were not shown in the chromatogram of the raw extract owing to sample overload.

According to the size-exclusion chromatographic calibration profile, fraction A may be composed of 11S globulin proteins, which are proteins composed of six identical monomers with a molecular weight of ~52 kDa. Fractions B, C and D may be composed of 2S proteins with a molecular weight of ~16.5 kDa. These results differ from the observed selenoprotein profiles in previous studies reported. In these earlier studies [25,41], denaturing conditions were used in the sample preparation and this may have dissociated the hexamers of 11S globulin proteins. According to Dernovics et al. [41], the protein precipitate from acetone precipitation of crude extract is not completely soluble upon re-solubilization, and 7S vicilins and 11S legumins have been suggested as the insoluble residue. In our study, the protein precipitate obtained by ammonium sulfate precipitation was dissolved in a minimum volume of 30 mM Tris–HCl (pH 7.5) buffer immediately after the precipitation and complete dissolution was observed.

In the determination of Se by quadrupole ICPMS without collision/reaction cell, formation of polyatomic interference ions in argon plasmas is one of the main difficulties. The formation of the dimers $^{38}Ar^{40}Ar$ and $^{40}Ar_2$ are the main interference ions and overlap with the two most abundant Se isotopes ^{78}Se (23.8% abundance) and ^{80}Se (49.6% abundance), respectively. Using ^{76}Se (9.4% abundance) for the determination of Se is not possible owing to the interference from $^{38}Ar_2$. In addition to the above-mentioned interferences caused by the carrier gas argon, the buffer used may contribute to interferences in chromatography–ICPMS coupling experiments. For example, chloride ion from Tris–HCl may generate some polyatomic interferences such as $^{40}Ar^{37}Cl$ hindering the detection of ^{77}Se. Therefore, it was decided to use ^{82}Se and ^{78}Se isotopes for the detection of Se and $^{78}Se/^{82}Se$ isotope ratio was determined and compared with the theoretical values in order to determine whether interference-free detection was achieved. These determinations showed that $^{78}Se/^{82}Se$ isotope ratio for both Se peaks (from 15 to 17.5 min and 21.5 to 27.5 min) are 2.54 and 2.56, respectively, and match well to the theoretical value of 2.40 considering the quality of isotope ratios available on this instrument (Supplementary information-Fig. 5).

3.5. MALDI-MS and LC–ESI-MS/MS analysis of tryptic digest (fraction A)

The high molecular-weight selenium-containing protein fraction was collected and concentrated by ultra filtration. A part of the concentrated fraction (60 μl) was subjected to in-solution tryptic digestion following the standard protocol used in proteomic research. A part of the tryptic digest was purified by C-18 Zip-Tip and analyzed by MALDI-TOF-MS in the reflectron mode. Several signals with isotopic patterns, which are different from normal ^{13}C isotope patterns, were observed at m/z values 1776.2, 1219.8, 2113.3 and 2130.5 (Fig. 3a–c). All these signals were from singly charged ions as revealed by space between adjacent lines. From these signals, the signal at *m/z* 1776.2 is the most intense one and the signal at *m/z* 1219.8 is moderate in intensity, while signals at *m/z* 2113.3 and 2130.5 are not well-resolved from base line. The isotope pattern of the signal at *m/z* 1776.2 shows that it may be due to a peptide containing one or two selenium atoms.

In order to examine the effect of number of selenium atoms and the contribution of other atoms such as ^{13}C to the isotope pattern, theoretical isotope patterns for two hypothetical peptides with comparable molecular weight were generated using MassLynx (v.4.0) software. $C_{80}H_{135}N_{19}O_{20}Se$ [VNVPILTFLQLAA(SeMet)KG] and $C_{77}H_{130}N_{18}O_{19}Se_2$ [VNVPILTFLQ(SeMet)AA(SeMet)K] are the empirical formulas of peptides used to generate isotope patterns shown in Figs. 4.1 a and b, and 4.2 a and b. When theoretical isotope patterns were generated, they (Figs. 4.1 a and b, and 4.2 a and b) were slightly different from the observed patterns (Fig. 3). One possible explanation for this can be presented by assuming that the selenium-containing peptide contains asparagines (Asn) and/or glutamine (Gln) residues and undergoes partial deamidation. If deamidation did not take place or complete deamidation did take place, isotope pattern in ~1500 Da mass range should be similar to that of Fig. 4.1 a or b, respectively [OR Fig. 4.2 a or b, respectively]. But, the observed pattern is similar to that of Fig. 4.1 c [OR Fig. 4.2 c].

