

Honey protein extraction and determination by mass spectrometry

Lee Suan Chua · Jun You Lee · Giek Far Chan

Received: 5 October 2012 / Revised: 29 November 2012 / Accepted: 5 December 2012 / Published online: 8 January 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract There are relatively limited studies on the protein of honey samples mainly because of the low amount of protein in honey (0.1–0.5 %), the difficulty in extracting honey protein from the sugar-rich environment, and the hindrance of protein characterization by conventional approaches. Several protein extraction methods such as mechanical (ultrafiltration and ultracentrifugation) and chemical (precipitation) techniques have been applied to different types of honey samples. Most of these studies reported the quantity and molecular size of honey protein from gel electrophoresis, but were unable to identify and characterize the protein. This limitation might be due to the low capacity of analytical equipment in those days. Although different precipitants have also been used, not all them are compatible with mass spectrometric methods during downstream analysis. As a result, the sample preparation step is essential in order to confidently characterize the low and varied amount of honey protein. Nowadays, honey protein is getting attention from researchers because of its potential activity in pharmacological applications. Therefore, honey protein extraction and determination by mass spectrometry are critically reviewed in order to stimulate further honey protein research.

Keywords Honey protein · Protein extraction · Mass spectrometry · Database search

Recent developments in honey protein research

The study of honey protein is extremely limited to date, even though the enzyme content in honey has received much attention as a quality indicator in Western countries. The information on honey protein has also been used to distinguish between honey and artificial blends. To the best of our knowledge, no scientific data are available in the literature regarding the proteome of honey from any origin. The limited study of honey protein is mostly attributed to the small and variable amount of protein in various types of honey samples. In the early 1900s, several investigators reported the presence of small amounts of proteins such as globulin, albumin, protease, and peptone in honey [1]. During that period, numerous biological tests including immunological methods using rabbit anti-serum and anti-bee serum were also developed to differentiate between genuine and artificial honey [1]. The first study that extracted at least 19 protein bands from honey samples was using silver staining of sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) [2]. However, the study did not report any protein identification on the detected spots perhaps because of the problem of compatibility between silver staining and mass spectrometry (MS) and the relative lack of availability of mass analyzers in those days [3].

Ferreres et al. [4] later reported the identification of two pollen proteins with a molecular size of 33 and 36 kDa in sunflower honey, as well as 38 kDa in eucalyptus honey based on SDS-PAGE analysis. They concluded that the pollen protein could be used as a marker for authentication of floral origin. Besides that, multiple protein bands ranging from 10.5 to 72 kDa were also detected in the honey samples of locust, sunflower, chestnut and forest using Coomassie Brilliant Blue R-stained 12 % SDS-PAGE [5]. The investigation characterized the allergenic components in honey by cross-reacting proteins in the immunoblotting assay. Although the assay detected pollen proteins and

L. S. Chua (✉) · J. Y. Lee
Metabolites Profiling Laboratory, Institute of Bioproduct Development, Universiti Teknologi Malaysia, 81310UTM, Johor Bahru, Johor, Malaysia
e-mail: lschua@ibd.utm.my

G. F. Chan
Nanoporous Materials for Biological Application Research Group, Faculty of Bioscience and Bioengineering, Universiti Teknologi Malaysia, 81310UTM, Johor Bahru, Johor, Malaysia

proteins derived from secretions of pharyngeal and salivary glands of honey bee heads, no further protein identification based on the MS approach was carried out. Indeed, the characterization of honey protein based on MS tools is lagging behind compared to the proteomic studies for other biological samples.

In another study, researchers from Korea reported their findings on the identification of major honey proteins with a molecular weight of 56 and 59 kDa produced by *Apis cerana* (Korean native honey bee) and *Apis mellifera* (European honey bee), respectively [6]. Ten years later, the same group characterized the major proteins using MALDI TOF MS after peptide N-glycosidase digestion [7]. They found that both proteins were identified as major royal jelly protein 1 (MRJP1) with a molecular weight of 49 kDa having a 91 % homology. The difference in the molecular weight of the honey major proteins was due to the N-linked oligosaccharide constituting a minimum of 3–5 % of the total molecular weight of the polypeptide. The difference was also noticed in the glycosylation pattern as shown in the peptide mass fingerprinting analysis. The finding also showed that both proteins were similar to MRJP1 with a molecular weight of 55 kDa, which was the most abundant protein in honey bee product [8]. Therefore, MRJP1 is a glycoprotein and it was reported to be identical to apalbumin-1 in the N-terminal amino acid sequencing [8].

Recently, Girolamo et al. [3] reported the detection of eight proteins from the honey samples of chestnut, acacia, eucalyptus, and orange. The detection was carried out by using SDS-PAGE for protein separation and using MS for protein identification. Only one protein was of plant origin (glyceraldehydes-3-phosphate dehydrogenase from *Mesembryanthemum crystallinum*) and the rest belonged to the *Apis mellifera* proteome. A total of five out of the seven belonged to the major royal jelly proteins, and the other two proteins were reported to be α -glucosidase and defensin-1. Girolamo et al.'s findings are considered a major achievement in honey proteome research.

Proteins and amino acids in honey

The identification and measurement of peptides and proteins in honey samples, particularly at the nanogram per milliliter scale level, become a critical area of research because of their importance as potential targets for pharmacological research [9] and clinical diagnostics [10]. Protein is present in honey only in minute amounts, approximately 0.1–0.5 % [6, 11], with molecular weights ranging from 22 to 75 kDa [12]. The amount of protein in honey is usually calculated by multiplying the total nitrogen content from the Kjeldahl method by a factor of 6.25. White and Kushnir [13] found that the nitrogen content of 490 honey samples in the USA

ranged from 0 to 0.133 % with an average of 0.041 %. However, not all nitrogen necessarily belongs to proteinaceous substances. Indeed, only 40–80 % of the total nitrogen is in the protein fraction and the rest might be attributed to free amino acids [3, 13, 14]. The existence of protein often causes honey to have a lower surface tension, hence having a marked tendency to produce foam and form scum and air bubbles [3].

The presence of honey protein has also been applied as an internal standard for stable carbon isotope ratio detection of adulteration [15]. The addition of low-cost sugars to pure honey changes the carbon isotope ratio, but not the protein composition [16]. Honey protein is naturally formed by bees through the enzymatic breakdown of pollen and nectar. Hence, the carbon isotope value of honey and protein must be virtually identical if no synthetic sugar is added. The difference between the carbon isotope ratio of honey and protein can be used to determine the level of adulteration. Cienfuegos et al. [17] compared the accuracy of two pyrolysis methods such as quartz (modified version based on Sofer's approach) and Pyrex (classical) methods for the carbon isotopic analysis. It was found that the quartz method produced more accurate results owing to the complete combustion of honey samples. This technique has been adopted as an Association of Official Analytical Chemists (AOAC) official method for detecting corn or cane sugar (C4 plant sugars) adulteration of honey. MS of the stable carbon isotope ratio technique is the first MS method accepted by the AOAC [18] to authenticate honey adulteration by synthetic sugars. The other analytical techniques including chromatography methods are claimed to be not sensitive enough to detect very low concentrations of C4 plant sugars in parallel with the advancement of honey adulteration techniques. Somehow, this technique is not suitable for the detection of honey adulteration with C3 plant sugar such as beet sugar [19].

