

Contents lists available at [ScienceDirect](http://ScienceDirect)

# Chemometrics and Intelligent Laboratory Systems

journal homepage: [www.elsevier.com/locate/chemometrics](http://www.elsevier.com/locate/chemometrics)

## Improved gel-enhanced liquid chromatography-mass spectrometry by chemometrics for halal proteomics



Mohd Hafis Yuswan<sup>a</sup>, Wan Mohd Aizat<sup>f</sup>, Mohd Nasir Mohd Desa<sup>e</sup>, Amalia Mohd Hashim<sup>c</sup>, Nur Ajreena Rahim<sup>a</sup>, Shuhaimi Mustafa<sup>c</sup>, Rozi Mohamed<sup>d</sup>, Dhilia Udie Lamasudin<sup>a,b,\*</sup>

<sup>a</sup> Laboratory of Halal Services, Halal Products Research Institute, Universiti Putra Malaysia, 43000, UPM Serdang, Selangor, Malaysia

<sup>b</sup> Department of Cell and Molecular Biology, Faculty of Biotechnology and Biomolecular Science, Universiti Putra Malaysia, 43000, UPM Serdang, Selangor, Malaysia

<sup>c</sup> Department of Microbiology, Faculty of Biotechnology and Biomolecular Science, Universiti Putra Malaysia, 43000, UPM Serdang, Selangor, Malaysia

<sup>d</sup> Forest Biotech Laboratory, Department of Forest Management, Faculty of Forestry, Universiti Putra Malaysia, 43000, UPM Serdang, Selangor, Malaysia

<sup>e</sup> Department of Biomedical Science, Faculty of Medicine and Health Science, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor, Malaysia

<sup>f</sup> Institute of Systems Biology (INBIOSIS), Universiti Kebangsaan Malaysia, 43600, UKM Bangi, Selangor, Malaysia

### ARTICLE INFO

#### Keywords:

Chemometrics  
GeLCMS  
Bottom-up proteomics  
Protein marker  
PCA

### ABSTRACT

Numerous analytical methods for the authentication of halal meat are now well established, with gel-enhanced liquid chromatography-mass spectrometry (GeLCMS) being a popular approach. However, the selection of potential protein markers on 1-dimensional gel electrophoresis (1DE) prior to LCMS is considered problematic, because using the optical density for the selection process could introduce human error. In this study, an improved GeLCMS method assisted by multivariate principal component analysis (PCA) was developed to identify the potential protein markers for non-halal pork among halal beef and chicken. The improved GeLCMS technique allowed for the confident excising of identified protein bands prior to in-gel tryptic digestion. The inferential protein markers (myofibrillar proteins), which might be present in the samples, were determined based on the identified sequence of peptides. This chemometric-assisted GeLCMS could potentially be used as a guideline to assist chemists in analysis of any gel-based separation of biomolecules, regardless of the field of study.

### 1. Introduction

Halal food is very important for Muslims, because it is clearly mentioned in the Holy Quran [1]. Halal meat refers to animal meat that fulfils the requirements for slaughter according to Islamic law, which specifically excludes pork [2]. However, halal meat products are commonly adulterated with pork because it is cheaper than beef and chicken [3]. From the non-Muslim's perspective, meat adulterated with pork results in a loss of quality, owing to unethical manufacturers trying to increase the quantity. Furthermore, a previous study [4] reported that pork is a carrier of various foodborne zoonoses, including protozoa (*Toxoplasma gondii* and *Sarcocystis* spp.) and helminths (*Trichinella* spp. and *Taenia* spp.).

Owing to the issues of meat adulteration, numerous analytical methods for the authentication of meat have recently been developed using either genomics [5–9] or proteomics [10–16] approaches. The application of proteomics for the authentication of meat can complement

genomics, because it provides an alternative analytical method. Moreover, because proteomics is based on the primary structure of peptides, which remains intact through chemical or mechanical processes, it has an advantage over genomics. In proteomics, proteins extracted from a biological sample are manipulated by a series of procedures, such as proteolytic cleavage, reduction of disulfide bonds, and alkylation of the thiol group of cysteine residues, to generate free peptides. Subsequently, the peptides are subjected to mass spectrometry (MS) analysis, with post-analysis using well-developed web-based search engines. These search engines, which are available either freely or as commercial packages, help identify candidate peptides and subsequently deduce the proteins that might be present in the samples [17]. This identification is based on the molecular weight search (MOWSE) score, for which a high score indicates a high confidence level [18].