Non-enzymatic deamidation of asparagines (Asn) and glutamine (Gln) in proteins and peptides may occur under physiological conditions. The rate of non-enzymatic deamidation is dependent on conditions such as pH, temperature, and the amino acid sequence and higher order structures of the protein or peptide [44–46]. The difference between the normal form and deamidated form is 1 Da.

$$\text{Normal form} \xrightarrow[+OH]{-NH_2} \text{Deamidated form}$$

In such a situation, the observed isotope pattern in the MALDI-TOF-MS reflectron spectrum is the combined effect of the two species [normal form and deamidated form]. The exact pattern is dependent on the ratio of the two species (Figs. 4.1 c and 4.2 c). Since Brazil nut has been reported to be rich in these amino acids (Asn and Gln), this suggests a possible explanation for the observed isotope pattern. When the tryptic digest was analyzed by LC–ESI-MS, a doubly charged ion at *m/z* 887 [Supplementary information-Fig. 6] with the same isotopic pattern was observed, but the intensity of that signal was very weak, which may be indirect evidence for the poor ionization of that peptide in the ESI process. If those two species, normal form and deamidated form, co-elute, then again, what is observed in the ESI-MS spectrum as well, is the combined isotope effect of those two species. In the LC–ESI-MS experiment, the elution conditions were ACN-H_2O with formic acid as the ion pairing agent; hence acidic conditions. It has been reported that acidic mobile phase conditions employed in reversed-phase chromatography cannot separate the deamidated form from the parent molecule (non-deamidated form) and neutral eluent systems such as 30 mM sodium phosphate at pH 6.5 or 20 mM ammonium acetate at pH 6.8, are necessary [47,48]. The intensity of the ion at *m/z* 887

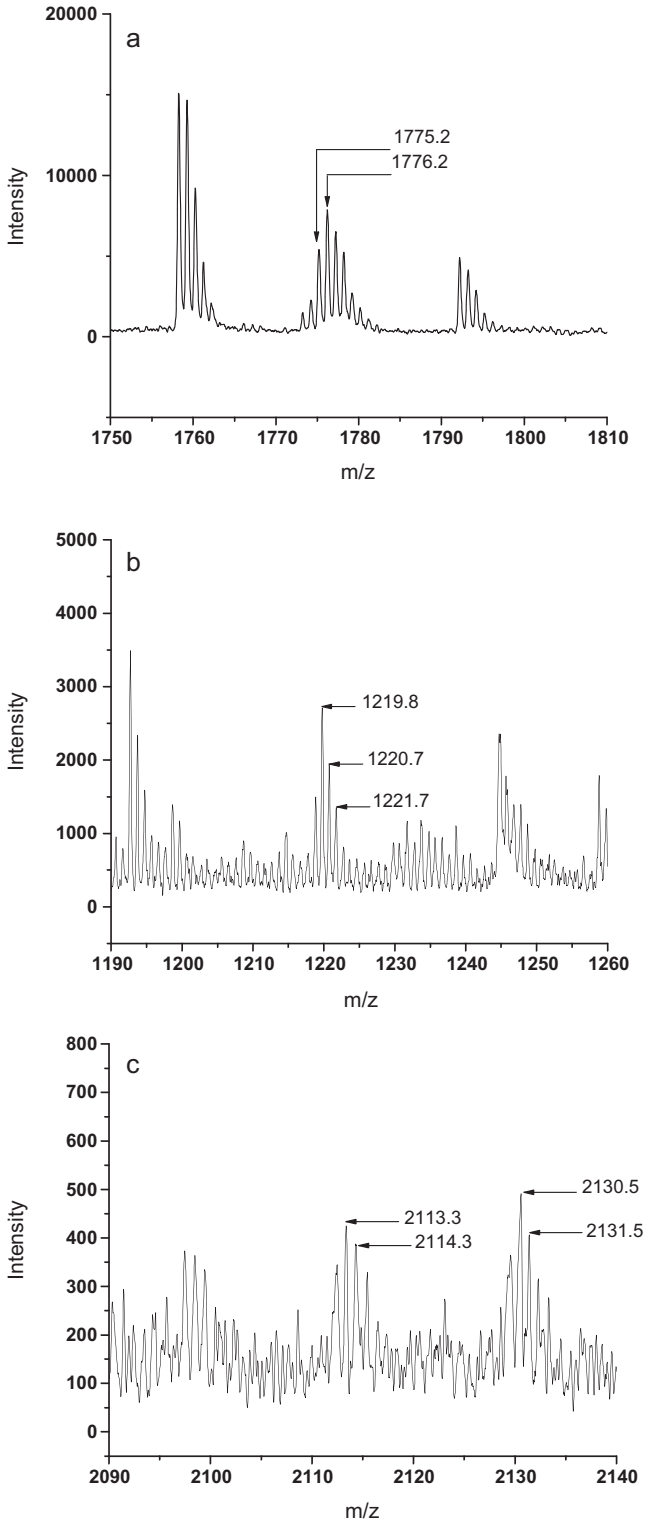

Fig. 3. (a) MALDI spectrum between *m*/*z* 1750 and *m*/*z* 1810. (b) MALDI spectrum between *m*/*z* 1190 and *m*/*z* 1260. (c) MALDI spectrum between *m*/*z* 2090 and *m*/*z* 2140 [Se-containing peptides from fraction A].

was below the pre-defined value for the ion selection with collision-induced dissociation. Therefore, there are no MS/MS data for this ion.

When the tryptic digest of glutathione peroxidase is of concern, the peptide fragment containing selenocysteine has an asparagine (N) unit. But, deamidation was not observed in tryptic digest of glutathione peroxidase when analyzed by MALDI-MS and LC–ESI-MS and the "classical" isotope pattern was observed. In this experiment, the standard sample of glutathione peroxidase, which was received as a powdered material was dissolved in doubly distilled water and subjected to tryptic digestion immediately after that. Therefore, under these experimental conditions, non-enzymatic deamidation of glutathione peroxidase does not take place, or the rate is very slow.