Twenty-six amino acids, which accounted for 0.3–1 % (w/w) of the total honey weight, have been reported in honey samples [20, 21]. Most of them are present in the bound form. Proline appears to be the most predominant amino acid in floral honey, followed by phenylalanine and glutamic acid [21, 22]. Since most of the amino acids are present in trace quantity, only the content of proline is included in the European Union food laws for the quality parameters of honey. The amount of proline could be used to represent the total amino acid content [23]. Work by von der Ohe et al. [24] also proved that proline came from the honey bees after comparing the result to the honey produced from sugar-fed colonies. As a consequence, the amino acid profile could be used to determine the geographical [21, 23, 25] or botanical origin [22, 26–28] of honey. In some cases, the amino acid composition is an important precursor for aroma and flavor of food products [29].

Origin of honey protein

Honey protein could originate from honey bees and/or plant origin. The plant origin of honey protein could be the nectars accumulated from different floral pollens [30]. Somehow, the amount of protein from plant origin was found to be lower than that from honey bees [30]. It is also interesting to note that the kind and amount of protein in honey always vary according to the bee species and environmental factors [6, 7]. For example, the honey protein content from *Apis cerana* and *Apis mellifera* is about 0.1–3.3 % and 0.2–1.6 %, respectively [6]. Bauer et al. [5] revealed that the three protein bands at the molecular masses of 54, 60, and 72 kDa in four types of honey samples, namely locust, sunflower, chestnut, and forest, were suggested to be of bee origin.

The variation in protein molecular weight and protein characteristics might be due to the difference in sugar chain length and glycosylation pattern (branching and anomericity). Mostly, protein is prompted to be glycosylated in honey because of the polar environment. On the basis of the finding by Won et al. [7], the major protein of European honey and Korean native honey (MRJP1 which is identical to apalbumin-1) is a glycoprotein constituting at least 3–5 % of N-linked oligosaccharides from the total molecular weight of the polypeptide (49 kDa). Several studies detected major royal jelly protein in honey samples [6–8]. The specific proteins are mostly royal jelly proteins secreted from the glands of young worker bees to feed the larvae and the queen [31]. The honey proteins resemble the royal jelly proteome, except with fewer species and in lower concentration (three to four orders of magnitude) [3].

Honey protein that originated from bees is mostly enzymes secreted from the salivary and hypopharyngeal glands of worker bees such as α - and β -glucosidase (invertase), α - and β -amylase (diastase), and glucose oxidase [7, 32–34]. The presence of these enzymes in honey samples has been known for many years. They have been extensively examined by ion exchange chromatography on DEAE cellulose, gel filtration through Sephadex G-200, and starch gel electrophoresis together with enzyme assays [35–37]. Other enzymes such as catalase and acid phosphatase have occasionally been reported, but have not received as much attention as the former enzymes. Jintanart et al. [38] purified the three kinds of α -glycosidases from honey samples which could be used to measure honey freshness. White and Kushir [35] reported that 13 out of the 18 isozymes of α -glucosidase had a constant ratio of hydrolysis activity upon disaccharides such as sucrose and maltose, hence confirming the nature of α -glucosidase as honey invertase. It is believed that this enzyme is responsible for most of the chemical changes taking places during the conversion of nectar to honey. This is because α -glucosidase is the most

abundant enzyme, approximately 50 % of the total protein in the hypopharyngeal gland of forager bees with 16 peptides [3, 32]. The enzyme has been purified and some of the enzyme characteristics have been investigated in previous studies [39, 40].

The rest of the enzymes, such as glucose oxidase and amylase, each might contribute 2–3 % of the total protein [41]. Glucose oxidase is another important enzyme added to honey by bees. It was partially purified and studied in detail by Schepartz and Subers [36]. White and Kushir [35] isolated two fractions of glucose oxidases from honey samples, corresponding to the molecular weights of 120 and 200 kDa. The enzyme exhibited antibiotic activity through the production of hydrogen peroxide and gluconic acid [42]. According to White [43], the presence of gluconic acid and gluconolactone contributed to the acidity of honey with the pH value being lower than 4 for most honey samples.

It is also important to highlight that amylase content in honey is getting attention from researchers nowadays. This is mainly because amylase can also be used as an indicator of honey freshness [44]. The origin of amylase was identified to be the salivary secretion of bees, based on the detection of amylase in honey produced by sugar-fed bees and the similarities between honey amylase and bee amylase [45]. As the level of amylase in honey varies greatly, the international standard in Codex Alimentarius does not set a limit of amylase in honey. In the works by Schepartz and Subers [36] and Ohashi et al. [41], α -amylase was from the hypopharyngeal glands of forager bees was purified by ion exchange chromatography. The molecular weight of the enzyme was estimated to be 57 kDa from the SDS-PAGE. However, White and Kushir [35] reported that the molecular weights of amylase and α -glucosidase were likely to be 24 and 51 kDa, respectively. They eluted the enzyme peaks using Sephadex gel filtration on dialyzed honey concentrates. The difference in the molecular weight of amylases could be explained by the ionic interaction between honey enzyme and plant protein present in honey, thus leading to the formation of multiple forms of amylases with different apparent rates of migration in gel electrophoresis [35].

To the best of our knowledge, the proteins isolated from honey samples could be of plant or bee origin. Additionally, plant protein in the nectar might be admixed and digested by proteases secreted by mandibular and hypopharyngeal glands, thus detecting as specific protein of bee origin. Therefore, highly expressed enzymatic proteins from the salivary system of bees might provide important insights into the diversified ecological lifestyle of insects on a molecular level [46]. Fujita et al. [46] reported 31 major house-keeping proteins from the two secretory glands of the bee salivary system, namely the postcerebral and thoracic glands. In addition, the pharyngeal gland located in the

honey bee head is also actively involved in the honey production [5].

Honey protein extraction

Usually, honey contains pollen after harvesting from the bee hives. The presence of pollen might increase the protein content in honey. Therefore, honey is centrifuged to remove pollen before protein extraction. Numerous protein extraction methods are available and these methods can be classified into physical and chemical methods. To date, protein extraction techniques remain a challenge in the accurate analysis of protein because of the presence of impurities that affect the performance of gel electrophoresis and MS data. Nevertheless, the ideal extraction method should not only capture the highest possible number of proteins, but also should be compatible with downstream analysis by liquid chromatography (LC) and MS. Significantly, the ongoing development in protein extraction, purification, and identification methods has facilitated the understanding of an increasing number of biological questions.

Physical method for protein extraction

Centrifugation is the most commonly used physical method for protein separation. This physical method is considered the simplest technique because it is carried out without the use of chemicals. Some chemicals might cause difficulty and complexity in analytical work, in addition to high operating costs. Centrifugation involves the use of mechanical force to separate protein from honey samples. For instance, Babacan and Rand [32] reported the application of ultrafiltration to concentrate and subsequently ultracentrifugation to isolate amylase in honey before purification by ion exchange chromatography. The supernatant can be concentrated by ultrafiltration at the molecular weight cutoff from 8,000 to 30,000 and then centrifuged at 5,000–10,000×*g* for 10 min [5, 32]. As a complementary step, ultrafiltration and ultracentrifugation could be coupled to the dialysis method to extract protein from highly complex sample matrixes. This dialysis method is used to remove low molecular weight and interfering compounds, namely sugars from macromolecules, by passive diffusion through a semi-permeable membrane selectively. Honey is normally dialyzed against distilled water overnight at 4 °C to remove sugar [7, 47]. In case of Korean native honey, it was dialyzed against 20 mM Tris-HCl at pH 8.5 for 12 h [7].