Prior to tandem MS analysis, the complex protein samples must be fractionated by either chromatography or electrophoresis, followed by proteolytic digestion, to reduce sample complexity and increase the

\* Corresponding author. Laboratory of Halal Science Research, Halal Products Research Institute, Universiti Putra Malaysia, 43000, UPM Serdang, Selangor, Malaysia.

E-mail address: [dhilia@upm.edu.my](mailto:dhilia@upm.edu.my) (D.U. Lamasudin).

<https://doi.org/10.1016/j.chemolab.2019.103825>

Received 21 September 2018; Received in revised form 25 May 2019; Accepted 12 August 2019

Available online 16 August 2019

0169-7439/© 2019 Elsevier B.V. All rights reserved.

dynamic range for peptide identification [19]. This procedure is also known as bottom-up proteomics [20]. If the complex protein samples are separated according to their molecular weight through gel electrophoresis, the technique is known as one-dimensional gel electrophoresis (1DE). Two-dimensional gel electrophoresis (2DE) includes an additional step, wherein the proteins are first separated by isoelectric focusing through an immobilized pH gradient strip.

The gel electrophoresis technique for the separation of proteins, using Coomassie- and silver-stained gels with an in-gel digestion protocol, was first demonstrated in 1996 by Shevchenko, Mann, and co-workers to enable protein identification by peptide mass fingerprinting using a mass spectrometer [21]. Later, in 2011, an advanced protocol known as gel-enhanced liquid chromatography-mass spectrometry (GeLCMS) was introduced by Lundby and Olsen. The protocol describes the sample preparation in detail as well as a downstream analysis using LCMS, after pre-separation by 1DE [19]. Nevertheless, 1DE remains a favorable method for protein separation for most chemists because the gel acts as an ideal container for handling, concentrating, and storing protein. Furthermore, the gel is efficient in filtering out low-molecular-weight impurities such as detergents and buffer components.

However, the selection of protein bands from 1DE relies solely on the measurement of relative protein abundance based on optical density prior to tandem MS analysis [19,21–23]. This conventional technique for excising protein bands is highly subjected to human error, resulting in inconsistent results when performed by different chemists. Therefore, this study presents a new, improved GeLCMS protocol that implements chemometrics, through multivariate principal component analysis (PCA), for selecting the protein bands from the 1DE. The PCA interprets the complexity of optical density data, originating from all protein bands on the 1DE, through dimensional reduction. The PCA model also retains maximum data variability, thus enabling the discrimination of latent variables (protein bands) in the samples.

Previous studies have shown that PCA can differentiate porcine gelatin from that of other sources, but the analysis was not extended to tandem MS, thus limiting the identification to the molecular weight level [24,25]. However, a study has been conducted on species identification of a moray eel meat in which the identified protein markers solely relied on the 1DE profile, but the proteins' identity was uncertain [22]. Therefore, this study provides a set of instructions to confidently identify proteins from gel pieces prior to LCMS analysis; some of the procedures are previously known to experienced researchers, while others are relatively novel. By harnessing the potential of PCA, in-depth identification of potential protein markers for pork (*Sus scrofa*) meat is realistic and does not require difficult protocols for the excising of protein bands.

## 2. Materials and methods

### 2.1. Materials

Analytical grade urea and thiourea were purchased from Friendemann Schmidt, Australia. Electrophoresis grade EDTA, sodium dodecyl sulfate (SDS), iodoacetamide (IAA), and dithiothreitol (DTT) were purchased from Bio-Rad, Shanghai, China. Ultrapure grade Tris-HCl, urea, and thiourea were purchased from 1st BASE, Shanghai, China. Mass spectrometry (MS) grade acetonitrile (ACN) and formic acid (FA) were purchased from Fisher Chemical, Fair Lawn, NJ, USA. Sequencing grade endoproteinase trypsin was obtained from Promega, Madison, WI, USA. A standard chemical kit with a low/high concentration of polypropylene glycol (PPG) for calibration of AB SCIEX instruments was purchased from AB SCIEX Pte. Ltd., Framingham, MA, USA. Water was freshly prepared from an ultrapure water system (Arium®611VF, Sartorius Stedim, Goettingen, Germany). Three types of raw meat samples, divided into halal and non-halal groups, were purchased randomly from a local market in the nearby area of Serdang Selangor, Malaysia. The halal certification was obtained from the Department of Islamic Development Malaysia (JAKIM). Raw beef ( $n = 5$ ) and chicken ( $n = 5$ ) were grouped as

halal meat, whereas pork ( $n = 5$ ) was grouped as non-halal meat. All the raw meats were stored at  $-80\text{ }^{\circ}\text{C}$  until further analysis.