3.6. cap-LC–ESI-MS studies of tryptic digest of fractions C and D

The low molecular-weight selenium-containing protein fractions (fractions C and D) were collected and concentrated by ultra filtration. A part of each concentrated fraction (60 μl) was subjected to in-solution tryptic digestion following the standard protocol used in proteomic research.

In order to identify the proteins which are present in fraction C and D, tryptic digests of fraction C and D were analyzed by cap-LC coupled with ESI-Q-TOF mass spectrometry and MS and MS/MS spectra were recorded. MS/MS spectra were used to identify the peptides present in the digest and, in turn, proteins in the starting fraction. For LC–ESI-MS, 70 min gradient time was applied to have a very shallow gradient and thereby minimize the number of peptides coming out of the column in order to minimize or eliminate signal suppression. MS/MS spectra were searched against the NCBI non-redundant protein data base using the Mascot program from Matrix Science. Tables 1 and 2 show the identified proteins in fraction C and D, respectively, with relevant data such as Mowse score, nominal mass and accession number.

According to Table 1, there are six proteins with Mowse scores greater than 65 in fraction C (according to Mascot search results, proteins with scores greater than 65 are significant). Out of those six, five are Brazil nut proteins. Four of five Brazil nut proteins belong to 2S albumins and their nominal masses are around 16–17 kDa (for the proteins S14946, CAA00798, BAA96554) and 3–4 kDa (for A25802). According to size exclusion column calibration, proteins in fraction C should be in that molecular weight range. The other Brazil nut protein, which is present in fraction C, according to Mascot search, is 11S globulin. 11S globulin is a hexamer made of identical monomers of 52 kDa. Therefore, such a protein cannot be among the proteins in fraction C as the fraction was collected under non-denaturing conditions. As a result, the presence of 11S in the fraction C is suspicious even though it has a high Mowse score. One possible explanation is the presence of a protein or proteins in the fraction C, which have identical amino acid sequences with part of amino acid sequence of 11S. Sequencing of just one polypeptide is enough to identify a protein if that polypeptide is unique to that protein. But, if several proteins generate identical polypeptide or polypeptides, under such a situation identifying those polypeptides imply only the probable presence of one or some or all of those proteins. Since the Brazil nut proteome has not been completely sequenced, it is difficult to explain the presence of 11S globulin in this fraction.

Table 2 shows the presence of seven proteins with Mowse scores greater than 74 in the fraction D (according to Mascot search results, proteins with scores greater than 74 are significant). Their nominal masses are around 16–18 kDa (for the proteins Q9LRC2, CAA00798, S14946, S14947) and 3–4 kDa (for A25208). Out of these seven, six are Brazil nut proteins and five of those represent the 2S albumin family and the other is 11S globulin. Even in this case, the presence of 2S albumin is in agreement with the size exclusion column calibration. But, again the presence of 11S in this fraction is difficult to explain. Further fractionation and simplification of the protein mixture and, ultimately isolating each protein in pure form and then sequencing (i.e., classical biochemistry approach) is essential to answer these questions.