Besides dialysis, honey protein can be absorbed and captured by beads with specific properties. Girolamo et al. [3] used ProteoMiner™ (combinatorial hexapeptide ligand library bead) and C18 beads to capture honey peptidome from honey samples of chestnut, sunflower, eucalyptus,

orange, and acacia. The honey peptidome was then filtered and eluted from the beads by using a suitable solvent system. This concept is actually similar to the principle of affinity chromatography.

SPE-based serum protein extraction prior to MS has been critically reviewed elsewhere [48]. The application of this SPE technique with different chromatographic properties of the stationary phases such as adsorption (reversed and normal phase), ion exchange, and affinity binding has proven advantageous for simultaneous sample cleanup and protein enrichment in a simple, scalable, and high-throughput manner. It is believed that the sensitivity and resolution of MS data are highly dependent on the quality of sample preparation.

Recently, microwave-assisted acid hydrolysis was applied to rapidly degrade protein into or peptides [49] before analysis by MS. Lill et al. [50] demonstrated that microwave-assisted acid hydrolysis was two orders of magnitude faster than the conventional direct oven heating method. Microwaves are high frequency electromagnetic radiation which is strongly absorbed by polar molecules such as proteins [51]. The reaction under microwave conditions can be performed at elevated pressures and temperatures which might be above the boiling point of the solvent submitted to microwave irradiation. A strong localized heating can occur at the polar groups and results in hydrolysis, degradation, and/or extraction of some constituents from the sample matrix. Hence, the microwave-assisted acid hydrolysis of protein will definitely generate some information about the chemical composition of the macromolecules.

Chemical extraction by precipitation method

Of the honey protein extraction techniques, chemical precipitation is the most widely used method because of its speed, simplicity, and applicability to a wide range of biological samples [52]. In most cases, combinations of both physical and chemical protein extraction processes are applied to optimize the honey protein extraction. Previously, the protein content in honey was estimated by determining the volume of precipitate with tannin, phosphotungstic acid, alcohol, or trichloroacetic acid [14, 53]. Nelson and Cohn [54] reported the first attempt to use 80 % ethanol to precipitate α -glycosidase from honey. Papadakis [55] later proved the detection of α -glucosidase activity in ethanol precipitate after separation using chromatographic alumina.

Recently, Padovan et al. [56] reported the use of 10 % sodium tungstate solution with 2–3 N of sulfuric acid to precipitate honey protein, in combination with the carbon isotopic ratio for the determination of honey purity. Before precipitation, the honey solution containing the precipitant was heated in a water bath with constant shaking at 70 °C for 4 h. The protein pellet was observed after centrifugation and washed repeatedly with distilled water until the

supernatant was clear and free from sugars. This work has been done by referring to the method described in the AOAC official method 998.12 [57].

Polson et al. [58] evaluated various protein precipitants for their effectiveness in plasma protein removal and the extent of ion suppression (matrix effect) in LC-MS/MS. They concluded that trichloroacetic acid (TCA) was good for protein precipitation and could be detected by LC-MS/MS using a mobile phase consisting of methanol in water or acetonitrile in water. Afzuni-Zadeh et al. [51] successfully used cold 5–10 % TCA to precipitate protein oxidation markers after centrifugation. The combination of 5 % TCA and 5 % sulfosalicylic acid (SSA) was used to precipitate protein from honey samples. The honey solution was boiled together with the precipitant solvent system for 30 min and cooled to room temperature before analysis. However, no protein from the gel could be detected after SDS-PAGE. This observation could be explained by the limited amount of protein precipitated through this method. It was found that the TCA protein precipitation method is commonly and effectively applied for plant protein extraction [59]. Besides that, TCA was likely to be effective for drone egg protein extraction [60]. Therefore, this method is presumed to be suitable for protein extraction from biological samples.

Other MS-compatible precipitants such as ammonium sulfate [61, 62] and ammonium acetate [63] are also widely used to precipitate protein. Ammonium sulfate was successfully used to precipitate α -glucosidase from honey produced by European honey bees (*Apis mellifera* L.). The honey samples were dialyzed against running tap water for 3 days before precipitation by ammonium sulfate. The precipitate was then mixed with Celite 545 and packed into a column under reduced pressure for salting-out chromatography analysis. It is difficult to compare the efficiency of ammonium sulfate and ammonium acetate for honey protein precipitation. This is because, to our knowledge, no study has been performed to precipitate honey protein by using ammonium acetate. Nevertheless, ammonium acetate had been reported to efficiently precipitate plant protein after phenol extraction. The protocol was successfully carried out on the plant samples from banana, apple, potato, and tomato [63, 64]. Sheoran et al. [63] reported the efficiency of plant protein extraction from tomato pollen using phenol in Tris-HCl buffer at pH 8.8 before protein precipitation by ammonium acetate. This method appears to be superior to the protein extraction method by isoelectric focusing buffer and Tris-HCl buffer, in terms of gel electrophoresis separation, MS detection, and peptide mass fingerprinting identification.

It is also a strong possibility to detect glycoprotein in honey because honey is a supersaturated sugar product produced by honey bees. For glycoprotein, Steinhorn et al. [61] used β -glucosyl Yariv reagent to precipitate arabinogalactan protein (~110 kDa) from New Zealand kanuka honey.

A wide range of arabinogalactones have been shown to specifically bind β -glucosyl Yariv reagent [65, 66]. Arabinogalactone is a glycoprotein that covalently links protein with sugar moieties consisting of galactose and arabinose [66]. They used the method described by Gane et al. [67] to decompose the red-colored Yariv phenyl glycoside dye.

Total protein content

There are two common methods widely used to determine the total protein content in honey samples, namely Lowry [68] and Bradford assays. The former assay involves the use of copper [69] and Folin-phenol reagent [70] to produce a protein complex for colorimetric measurement. The latter depends on the formation of UV-absorbing protein–dye complex from Coomassie Brilliant Blue. Bovine serum albumin (BSA) is usually used as the external standard in the total protein assay.

Although the Lowry method is a commonly accepted colorimetric approach in the total protein determination, it is reported to have a relatively slow reaction rate and non-linearity in the standard curve preparation. The other significant disadvantages include the lack of specificity and stability of some reagents [71]. The reagents that might interfere with the result of this assay include EDTA [72], Tris [73], thiol reagents [74], and carbohydrate [75]. Besides these organic substances, the presence of potassium and magnesium in the sample could also interfere in the Lowry assay. On the other hand, several researchers reported the advantages of the method after having slightly modified and increasing the time of analysis [76–78].

The Bradford method does not suffer from the limitations described in the Lowry assay. Coomassie Brilliant Blue G-250 reagent will bind to the protein molecules and transform them into a blue protein–dye complex, which can be measured at 595 nm. This method is reported to have higher sensitivity, fast reaction rate (approximately 2 min), and more stable protein–dye complex (approximately 60 min) for spectrophotometric measurement. It is also suitable for automation, because the Bradford assay is easily applied and reliable for a number of sample matrices.

SDS polyacrylamide gel electrophoresis (SDS-PAGE)

One of the most popular approaches in protein characterization is an immunological method. Peptides and proteins in biological matrices have traditionally been measured by immunological methods such as radioimmunoassay (RIA) and enzyme-linked immunosorbent assay (ELISA). However, these methods have several drawbacks including the need for antibody reagents, low selectivity to differentiate

the peptide and its fragments, as well as low dynamic range. Similarly, the chromatographic separation technique, namely HPLC combined with typical detectors like a diode array detector and fluorescent detector, also needs standard chemicals for compound identification. Nevertheless, the immunoaffinity assay, namely stable isotope standards and captured by anti-peptide antibodies (SISCAPA), has shown its capacity in the enrichment of low abundance proteins before quantification [79]. The results highlighted that the selection of protein identification method is largely dependent on the chemical nature of protein and its quantity in the sample.