### 2.2. Sample extraction

The raw meats were diced to approximately  $2 \times 2 \times 2$ -cm cubes and then washed according to a protocol described in a previous study, prior to protein extraction [26]. Subsequently, approximately 20–30 g of washed meat samples was transferred into a 50-mL Falcon tube for lyophilization (FreeZone 1 L benchtop freeze dry system, LabConco, Kansas City, MO, USA). Then, each lyophilized raw meat sample was ground into a fine powder and pooled according to its species for randomization prior to protein extraction, with three replicates each. The extraction procedure was based on that described in a previous study, with some modification [19]. Briefly, for each fine powder sample, an extraction buffer (6 M urea, 1 M thiourea, 0.1 mM EDTA, 1 mM DTT, and 50 mM Tris-HCl) was added into the 50-mL Falcon tube at a 1:1 ratio (w/v) and gently mixed. The mixture was incubated at  $25\text{ }^{\circ}\text{C}$  for 1 h prior to centrifugation at 12,000 rcf at  $4\text{ }^{\circ}\text{C}$  for 60 min. The supernatant was aliquoted into a 2-mL Eppendorf tube and stored at  $-80\text{ }^{\circ}\text{C}$  until further analysis.

### 2.3. 1DE protein separation

Extracted protein separation was performed according to the method described in a previous study, with some modifications [19]. Briefly, approximately 10 mg of each sample was mixed with 10 mM DTT-SDS buffer (62 mM Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, 5% DTT, and 0.00125% bromophenol blue in ultrapure water) at a 1:1 ratio (v/v) and incubated at  $70\text{ }^{\circ}\text{C}$  for 5 min prior to loading into the well. The samples were separated on a gel consisting of 8% and 4% resolving and stacking acrylamide gel, respectively. The separation was carried out in a Mini-Protean II Tetra Cell (Bio-Rad Laboratories, Hercules, CA, USA) at a constant voltage of 120 V for 90 min. Bio-Rad Precision Plus Protein All Blue Prestained Protein Standards were used as protein marker, ranging from 37 to 250 kDa to estimate the molecular weight of the protein bands for each sample. Subsequently, the gel was stained with Coomassie blue staining solution for 24 h, destained prior to the scan, and visualized using a densitometer (GS-800 Calibrated Densitometer, Bio-Rad). The data were analyzed using Quantity One® software (Bio-Rad) to determine the molecular weight and optical density of the polypeptide bands. All samples were run in triplicate, and data from the 1DE were subjected to statistical and chemometric analyses.

### 2.4. In-gel tryptic digestion and desalting

The in-gel tryptic digestion and desalting procedures were performed as described in a previous study, with some modification [19]. Bands for the proteins of interest, chosen based on the chemometric analysis, were excised and transferred into a separate 2-mL Eppendorf tube. Then, 500  $\mu\text{L}$  each of 0.05 M ammonium bicarbonate ( $\text{NH}_4\text{HCO}_3$ ) and 50% ACN were added to wash the gels, and the supernatant was discarded after being centrifuged at 1200 rpm for 20 min. The sample was dehydrated by incubating with 500  $\mu\text{L}$  of 100% ACN for 15 min. The supernatant was discarded prior to reducing the proteins by adding 100  $\mu\text{L}$  each of 0.01 M DTT and 0.1 M  $\text{NH}_4\text{HCO}_3$  for 30 min at  $60\text{ }^{\circ}\text{C}$ . The supernatant was separated by centrifugation at 1200 rpm for 45 min and discarded. Subsequently, protein alkylation was performed by adding 100  $\mu\text{L}$  each of 0.05 M IAA and 0.1 M  $\text{NH}_4\text{HCO}_3$  and incubating at  $27\text{ }^{\circ}\text{C}$  for 30 min, followed by removing the supernatant after centrifuging at 1200 rpm for 45 min. The digested proteins were washed by adding 100  $\mu\text{L}$  each of 0.1 M  $\text{NH}_4\text{HCO}_3$  and 50% ACN and then removing the supernatant by centrifuging at 1200 rpm for 10 min. The washing step was repeated twice. To digest the protein, 5  $\mu\text{L}$  of 1 mg/mL trypsin was added and incubated overnight. Next, 100  $\mu\text{L}$  each of 0.05 M  $\text{NH}_4\text{HCO}_3$  and 100% ACN were added prior to desalting. The digested protein was desalted using Agilent Bond Elut OMIX C18 tip (Waldbronn, Germany).