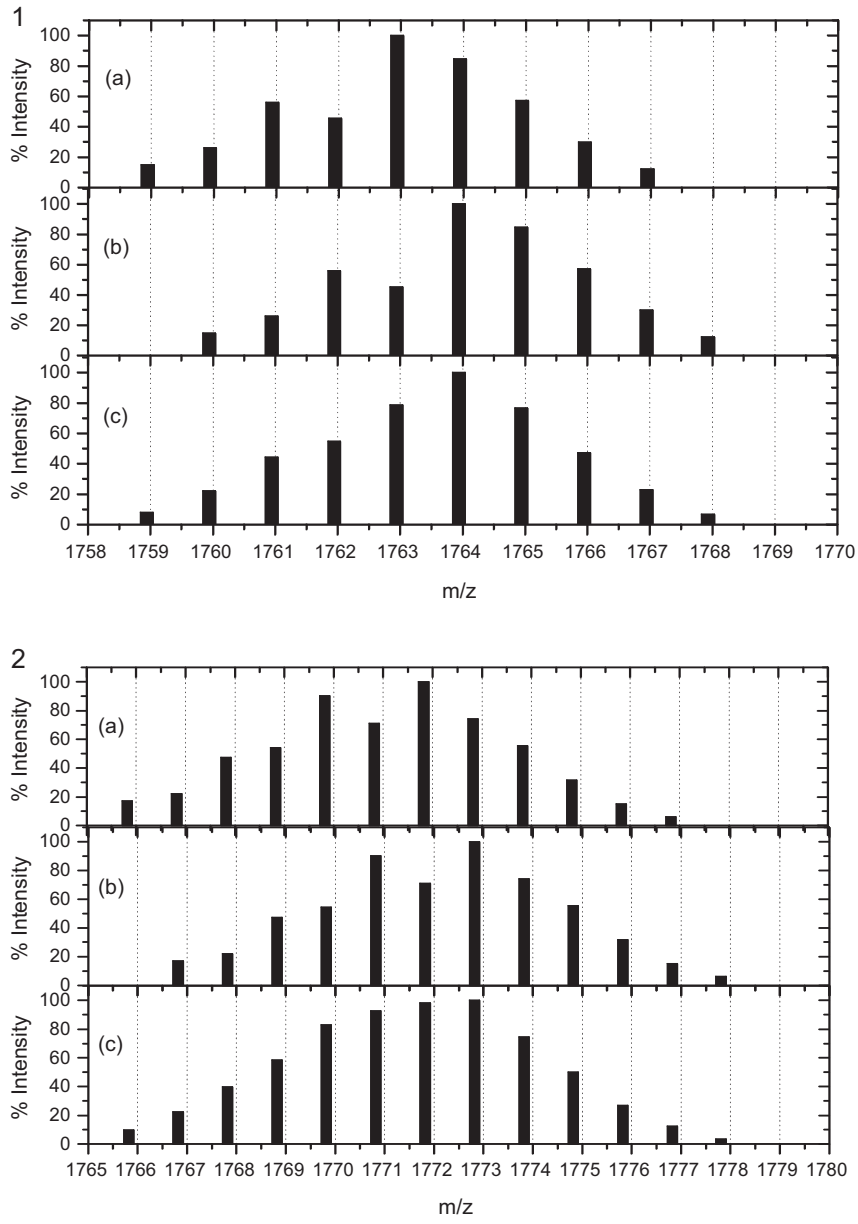

Fig. 4. (1) (a) MALDI spectrum of $C_{80}H_{135}N_{19}O_{20}Se$—100% normal form; (b) MALDI spectrum of $C_{80}H_{134}N_{18}O_{21}Se$—100% deamidated form; (c) MALDI spectrum of the mixture of 50% $C_{80}H_{135}N_{19}O_{20}Se$ [normal form] and 50% $C_{80}H_{134}N_{18}O_{21}Se$ [deamidated form]. (2) (a) MALDI spectrum of $C_{77}H_{130}N_{18}O_{19}Se_2$—100% normal form; (b) MALDI spectrum of $C_{77}H_{129}N_{17}O_{20}Se_2$—100% deamidated form; (c) MALDI spectrum of the mixture of 50% $C_{77}H_{130}N_{18}O_{19}Se_2$ [normal form] and 50% $C_{77}H_{129}N_{17}O_{20}Se_2$ [deamidated form].

According to Mascot search results, both fractions (fractions C and D) consist of similar proteins (except BAA96954 in fraction C and Q0LRC2 in fraction D). But, identified proteins in fraction D have high Mowse score and which indicates identification with high confidence. Therefore, fraction D was selected for MALDI-MS-directed LC–ESI-MS study.

3.7. SEC–ICPMS and MALDI-MS studies of tryptic digest of fraction D

Tryptic digest of fraction D was desalted by C-18 Zip-Tip™ micropipette tips according to the manufacturer's instructions and analyzed by MALDI-MS in the reflectron mode using cyano-

Table 1
Identified proteins in fraction C.

Accession number	Nominal mass	Mowse score	Description
S14946	16,899	308	2S seed storage protein large chain—Brazil nut
CAA00798	16,269	308	Brazil nut 2S—albumin (fragment)
Q84NDZ_BEREX	52,264	250	11S globulin—*Bertholletia excelsa*
BAA96954	16,898	227	AB044391 NID—*Bertholletia excelsa*
A25802	3631	71	2S seed storage protein small chain—Brazil nut
Q8M9S4_9ASTE	13,842	68	ATP synthase epsilon subunit (fragment)—*Desfontainia spinosa*

Table 2
Identified proteins in fraction D.

Accession number	Nominal mass	Mowse score	Description
Q9LRC2	16,898	528	2S albumin—*Bertholletia excelsa*
CAA00798	16,269	434	Brazil nut 2S—albumin (fragment)—synthetic construct
S14946	16,899	433	2S seed storage protein large chain—Brazil nut
S14947	18,242	433	2S seed storage protein large chain—Brazil nut
Q84NDZ	52,264	88	11S globulin—*Bertholletia excelsa*
A25802	3631	79	2S seed storage protein small chain—Brazil nut
AF2694	31,011	79	Hypothetical protein Atu0958

4-hydroxycinnamic acid (HCCA) as the matrix. No peaks with selenium isotope patterns were observed, but, not seeing a signal in MALDI-MS does not necessarily mean that the corresponding fragment is not present in the tryptic digest. It was decided to fractionate the tryptic digest of fraction D by chromatographic methods and analyze the fractions containing Se (identified by ICPMS) by MALDI-MS. After considering the factors related to different chromatographic methods, it was decided to employ size exclusion chromatography with 30 mM Tris–HCl (pH 7.5) as the eluent using a Superdex peptide column. ^{82}Se isotope was selected as detection signal. Fig. 7, in Supplementary information shows SEC–ICPMS profile of tryptic digest D. This profile demonstrates the presence of several Se-containing peptides in the digest and those are distributed throughout the fractionation window of the column. Moreover, comparison of SEC–ICPMS profiles of fraction D before and after digestion indicates that the initial proteins have been broken down to smaller peptides (Supplementary information Fig. 7). Fractions were collected and concentrated by vacuum centrifugation at 37 °C. The concentrated fractions were desalted using Zip-tip pipette tip according to manufacturer's instructions and analyzed by MALDI-MS. But, again, no signal with a Se isotope pattern was observed, even though SEC–ICPMS profile indicated the presence of Se in the digest. As Tris–HCl buffer is not volatile, the concentration of buffer increases as a result of vacuum evaporation. Therefore, the sample may not be completely desalted when it is purified by Zip-tip before MALDI-MS analysis. Though MALDI-MS analysis tolerates salt impurities to some extent, very high salt content suppresses ion formation.