Another conventional technique for protein and peptide separation is gel electrophoresis. This method has limited mass accuracy and resolution [80]. Nowadays, a two-dimensional gel electrophoresis is widely used for resolving complex protein extracts. It is considered the most powerful proteomic tool for the separation of proteins. BSA or standard protein marker mixture is usually used as the external standard in SDS-PAGE analysis. The external standard needs to be dissolved in buffer solution and denatured by heating in a boiling water bath for 4 min [34]. For instance, BSA (15 µg) is added into sample buffer solution (125 mM Tris-HCl, pH 6.8) containing glycerol (10% v/v), SDS (2 % w/v), mercaptoethanol (5 % v/v), and bromophenol blue (0.01 % w/v). The protein spot is separated on the polyacrylamide gel according to protein mass and stained with Coomassie Brilliant Blue for visual detection.

In gel digestion of protein

Since the protein band in the SDS gel is stained for visual detection, the band must be destained in ammonium bicarbonate (25 mM) in 50 % acetonitrile before digestion [34, 60]. The destaining solution can also be a solvent system containing potassium ferricyanide (15–30 mM) and sodium thiosulfate (50–100 mM) [34, 81]. The destaining solution might need to be changed every 10 min and the process might take up to overnight.

MS-compatible detergents are usually applied to ensure full solubilization and unfolding of protein prior to digestion. The most frequently used detergents, such as 3-[(3-choramidopropyl)dimethylamino]propane-1-sulfonate (CHAPS), Brij35, Triton X-100, dimethyl sulfoxide (DMSO), glycerol, Tween, and Lubrol, were reported to likely interfere with MALDI and ESI MS analysis [82]. Yeung et al. [83] mentioned the requirement of hydrophobic protein detergent removal prior to MS analysis. As an alternative to classic detergents, acid-cleavable detergents have recently been proposed for proteomics, primarily because of their minimal interference with LC separation and MS analysis [84, 85]. Nowadays, a number of effective MS-compatible protein solubilizers such as Invitrosol™ (Invitrogen), RapiGest™

(Waters), and Protease MAX™ (Promega) are also commercially available.

For digestion, the protein band is often excised into small pieces (1 mm) for effective enzymatic digestion, usually using modified porcine trypsin or chymotrypsin in ammonium bicarbonate solution (25–50 mM) overnight at 37 °C [84, 85]. Trypsin preferably cleaves the protein on the C-terminal side of lysine or arginine residues. It is considered the most prevalently used enzyme in the digestion because lysine and arginine are present at about 11.7 % of all peptide residues. Consequently, peptides that are digested by trypsin typically contain between 6 and 25 amino acid residues with an average of about 10 amino acid residues [86]. Such relatively short peptides can be easily ionized and detected by most mass spectrometers. Furthermore, the presence of the basic lysine or arginine at the C-terminus, in conjunction with the basic N-terminus yields doubly charged ions during ionization. These doubly charged precursor ions will collide with inert gas molecules such as nitrogen or argon in the collision cell of a mass spectrometer before being fragmented in a predictable way that provides useful information in protein identification. Li et al. [60] removed the excessive trypsin to avoid auto-digestion of trypsin itself before incubation overnight. The purpose of the digestion is to break down the large protein into smaller peptide fragments to allow protein identification. After fragmentation, there is no intact protein left which could be soluble. The smaller peptide fragments increase the solubility of protein and can be easily ionized in the ion source of a mass spectrometer [87]. This proteolytic digestion can be quite variable because the digestion is highly dependent on the individual protein structures such as disulfide bridges, specific fold, solubility, and glycosidation site [79]. In order to obtain reproducible results, the conditions including the digestion time need to be kept constant. According to Arsene et al. [88] and Agger et al. [89], the maximum peptide signal can be observed in a short period of digestion time before reaching steady state conditions.

Even though different reducing agents (2.5 % dithiothreitol (DTT); 50 mM DTE in 25 mM ammonium bicarbonate), alkylation (100 mM iodoacetamide (IAA) in 25 mM ammonium bicarbonate) agents, and digestion conditions have been used in the previous studies, there is still limited research comparing the different protocols with regard to the absolute quantification, as well as optimizing the digestion process. Somehow, Proc et al. [90] recommended the use of sodium deoxycholate with 9 h of digestion time as the optimum condition for human plasma proteins. Besides that, Matysiak et al. [91] used 2.5 % DTT to reduce disulfide bonds in the peptides of honey bee venom. They identified 16 known peptide sequences and discovered a new peptide from bee venom samples. Therefore, it is important to perform the chemical reduction and alkylation in order to

cleave non-specific disulfide bonds of cysteine residues, to improve MS signal quality and to simplify MS analysis [80]. The reduction can prevent both intramolecular and intermolecular disulfide bonds formation [79]. The alkylation of protein aims to modify the sulfhydryl groups from the reformation of disulfide bonds [79]. To ensure that the reduction and alkylation were completely carried out, the intact proteins from prokaryotic and eukaryotic microorganisms were reduced and alkylated by using both DTT and IAA subsequently in the dark. The excess IAA could then be quenched with the addition of DTT reducing solution [92]. The tryptic peptides are desalted by a C18 Zip Tip (Millipore Corporation, Bedford, MA, USA). The desalted sample can be further purified by using resin to wash out the unbound peptide in the eluant, 0.1 % trifluoroacetic acid in water.

The peptides can then be extracted from the gel using a suitable solvent system. Peiren et al. [31, 93] used 60 % acetonitrile containing 0.1 % formic acid for the extraction of protein from samples of honey bee worker venom. Pontoh and Low [34] used 50 % acetonitrile containing 5 % trifluoroacetic acid to extract β -glucosidase from the ventriculus and honey sac of *Apis mellifera*. Fang et al. [94] used 50 % acetonitrile containing 5 % formic acid to extract human BK virus structural protein. However, Sundqvist et al. [80] used only 5 % acetonitrile containing 0.1 % formic acid to extract human proteome resource protein epitope signature tags. Li et al. [60] extracted the trypsinized peptides from the drone embryo samples in two stages. Initially, the soluble proteins in the samples were extracted with aqueous 5 % (v/v) trifluoroacetic acid. Then, the remaining peptides were extracted again but under less acidic conditions with 50 % (v/v) acetonitrile containing 2.5 % (v/v) trifluoroacetic acid. Hence, the peptide extraction system is varied depending on the nature of protein as well as the source of the sample.

Protein separation and identification by LC-MS/MS

Recent advancement and innovation in LC-based separation have helped achieve high-throughput analysis in proteomics studies. The chromatographic performance has dramatically increased over the last decade, from the introduction of reversed-phase capillary columns (3.5 μ m) to the development of sub-2- μ m stationary phase ultra-performance liquid chromatography (UPLC) systems without compensation in separation efficiency. It was found that a 15-min run in UPLC can have a similar peak capacity as generated from a 60-min run in a traditional HPLC system [95]. Later, the micro-scale separation monolithic stationary phase consisting of a piece of solid allowing for high flow rate at low column back pressure [82, 96, 97] was introduced as the

most promising chromatographic design for complex peptide mixture separation. It is important to note that the large peak capacity should be able to provide higher protein coverage.

Development has progressively improved with the introduction of multidimensional LC systems based on the sequential separation in a biphasic column involving stationary phases with different properties. This two-dimensional LC setup prior to MS analysis was introduced into the field of proteomics nearly 10 years ago [82]. This technique is also known as multidimensional protein identification technology (MudPIT) [87]. Delahunty and Yates III [87] applied this technique by placing a strong cation exchange (SCX) resin directly upstream from the C18 resin in the nano-column for peptide separation. The SCX resin was used to store peptide that was then separated in the reversed-phase (RP) column using an acetonitrile gradient. After re-equilibrium, another fraction of peptides was transferred from the SCX resin onto the RP resin with an increase in salt concentration. Complex mixtures of peptides frequently contain salts which can be desalted before separation. This is because high contents of salt might interfere with the interaction between the peptide and SCX resin. This iterative process was repeated until the reserve of peptides on the SCX resin had been exhausted. Washburn et al. [98] and Wolters et al. [99] demonstrated excellent results in using a multidimensional RP ion exchange LC-MS/MS system to quantitatively analyze the proteome of all functional and physical classes from the most abundant to the least abundant proteins/peptides in a complex peptide mixture with a dynamic range of 10,000 to 1. Link et al. [100] achieved a higher resolution and capacity in two-dimensional separation of protein components in the yeast of *Saccharomyces cerevisiae* ribosome using an in-line system with a biphasic nano-column. Hence, MudPIT could increase the number of low abundance proteins detected from the mixture with equal sensitivity for high abundance protein. It has been applied to identify proteins from a wide variety of samples, including membrane proteins, post-translational modification, and protein expression using stable isotope labelling. To the best of our knowledge, no study has been conducted to identify honey protein using the MudPIT method.

An analytical platform for a two-dimensional RP-RP nano-UPLC system is currently being developed by Waters (Manchester, UK). This platform aims to provide highly reproducible and excellent separation performance for RP chromatography in both dimensions at different pH values [101]. The application of a chromatographic gradient to a highly complex mixture is of great importance to reduce ion suppression during the ionization process for peptide identification. Therefore, the success of protein identification is highly dependent on the chromatographic gradient used to

introduce the sample into the mass spectrometer. The application of multidimensional liquid chromatographic platforms combined with other analytical tools in proteomics has been gaining attention from researchers recently. LC integrated with a mass spectrometer is considered a hybrid system for comprehensive protein separation and identification.

LC integrated with a single mass separation device has its limitation in detecting lower concentrations of proteins in the presence of more abundant proteins [102]. Therefore, LC coupled with a tandem mass spectrometer is the technique of choice to improve both selectivity and sensitivity up to femtomole level [9]. This multiple mass-resolving device is crucial to identify complex biomolecules, particularly proteins in honey samples. Numerous successful studies have proven that the integration of LC with a tandem mass spectrometer provides a complementary technique to gel-based methods, mainly because of the high speed and accuracy of compound identification. The high resolution, sensitivity, and wide dynamic range of the tandem mass spectrometer generate vast amounts of mass spectral data, thus enabling the direct analysis of complex mixtures for entire proteomes in a single technology platform [87].

The significant advantage of tandem MS for protein identification is the inherent ability to sequence peptides directly from the mixture without any sample pretreatment prior to injection [87]. It allows the direct identification of individual peptides from protein complex in a wide range of physiological functions. Nowadays, MS has become an indispensable analytical tool for the analyses of complex biological samples, as a consequence of the development of two soft ionization techniques, namely matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) [103–105]. Therefore, these two major ionization techniques have been widely used for protein identification in MS: MALDI TOF MS for protein fingerprinting and ESI MS/MS for peptide sequencing.

The powerful MS detector is certainly required for the identification of many peptides from a highly complex mixture of honey samples. MS-based methods for the identification of proteins have become a standard platform in proteomics nowadays [87]. Isotope-labelled standard is no more as a prerequisite for exact quantification [79, 82]. Although isotope labelling compensates matrix effects and increases the dynamic range, it might also increase the complexity and the cost of the assay.

Although the analytical capacity of the mass spectrometer is growing and generating ever larger data sets, scientific information about the honey proteome is scarcely found in the literature. The difficulty in honey protein identification might be mainly due to the complication of the protein being in a super viscose sugar-rich environment. The application of proteomic approaches to honey samples is restricted by several factors such as protein abundance, size, hydrophobicity,

and other electrophoretic properties [106]. The use of MS data for quantification of both the relative and absolute amounts of particular proteins/peptides is not straightforward, and a number of strategies have therefore been elaborated to address this issue [107, 108].

MALDI TOF MS

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI TOF MS) has proven to be a powerful tool, particularly in a typical bottom-up proteomic approach [93]. The wide application of MALDI TOF MS in protein and peptide characterization is mainly due to the advantages such as high sensitivity, high speed of analysis, robustness, and the prevalence of singly charged ions [109]. It is mainly used for semi-quantitative analyses because the last factor simplifies the interpretation of the spectra [110]. Furthermore, the measurement of individual samples can be easily automated and adapted for higher throughput with the use of MALDI as the ionization method [91]. Operating a MALDI TOF MS system is a much easier task than running a sample using ESI MS/MS.

In a TOF mass analyzer, all ions are accelerated into either an axial or orthogonal flight tube. The resolving power of the TOF system is dependent on the accelerating voltage and length of the flight tube, which are typically one to two orders of magnitude better than a quadrupole system [9]. Since the same kinetic energy is applied to all ions, the time taken to reach the detector is closely related to the mass of the compound.

A wide variety of samples can be analyzed by MALDI TOF MS dependent upon the type of matrixes [111, 112]. The most common UV-absorbing matrix system is cyano-4-hydroxycinnamic acid in a solution containing 30–70 % acetonitrile and 0.1 % trifluoroacetic acid at a concentration of 5–20 g/L [93, 113]. The matrix solution is mixed with the digested protein and then directly spotted on the MALDI target plate; the sample is dried with a gentle stream of air (~35 °C). The mass spectra are registered in a reflection or linear positive ion mode by accumulating data from 200 to 1,000 laser shots for each spectrum with a mass accuracy between 50 and 75 ppm for peptide mass fingerprinting [34, 91].

Matysiak et al. [91] reported the use of two standard matrix systems, namely cyano-4-hydroxycinnamic acid in acetonitrile containing 0.1 % trifluoroacetic acid and sinapinic acid in the same solvent system, for honey bee venom characterization. Both matrix systems gave complementary results. Somehow, the number of peptides detected in the sinapinic acid was found to be higher than that detected in the cyano-4-hydroxycinnamic acid matrix system [91].

Typically, the protein sample is digested by proteolytic enzyme and a mass spectrum is then acquired which contains the masses of all peptides. These masses are used as a

fingerprint to match protein in a database. The identified protein is usually the protein that hits the highest score in the ranking. Therefore, the results from both MALDI and ESI measurements are often complementary, thus applying both techniques in parallel would produce more comprehensive analysis of the samples.

ESI MS/MS

The principle of ESI MS/MS analysis in protein identification is different from MALDI TOF MS. The peptide mixture from the proteolytic digestion is separated by an HPLC system and then fragmented by collision-induced dissociation in a tandem mass spectrometer. The tandem mass spectrometer is mostly a triple quadrupole MS/MS which consists of a vacuum chamber containing two quadrupole mass filters, separated by a collision cell in between both quadrupoles. The first quadrupole is used in mass-resolving mode to select the right precursor ion, whereas the second quadrupole is also used in mass-resolving mode, but to select the daughter ions fragmented from the collision cell. Often, several thousand MS/MS spectra are generated for each fragmented peptide. The MS data are used to search for the matched protein in a database. A protein hit is identified by multiple independently sequenced peptides from the same protein.