Therefore, it was decided to use 30 mM NH_4Ac (pH 6.8), which is a volatile buffer, as the eluent in the peptide separation of tryptic digests. Fig. 5 shows the SEC–ICPMS profiles of tryptic digest of fraction D when NH_4Ac was used as eluent. This profile also demonstrates the presence of several Se-containing peptides in the digest and those are distributed throughout the fractionation window of the column, but resolution with this eluent is not as good as with 30 mM Tris–HCl. Based on the SEC–ICPMS profile (Fig. 5), five fractions (fractions E, F, G, H and I), were collected and concentrated by vacuum centrifugation at 37 °C.

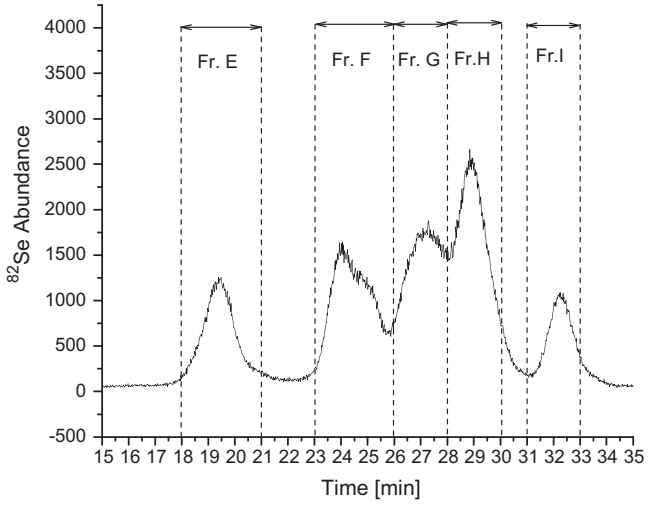

Fig. 5. SEC–ICPMS profiles of tryptic digest of fraction D [eluent—30 mM NH_4Ac].

Out of these fractions, fraction F, G, H and I were analyzed by MALDI-MS (Fig. 6a–g). Observed signals and corresponding fraction numbers are shown in Table 3.

As the space between adjacent lines shows, all these signals are singly charged ions. According to isotope pattern of these signals, deamidation is a possible phenomenon in all ions except the ions at m/z 1004.1 and 1102.2. According to previous studies, Brazil nut 2S proteins contain a high content of Asn and Gln. Therefore, these isotope patterns can be explained by assuming the presence of a mixture of 1:1 deamidated and normal form as discussed earlier.

Even though the tryptic digest of fraction D was analyzed by MALDI-MS-directed LC-EST MS and MS/MS, identification of Se-containing signals through the analysis of extracted ion chromatograms was not successful.

When the non-fractionated tryptic digest of fraction A was analyzed by MALDI-MS, a well-resolved, high intensity signal with Se isotope pattern (though modified to some extent owing to the deamidation phenomenon) was observed (Fig. 3a and b). But, when the same mixture was analyzed by LC–ESI-QTOF-MS, a good signal was not observed even though a significant reduction of sample complexity was achieved by the chromatographic separation, thereby reducing or eliminating signal suppression. Ionization mechanisms in MALDI and ESI are different and a compound which shows a high degree of ionization in MALDI-MS may poorly ionize in ESI-MS and vice versa. Some molecules may ionize well in both ionization methods. This might be a reason for not observing the expected, high intensity Se signal in the LC–ESI-MS analysis of tryptic digestion of fraction A. When the non-separated tryptic digest of fraction D was analyzed by MALDI-MS, no signals were observed. When the fractionated fractions of tryptic digest of fraction D were analyzed by MALDI-MS, several signals with characteristic Se isotope pattern were observed. Therefore, signal suppression should be the reason for not seeing the signals when the non-fractionated tryptic digest of fraction D was analyzed my MALDI-MS.

As mentioned above, when the corresponding extracted ion chromatograms for MALDI-MS signals of tryptic digest of fraction D were constructed from the LC–ESI-MS of the non-fractionated tryptic digest of fraction D, it was observed that all those extracted chromatograms consisted of peaks, which are broad and spread throughout the entire chromatogram, implying that those cannot be used in MALDI-MS-directed ESI-MS and MS/MS studies. This can be solved using simplified fractions of tryptic digests (by SEC or IEC)

Table 3
Fraction number and observed MALDI MS signals in those fractions.