Usually, nanoLC-ESI MS/MS is applied especially for identification purposes owing to its high sensitivity and reliability [91]. The high sensitivity of this system is because of the small diameter (75 μm) of the nanoLC reversed-phase C18 column which can effectively concentrate peptides from 50- to 200-fold before MS detection. The increase in peptide concentration translates into an increase in MS signal. Owing to the high sensitivity, protein identification is always confidently determined on the basis of independent MS/MS sequencing of only one peptide. Protein identification based on one peptide sequence with 70–80 % coverage by MS/MS will be at the 90 % confidence level. The confidence level reaches 99 % if protein identification is based on MS/MS sequencing of two independent peptides. The confidence level will become 100 % if three peptides are sequenced. In contrast to MALDI MS fingerprinting, 10 to 20 peptides are required for confident protein identification. The utilization of samples that are loaded on a MALDI plate is relatively small (1/1,000 to 1/10,000), whereas almost all the samples that are injected into the electrospray ion source are used for identification. Therefore, the higher capacity in utilizing more samples will result in higher sensitivity.

Proteins and peptides are protonated in the ESI source and form multiply charged ions with mass to charge ratios averaging from 200 to 2,000 amu. Different degrees of protonation might coexist in collision-induced dissociation

and multiple peaks would be observed for each peptide because peptide bonds are relatively weak [79]. Thus, the selection of specific transitions with the most intense signal for the fragments is essential for protein identification. Usually, the best two to four transitions per peptide are selected for quantitative determination [79]. This selected reaction monitoring (SRM), also called multiple reaction monitoring (MRM), approach was proven to be reproducible and reliable in protein quantification across several laboratories by Addona et al. [114]. They revealed that the main variance in protein identification was due to the sample preparation process. It means that efficient upstream sample preparation is equally important for successful shotgun analysis of protein [115]. Many low abundance proteins in the range of nanogram per milliliter are still often not detectable without additional sample processing steps such as sample fractionation, solid-phase extraction, immuno-depletion of the abundant proteins, or antibody enrichment of the targeted protein [79]. It is essential to make the protein completely solubilized and denatured throughout the proteolysis step before identification [116].

It is also important to note that the fragmentation pattern of a particular peptide might vary from instrument to instrument. The variance is mostly due to the difference in instrument settings such as collision energy and ion source voltage. It is common to choose doubly and triply charged precursor ions with the fragment ions at greater m/z ratio than the precursor ion. Fragmentation usually results in the production of two complementary peptide ions which are termed the y-ion and b-ion. The y-ion retains a positive charge at the C-terminal end, whereas the b-ion retains the charge at the N-terminal end. The fragmentation pattern provides amino acid sequence information.

In the last few years, the MRM³ approach has become a promising strategy to improve the sensitivity, accuracy, and scanning speed of mass spectrometers by using a hybrid triple quadrupole and linear ion trap system. This technology allows secondary fragmentation in the third quadrupole by trapping the daughter ions for a certain period of time before releasing them for detection. Fortin et al. [117] showed that MRM³ coupled to conventional LC resulted in a three- to fivefold improvement in the limit of detection and quantification for serum prostate specific antigen (PSA) level determination.

Database search for matching protein

There are many search engines on the Web, both free and commercially available, for performing database searches by using MS/MS spectra. In the first step of a database search, the experimental MS spectrum is compared to the theoretical spectrum in the database. Mostly, the search

engines use the ubiquitous FASTA format for supporting protein sequence databases. The two major repositories for sequence data are hosted by the National Center for Biotechnology Information (NCBI) in the USA (<http://www.ncbi.nlm.nih.gov>) and the European Bioinformatics Institute (<http://www.ebi.ac.uk>). A protein sequence database (~6,681 sequences) for *Apis mellifera* is available from the NCBI website. The most commonly used databases include the universal protein resource (Uniprot) and Flybase databases. Uniprot is a high-quality, curated, and non-redundant protein database. It contains a consensus sequence for each distinct protein and known variants have been combined into a single entry. Hence, the size of Uniprot is relatively small and the search time is significantly reduced. The searching process might take from a few minutes to a few hours and sometimes up to a few days depending on the complexity of data and the size of database being searched [87]. The protein can also be grouped on the basis on its biochemical functions from the combined results of both databases.

Sequence similarity is investigated by using the protein–protein BLAST algorithm which is available at the website of <http://www.ncbi.nlm.nih.gov>. Usually, the homology search is carried out with a mass tolerance of 100 ppm in the MS mode and 0.4 Da in the MS/MS mode for data generated from MALDI MS, although a mass tolerance of 0.4 Da in the MS mode and 0.8 Da in the MS/MS mode are permitted for data obtained from LC-MS/MS. It is advisable to set the mass tolerance to 1–2 Da, because a narrow mass tolerance might fail to obtain a good match, unless the search engine performs calibration before search. If one uses the MASCOT search engine (<http://www.matrixscience.com>), the data file first needs to be converted into MASCOT format and only one missed cleavage is allowed in the search. Often, carbamidomethyl of cysteine and oxidation of methionine are selected as variable modifications [59].

Once peptide hits are scored, manual validation is required to verify the identified protein in order to avoid false positive identification. Indeed, database matching using the peptide sequences is surprisingly difficult because many peptide sequences can be assigned to more than one protein. Therefore, the quality of database being searched is a critical issue. Therefore, an algorithms for the estimation of false discovery rate have been developed such as PeptideProphet and ProteinProphet software [118, 119] and PLGS software (Waters, Manchester, UK). One of the most reliable and simple ways to validate the false discovery rate is by using a target-decoy search [120]. This can be done by using the identical search parameters against a database in which the sequence has been reversed or shuffled. The number of matches from the decoy database is used to estimate the number of false positives in the result. However, this

validation method is not suitable for giving a good estimation for a small number of spectra [120].

Protein identification is typically acceptable if the probability is higher than 95 % and contains at least two identified peptide sequences having the maximum peptide coverage (70–80 %). A score which is higher than 30 indicates a confidence of protein identification at 95 % [59]. A database search score lower than 95 % probability would produce unambiguous results [60].

Conclusion

The findings summarized in this review revealed that honey protein mostly originates from the salivary and hypopharyngeal glands of worker bees. A small portion of honey protein might also be of plant origin or result from the enzymatic reaction between bee saliva and plant pollen. Usually, a multistep sample treatment was required such as ultrafiltration, ultracentrifugation, and followed by chemical precipitation in order to concentrate honey protein from a highly viscous and sugar supersaturated sample matrix. The advancement and innovation of analytical instruments, particularly LC-MS/MS, have made protein identification and characterization simpler, easier, and possible even for low abundant protein. Recently, the combination of gel electrophoresis and LC-MS/MS has provided a powerful tool for proteome mapping. Mass spectra generated from MALDI TOF MS and/or ESI MS/MS could be matched with a few search engines to allow a high confidence level of protein identification.