Fraction number	m/z value of the signals
F	No signal
G	1075.6, 1211.6
H	1084.1
I	868.2, 933.3, 1004.1, 1018.2, 1084.2, 1102.2

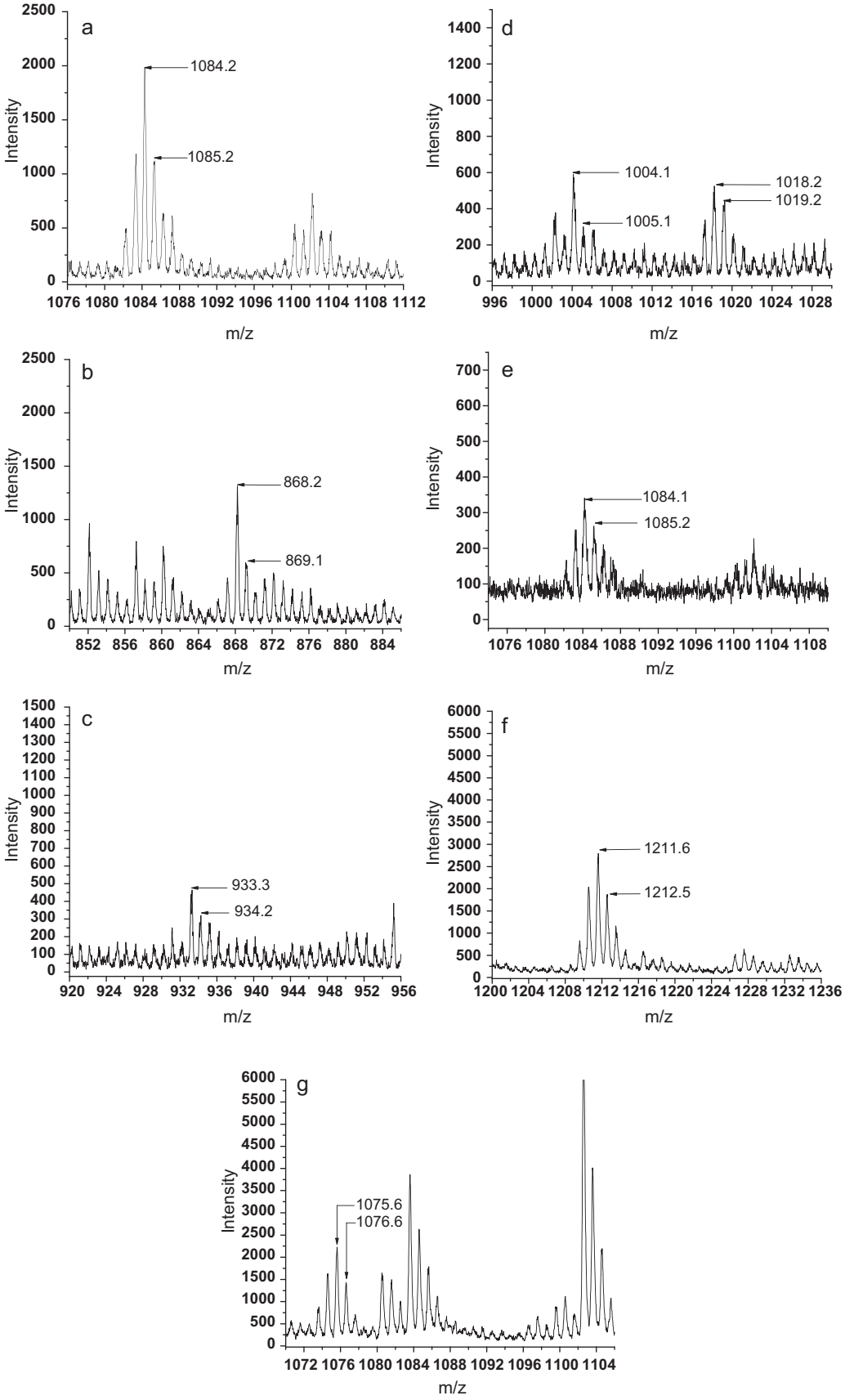

Fig. 6. (a) MALDI spectrum of fraction I between *m*/*z* 1076 and *m*/*z* 1112—[fraction I]. (b) MALDI spectrum of fraction I between *m*/*z* 850 and *m*/*z* 886—[fraction I]. (c) MALDI spectrum of fraction I between *m*/*z* 920 and *m*/*z* 956—[fraction I]. (d) MALDI spectrum of fraction I between *m*/*z* 996 and *m*/*z* 1030—[fraction I]. (e) MALDI spectrum of fraction I between *m*/*z* 1074 and *m*/*z* 1110—[fraction H]. (f) MALDI spectrum of fraction I between *m*/*z* 1200 and *m*/*z* 1236—[fraction G]. (g) MALDI spectrum of fraction I between *m*/*z* 1070 and *m*/*z* 1106—[fraction G].

for LC–ESI-MS step. But, then identification of the proteins is less confident or not possible owing to reasons such as low sequence coverage as a result of the distribution of peptide fragments of a protein in several different fractions. Therefore, it is essential to analyze a non-fractionated fraction of tryptic digest by LC–ESI-MS in order to identify the proteins present in the fraction under study.

In this study, well resolved, high intensity MALDI-MS signals of Se-containing peptides were observed in non-fractionated tryptic digest (fraction A) as well as in fractionated tryptic digest (fraction D). Therefore, sequencing of Se-containing peptides by MALDI-TOF-TOF MS is the most logical step, since some peptides, which have high ionization efficiency in MALDI-MS and produce MALDI-MS signals of high intensity, may poorly ionize in ESI-MS, as observed in tryptic digest of fraction A.

But, only some peptides in a protein digest ionize well in MALDI-MS. Therefore, LC–ESI-MS step is essential in order to identify the peptides which poorly ionize or do not ionize in MALDI-MS. The higher the sequence coverage, the higher the confidence in protein identification. Therefore, it is better to design the experimental and analytical procedures in order to get the maximum sequence coverage, in addition to identifying and sequencing Se-containing peptides. Then, data from these two sources can be utilized to obtain a more detailed knowledge of the Se-containing proteins and proteins in a sample. The disadvantage of ICP MS and MALDI-MS approach is that this method is not suitable to study Se-containing peptides (or other peptides) which poorly ionize in MALDI-MS but show high ionization efficiency in ESI–MS. Another experimental approach involving chemical modifications, ion-ion chemistry and different versions of ESI-MS is in designing stage for Se-containing peptides which poorly ionize in MALDI-MS, but exhibit a high ionization efficiency in ESI-MS.

4. Conclusion

Asccording to SEC-UV-ICPMS profile, selenium-containing proteins in water-soluble protein fraction of Brazil nut are distributed in high molecular weight region as well as in low molecular weight region. According to SEC calibration markers, 11S protein may be a possible candidate for the high molecular selenium-containing protein, while 2S proteins may constitute low molecular weight selenium-containing proteins. Also, this study demonstrated the feasibility of fractional precipitation by ammonium sulfate for achieving the matrix reduction and sample concentration simultaneously under non-denaturing conditions in selenoprotemics studies.

As far as protein and peptide identification is concerned, several Brazil nut proteins (S14946, CAA00789, A25802, BAA96554, Q9LRC2, S14947 and Q84ND2_BEREX [though questionable]) were identified with high confidence in low molecular weight Se-containing protein fractions.

The most important observation in this research is the slightly modified isotope pattern of selenium-containing peptides, most probably owing to the occurrence of partial deamidation of selenium-containing peptides. Anyway, this observed isotope pattern is only an indirect evidence for such a phenomenon and it should be confirmed by separating each species, recording mass spectrum of individual species and sequencing those. To our knowledge, this is the first report of the effect of partial deamidation of selenium-containing peptides on the observed selenium isotope pattern, though the evidences is indirect. Moreover, this observation, that is, effect of deamidation on the observed isotope pattern in MALDI-MS and LC–ESI-MS (under acidic elution conditions) is very important when investigation of metalloproteins by MALDI-MS and LC–ESI-MS is considered. In such studies, metal-containing peptides are identified by their characteristic isotope patterns. Therefore, the possibility of observing modified isotope patterns should be taken into consideration in analyzing mass spectra of such peptides. Therefore, this observation is considered as the most important outcome of this study owing to its effect on the metallomics-related research methodologies which are based on identification of peptides through isotope patterns.

Acknowledgements

The authors are grateful to NIEHS for partial support through Grant Number ES04908 and to Agilent Technologies for loan of the Agilent 7500ce ICPMS instrument. The expert assistance of Dr. Stephen Macha, Mass Spectrometry Facility, Department of Chemistry, University of Cincinnati, with MALDI-MS and ESI-MS is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ijms.2010.12.005.