References

1. White JW (1957) *Bee World* 38:57–66
2. Marshall T, Williams KM (1987) *Anal Biochem* 167:301–303
3. Girolamo FD, D'Amato A, Righetti PG (2012) *J Proteomics* 75:3688–3693
4. Ferreres F, Garcia-Viguera C, Tomas-Lorente F, Tomas-Barberan FA (1993) *J Sci Food Agric* 61:121–123
5. Bauer L, Kohlich A, Hirschwehr R, Siemann U, Ebner H, Scheiner O, Kraft D, Ebner C (1996) *J Allergy Clin Immunol* 97:65–73
6. Lee DC, Lee SY, Cha SH, Choi YS, Rhee HI (1998) *Korean J Food Sci* 30:1–5
7. Won SR, Lee DC, Ko SH, Kim JW, Rhee HI (2008) *Food Res Int* 41:952–956
8. Simuth J, Bilikova K, Kovacova E, Kuzmova Z, Schroder W (2004) *J Agric Food Chem* 52:2154–2158
9. Thomas R (2001) *Spectroscopy* 16:28–37
10. Ji QC, Rodila R, El-Shourbagy TA (2007) *J Chromatogr B* 847:133–141
11. Jagdish T, Joseph I (2004) *J Agric Food Chem* 52:3237–3243
12. Yilmaz H, Kufrevloglu OI (2003) *GIDA* 28:155–157
13. White JW, Kushnir I (1967) *J Apicult Res* 6:163–178

14. White JW, Rudyj ON (1978) *J Apicult Res* 17:234–238
15. White JW, Winters K (1989) *J Assoc Off Anal Chem* 72:907–911
16. White JW (1992) *J Assoc Off Anal Chem* 75:543–548
17. Cienfuegos E, Casar I, Morales P (1997) *J Apicult Res* 36:169–179
18. Association of Official Analytical Chemists (1990) AOAC official method 991.41. Official methods of analysis, 15th edn. AOAC, Washington DC
19. Korth W, Ralston J (2002) Techniques for the detection of adulterated honey, Rural Industries Research and Development Corporation. Barton, Australia
20. Gonzalez Paramas AM, Gomez Barez JA, Marcos CC, Garcia-Villanova RJ, Sanchez JS (2006) *Food Chem* 95:148–156
21. Rebane R, Herodes K (2008) *J Agric Food Chem* 56:10716–10720
22. Iglesias MT, Lorenzo CD, Polo MDC, Martin-Alvarez PJ, Pueyo E (2004) *J Agric Food Chem* 52:84–89
23. Anklam EA (1998) *Food Chem* 63:549–562
24. von der Ohe W, Dustmann JH, von der Ohe K (1991) *Dtsch Lebensm Rundsch* 87:383–386
25. Cometto PM, Faye PF, Naranjo RD, Rubio MA, Aldao MAJ (2003) *J Agric Food Chem* 51:5079–5087
26. Hermosin I, Chicon RM, Cabezudo MD (2003) *Food Chem* 83:263–268
27. Nozal MJ, Bernal JL, Toribio ML, Diego JC, Ruiz A (2004) *J Chromatogr A* 1047:137–146
28. Cotte JF, Casabianca H, Giroud B, Albert M, Lheritier J, Grenier-Loustalot MF (2004) *Anal Bioanal Chem* 378:1342–1350
29. Berlitz HD, Grosch W, Schieberle P (2009) *Food chemistry*, 4th edn. Springer, Leipzig
30. Baroni MV, Chiabrando GA, Costa C, Wunderlin DA (2002) *J Agric Food Chem* 50:1362–1367
31. Peiren N, de Graaf DC, Vanrobaeys F, Danneels EL, Devreese B, Beeumen JV, Jacobs FJ (2008) *Toxicon* 52:72–83
32. Babacan S, Rand AG (2005) *J Food Sci* 70:C413–C418
33. Vit P, Pulcini P (1996) *J Apicult Res* 35:57–62
34. Pontoh J, Low NH (2002) *Insect Biochem Mol Biol* 32:679–690
35. White JW, Kushnir I (1967) *J Apicult Res* 6:69–89
36. Schepartz AI, Subers MH (1964) *Biochim Biophys Acta* 85:228–237
37. White JW, Kushnir I (1966) *Anal Biochem* 16:302–313
38. Jintanart W, Takeshi Y, Hiroyuki N, Kim YM, Natsuko S, Mamoru N, Okuyama M, Mori H, Saji O, Chanchao C, Wongsiri S, Surarit R, Svasti J, Chiba S, Kimura A (2006) *Biosci Biotechnol Biochem* 70:2889–2898
39. Huber RE, Mathison RD (1976) *Can J Biochem* 54(1976):153–164
40. Cho NC (1994) *Korean Biochem J* 27:509–513
41. Ohashi K, Natori S, Kubo T (1999) *Eur J Biochem* 265:127–133
42. Sporns P (1992) In: Hui YH (ed) *Encyclopedia of food science and technology*, vol 2. Wiley, New York
43. White JW (1978) *Adv Food Res* 24:287–364
44. White JW (1975) In: Crane E (ed) *Honey: a comprehensive study*. Heinemann, London
45. Rinaudo MT, Ponzetto C, Vidano C, Marletto F (1973) *Comp Biochem Physiol* 46B:253–256
46. Fujita T, Kozuka-Hata H, Uno Y, Nishikori K, Morioka M, Oyama M, Kubo T (2010) *Biochem Biophys Res Comm* 397:740–744
47. Won SR, Li CY, Kim JW, Rhee HI (2009) *Food Chem* 113:1334–1338
48. Callesen AK, Madsen JS, Vach W, Kruse TA, Mogensen O, Jensen ON (2009) *Proteomics* 9:1428–1441
49. Reiz B, Li L (2010) *J Am Soc Mass Spectrom* 21:1596–1605
50. Lill JR, Ingle ES, Liu PS, Pham V, Sandoval WN (2007) *Mass Spectrom Rev* 26:657–671
51. Afiuni-Zadeh S, Guo X, Gholamhassan A, Lankmayr E (2011) *Talanta* 85:1835–1841
52. Xue YJ, Liu J, Pursley J, Unger S (2006) *J Chromatogr B* 831:213–222
53. Mitchell TJ, Irvine L, Scoular RHM (1955) *Analyst* 80:620–622
54. Nelson JM, Cohn DJ (1924) *J Biol Chem* 61:193–224
55. Papadakis PE (1929) *J Biol Chem* 83:561–568
56. Padovan GJ, Rodrigues LP, Leme IA, Jong DD, Marchini JS (2007) *Eurasian J Anal Chem* 2:134–141
57. AOAC (2010) Official method 998.12. Official methods of analysis. Association of Official Analytical Chemists, Washington DC
58. Polson C, Sarkar P, Incledon B, Raguvaran V, Grant R (2003) *J Chromatogr B* 785:263–275
59. Kong FJ, Oyanagi A, Komatsu S (2010) *Biochim Biophys Acta* 1804:124–136
60. Li J, Fang Y, Zhang L, Begna D (2011) *J Insect Physiol* 57:372–384
61. Steinhorn G, Sims IM, Carnachan SM, Carr AJ, Schlothauer R (2011) *Food Chem* 128:949–956
62. Kubota M, Tsuji M, Nishimoto M, Wongchawalit J, Okuyama M, Mori H, Matsui H, Surarit R, Svasti J, Kimura A, Chiba S (2004) *Biosci Biotechnol Biochem* 68:2346–2352
63. Sheoran IS, Ross ARS, Olson DJH, Sawhney VK (2009) *Plant Sci* 176:99–104
64. Carpentier SC, Witters E, Laukens K, Deckers P, Swennen R, Panis B (2005) *Proteomics* 5:2497–2507
65. Jermyn MA, Yeow YM (1975) *Aust J Plant Physiol* 2:501–531
66. van Holst GJ, Clarke AE (1985) *Anal Biochem* 148:446–450
67. Gane AM, Craik D, Munro SLA, Hoelett GJ, Clarke AE, Bacic A (1995) *Carbohydr Res* 277:67–85
68. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) *J Biol Chem* 193:265–275
69. Herriott RM (1941) *Proc Soc Exp Biol Med* 46:642–644
70. Wu H (1920) *J Biol Chem* 43:189–220
71. da C Azeredo L, Azeredo MAA, de Souza SR, Dutra VML (2003) *Food Chem* 80:249–254
72. Neurath AR (1966) *Experientia* 22:290
73. Kuno H, Kihara HK (1967) *Nature* 215:974–975
74. Vallejo CG, Lagunas R (1970) *Anal Biochem* 138:141–143
75. Lo C, Stelson H (1972) *Anal Biochem* 45:331–336
76. Wessel D, Flugge UI (1984) *Anal Biochem* 138:141–143
77. Makkar HPS, Sharma OP, Negi SS (1980) *Anal Biochem* 104:124–126
78. Horikawa S, Ogawara H (1979) *Anal Biochem* 97:116–119
79. Rauh M (2012) *J Chromatogr B* 883–884:59–67
80. Sundqvist G, Stenvall M, Berglund H, Ottosson J, Brumer H (2007) *J Chromatogr B* 852:188–194
81. Gharahdaghi F, Weinberg CR, Meagher DA, Imai BS, Mische SM (1999) *Electrophoresis* 20:601–605
82. Matros A, Kaspar S, Witzel K, Mock HP (2011) *Phytochem* 72:963–974
83. Yeung YG, Nieves E, Angeletti RH, Stanley ER (2008) *Anal Biochem* 382:135–137
84. Norris JL, Porter NA, Caprioli RM (2003) *Anal Biochem* 75:6642–6647
85. Chen EI, McClatchy D, Park SK, Yates JR III (2008) *Anal Biochem* 80:8694–8701
86. Perkins PD, Fischer SM (2008) Peptide derivatization method to increase fragmentation information from MS/MS spectra. US Patent Application US 2008/0026479
87. Delahunty C, Yates JR III (2005) *Methods* 35:248–255
88. Arsene CG, Ohlendorf R, Burkitt W, Pritchard C, Henrion A, O'Connor G, Bunk DM, Guttler B (2008) *Anal Biochem* 80:4154–4160
89. Agger SA, Marney LC, Hoofnagle AN (2010) *Clin Chem* 56:1804–1813