References

[1] J.J. Berzelius, Acad. Handl. Stockholm 39 (1818) 13.
[2] K. Schwarz, C.M. Foltz, J. Am. Chem. Soc. 79 (1957) 3292–3293.
[3] P.C. Uden, Speciation of selenium, in: R. Cornelis, J. Caruso, H. Crews, K. Heumann (Eds.), Handbook of Elemental Speciation, John Wiley & Sons, Ltd., 2005, pp. 350–351.
[4] L. Flohe, W. Gunzler, H.H. Schock, FEBS Lett. 32 (1973) 132–134.
[5] J.W. Forstrom, J.J. Zakowski, A.L. Tappel, Biochemistry 17 (1978) 2639–2644.
[6] R.J. Shamberger, D.V. Frost, Can. Med. Assoc. 100 (1969) 682–1682.
[7] G.N. Schrauzer, D.A. White, C.J. Schneider, Bioinorg. Chem. 7 (1977) 23–34.
[8] L.C. Clark, G.F. Combs Jr., B.W. Turnbull, E.H. Slate, D.K. Chalker, J. Chow, L.S. Davis, R.A. Glover, G.F. Graham, E.G. Gross, A. Krongrad, J.L. Lesher Jr., H.K. Park, B.B. Sanders Jr., C.L. Smith, J.R. Taylor, JAMA 276 (1996) 1957–1985.
[9] M.P. Rayman, Lancet 356 (2003) 233–241.
[10] C.B. Allan, G.M. Lacourciere, T.C. Stadtman, Annu. Rev. Nutr. 19 (1999) 1–16.
[11] D.B. Cowie, G.N. Cohen, Biochim. Biophys. Acta 26 (1957) 252–261.
[12] K.P. McConnell, J.L. Hoffman, FEBS Lett. 24 (1972) 60–62.
[13] M.X. Sliwkowski, T.C. Stadtman, J. Biol. Chem. 260 (1985) 3140–3144.
[14] M.P. Bansal, R.G. Cook, K.G. Danielson, D. Medina, J. Biol. Chem. 264 (1989) 13780–13784.
[15] M.P. Bansal, T. Mukhopadhyay, J. Scott, R.G. Cook, D. Medina, Carcinogenesis 11 (1990) 2071–2073.
[16] V.N. Gladyshev, S.V. Khangulov, T.C. Stadtman, Proc. Natl. Acad. Sci. U.S.A. 91 (1994) 232–236.
[17] A.J. Antunes, P.J. Markakis, J. Agric. Food Chem. 25 (1977) 1096–1098.
[18] T.B. Osbern, Vegetable Proteins, Longmans, London, 1924.
[19] A.J. Antunes, Food Science and Human Nutrition, Michigan State University East Lansing, MI, 1975.
[20] R.J. Youle, A.H.C. Huang, Am. J. Bot. 68 (1981) 44–48.
[21] P.I. Payne, Seed Proteins, Academic, London, 1983.
[22] S.S.M. Sun, S.B. Altenbach, F.W. Leung, Eur. J. Biochem. 162 (1987) 477–483.
[23] S.S.M. Sun, F.W. Leung, J.C. Tomic, J. Agric. Food Chem. 35 (1987) 232–235.
[24] J.C. Chang, W.H. Gutenmann, C.M. Reid, D.J. Lisk, Chemosphere 30 (1995) 801–802.
[25] S.S. Kannamkumarath, K. Wrobel, K. Wrobel, A. Vonderheide, J.A. Caruso, Anal. Bioanal. Chem. 373 (2002) 454–460.
[26] A.P. Wonderheide, K. Wrobel, S.S. Kannamkumarath, C. B'hymer, M. Montes-Bayon, C. Ponce De Leon, J.A. Caruso, J. Agric. Food Chem. 50 (2002) 5722–5728.
[27] T. Chunhieng, K. Petritis, C. Elfakir, J. Brochier, T. Goli, D. Montet, J. Agri. Food Chem. 52 (2004) 4318–4322.
[28] S. Ma, K.E. Hill, R.F. Burk, R.M. Caprioli, J. Mass Spectrom. 40 (2005) 400–404.
[29] S. Ma, K.E. Hill, R.M. Caprioli, R.F. Burk, J. Biol. Chem. 277 (2002) 12749–12754.
[30] L. Fu, X. Wang, Y. Eyal, Y. She, L.J. Donald, K.G. Standing, G. Ben-Hayyim, J. Biol. Chem. 277 (2002) 25983–25991.
[31] J.R. Encinar, L. Ouerdane, W. Buchmann, J. Tortajada, R. Lobinssk, J. Szpunar, Anal. Chem. 75 (2003) 3765–3774.
[32] S. Mcsheely, J. Kelly, L. Tessier, Z. Mester, Analyst 130 (2005) 35–37.
[33] M. Dernovics, L. Ouerdane, L. Tastet, P. Giusti, H. Preud'homme, R. Lobinski, J. Anal. At. Spectrom. 21 (2006) 703–707.
[34] P. Giusti, D. Schaumloffel, H. Preud'homme, J. Szpunar, R. Lobinski, J. Anal. At. Spectrom. 21 (2006) 26–32.
[35] H.G. Infante, G. O'Connor, M. Rayman, R. Hearn, K. Cook, J. Anal. At. Spectrom. 21 (2006) 1256–1263.
[36] L. Tastet, D. Schaumloffel, B. Bouyssiere, R. Lobinski, Talanta 75 (2008) 1140–1145.
[37] J.R. Encinar, R. Ruzik, W. Buchmann, J. Tortajada, R. Lobinski, J. Szpunar, Analyst 128 (2003) 220–224.

[38] H. Chassaigne, C.C. Cherry, G. Bordin, F. Vanhaecke, R. Rodriguez, J. Anal. At. Spectrom. 19 (2004) 85–95.
[39] L. Tastet, D. Schaumloffel, B. Bouyssiere, R. Lobinski, Anal. Bioanal. Chem. 385 (2006) 948–953.
[40] L. 40 Tastet, D. Schaumloffel, R. Lobinski, J. Anal. At. Spectrom. 23 (2008) 309–317.
[41] M. Dernovics, P. Giusti, R. Lobinski, J. Anal. At. Spectrom. 22 (2007) 41–50.
[42] R.K. Scopes, Protein Purifications—Principles and Practice, Springer-Verlag, New York, 1987.
[43] K.L. Stone, K.R. Williams, The Protein Protocols, Handbook Humana Press, 1996.
[44] T.V. Brennan, S. Clarke, Deamidation and Isoaspartate Formation in Model Synthetic Peptides, CRC Press, Boca Raton, FL, 1995, pp. 65-90.
[45] R.D. Li, A.J. 'Souza, B.B. Laird, R.L. Schowen, R.T. Borchardt, E.M. Topp, J. Pept. Res. 56 (2000) 326–334.
[46] T.H. Wright, Crit. Rev. Biochem. Mol. 26 (1991) 1–52.
[47] R. Bischoff, H.V.J. Kolbe, J. Chromatgr. B 662 (1994) 261–278.
[48] H. Sarioglu, F. Lottspeich, T. Walk, G. Jung, C. Eckerskorn, Electrophoresis 21 (2000) 2209–2218.