90. Proc JL, Kuzyk MA, Hardie DB, Yang J, Smith DS, Jackson AM, Parker CE, Borchers CH (2010) *J Proteome Res* 9:5422–5437
91. Matysiak J, Schmelzer CEH, Neubert RHH, Kokot ZJ (2011) *J Pharm Biomed Anal* 54:273–278
92. Ren D, Ratnaswamy G, Beierle J, Treuheit MJ, Brems DN, Bondarenko PV (2009) *Int J Biol Macromol* 44:81–85
93. Peiren N, Vanrobaeys F, de Graaf DC, Devreese B, Beeumen JV, Jacobs FJ (2005) *Biochim Biophys Acta* 1752:1–5
94. Fang CY, Chen HY, Wang M, Chen PL, Chang CF, Chen LS, Shen CH, Ou WC, Tsai MD, Hsu PH, Chang D (2010) *Virol* 402:164–176
95. Johnson D, Orlando R (2011) *J Biomol Tech* 22(Suppl):S57–S58
96. Wu R, Hu L, Wang F, Ye M, Zou H (2008) *J Chromatogr* 1184:369–392
97. Tanaka N, Kimura H, Tokuda D, Hosoya K, Ikegami T, Ishizuka N, Minakuchi H, Nakanishi K, Shintani Y, Furuno M, Cabrera K (2004) *Anal Biochem* 336:1273–1281
98. Washburn MP, Ulaszek R, Deciu C, Schieltz DM, Yates JR (2002) *Anal Chem* 74:1650–1657
99. Wolters DA, Washburn MP, Yates JR (2001) *Anal Chem* 73:5683–5690
100. Link AJ, Eng J, Schiartz DM, Carmack E, Mize GJ, Morris DR, Garvick BM, Yates JR III (1999) *Nat Biotechnol* 17:676–682
101. Gilar M, Olivova P, Daly AE, Gebler JC (2005) *J Sep Sci* 28:1694–1703
102. Scevchenko A (1996) *Proc Natl Acad Sci* 93:14440–14445
103. Alomirah HF, Alli I, Konishi Y (2000) *J Chromatogr A* 893:1–21
104. Baggerman G, Verleyen P, Clynen E, Huybrechts J, De Loof A, Schoofs L (2004) *J Chromatogr B* 803:3–16
105. John H, Walden M, Schafer S, Genz S, Forssmann WG (2004) *Anal Bioanal Chem* 378:883–897
106. Bhushan D, Pandey A, Choudhary MK, Datta A, Chakraborty S, Chakraborty N (2007) *Mol Cell Proteomics* 6:1868–1884
107. Bachi A, Bonaldi T (2008) *J Proteomics* 71:357–367
108. Wilm M (2009) *Proteomics* 9:4590–4605
109. Schmelzer CE, Schops R, Ulbrich-Hofmann R, Neubert RH, Raith K (2004) *J Chromatogr A* 1055:87–92
110. Tang K, Allman SL, Jones RB, Chen CH (1993) *Anal Chem* 65:2164–2166
111. Gonnet F, Lemaitre G, Waksman G, Tortajada J (2003) *Proteome Sci* 1:1–7
112. Karas M, Kruger R (2003) *Chem Rev* 103:427–440
113. Francese S, Lambardi D, Mastrobuoni G, la Marca G, Moneti G, Turillazzi S (2009) *J Am Soc Mass Spectrom* 20:112–123
114. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, Spiegelman CH, Zimmerman LJ, Ham AJ, Keshishian H, Hall SC, Allen S, Blackman RK, Borchers CH, Buck C, Cardasis HL, Cusack MP, Dodder NG, Gibson BW, Held JM, Hiltke T, Jackson A, Johansen EB, Kinsinger CR, Li J, Mesri M, Neubert TA, Niles RK, Pulsipher TC, Ransohoff D, Rodriguez H, Rudnick PA, Smith D, Tabb DL, Tegeler TJ, Variyath AM, Vega-Montoto JL, Wahlander A, Waldemarson S, Wang M, Whiteaker JR, Zhao L, Anderson NL, Fisher SJ, Liebler DC, Paulovich AG, Regnier FE, Tempst P, Carr SA (2009) *Nat Biotechnol* 27:633–641
115. Blonder J, Conrads TP, Veenstra TD (2004) *Expert Rev Proteomics* 1:153–163
116. Blonder J, Chan KC, Issaq HJ, Veenstra TD (2006) *Nat Protoc* 1:2784–2790
117. Fortin T, Salvador A, Charrier JP, Lenz C, Bettsworth F, Lacoux X, Choquet-Kastylevsky G, Lemoine J (2009) *Anal Chem* 81:9343–9352
118. Keller A, Nesvizhskii AI, Kolker E, Abersold R (2002) *Anal Chem* 74:5383–5392
119. Nesvizhskii AI, Keller A, Kolker E, Abersold R (2003) *Anal Chem* 75:4646–4658
120. Cottrell JS (2011) *J Proteomics* 74:1842–1851