## 2.5. Data-dependent acquisition of LCMS

The digested protein sample was separated using high-performance liquid chromatography (Agilent 1200 Series), consisting of an autosampler (G1367D), a binary pump (G1312B), and a column oven (G1316A), which was coupled to AB SCIEX 4000 QTrap MS (Singapore). The sample was injected into a Phenomenex Kinetex Core-shell C18 reversed phase column (Torrance, CA;  $100 \times 2.1$  mm,  $2.6 \mu\text{m}$ ,  $100 \text{ \AA}$ ) at  $45^\circ\text{C}$ , with the flow rate of  $300 \mu\text{L}/\text{min}$  for 25 min. Mobile phases A and B were 0.1% FA and 100% ACN, respectively. The column was equilibrated for  $10 \times$  column volume prior to sample injection at a volume of  $15 \mu\text{L}$ . The concentration of mobile phase A was maintained at 97% for 1 min, gradually dropped to 3% after 19 min, and maintained at 3% until 22 min. Finally, at 22.1 min, the concentration of mobile phase A was brought back to its initial concentration of 97% and maintained at this concentration until 25 min. Prior to MS analysis, a Standard Chemical Kit (AB SCIEX) containing polypropylene glycol at  $2 \times 10^{-6}$  M was used to calibrate the Q1 and Q3 of the MS. The positive mode of Turbo Spray (ESI) ionization was used at  $400^\circ\text{C}$ , 5200 V of nebulizer voltage, and 40 psi of pressure. Curtain gas was maintained at 30 psi, whereas the pressures of source gases 1 and 2 were 40 and 30 psi, respectively. The declustering potential (DP) was at 50 V to minimize in-source fragmentation. Data-dependent acquisition (DDA), consisting of enhanced MS (EMS), enhanced resolution (ER), and enhanced product ion (EPI), was used for the analysis of all samples. For EMS, the mass range of the instrument was  $m/z$  400–2000 at 4000 Da/s scan speed to obtain the highest sensitivity. To calculate the charge and exact mass of the peptides, the ER scan mode was carried out at 250 Da/s scan speed for best resolution. For tandem MS, the EPI scan mode was used with a scan range of  $m/z$  100–2000, divided into two segments ( $m/z$  100–280 and  $m/z$  275–2000), at 4000 Da/s scan speed. Collision-induced dissociation (CID) analysis was performed at a collision energy (CE) of 35 V, with CE spread (CES) of 15 V. The selection of ions for tandem MS was limited to the 1–4 most intense peaks with charge states of 1+, 2+, 3+, and 4+. Instrument control, data acquisition, and data processing were performed using Analyst 1.5.1 software. For protein identification, the peak lists were searched via MS/MS ion search using MS-TaQ (ProteinProspector 5.20.0, University of California, San Francisco, USA) against a decoy, randomized, and concatenated SwissProt database. The database parameters were as follows: enzyme, trypsin; taxonomy, *Sus scrofa*; miss cleavage, one; non-specific cleavage, at 0 terminus; parent tolerance, 1 Da; fragment tolerance, 1.5 Da; precursor charge, automatic; modification, carbamidomethyl and oxidation of methionine.

## 2.6. Statistical and multivariate analyses

Results from the optical density on 1DE for each sample were subjected to analysis of variance (ANOVA;  $H_0$ : all the samples have the same mean value;  $H_a$ : at least one of the samples has a different mean value). Subsequently, a pairwise multiple comparison (post hoc) was performed using Tukey's honest significant difference (HSD) to determine the significant difference between each sample, if  $H_a$  was accepted. This statistical analysis was performed using SPSS 15.0 software, and the mean difference for each sample was considered significant if  $p$ -value  $< 0.05$ . Results from qualitative and quantitative comparisons of 1DE were subjected to PCA by setting the molecular weight based on the relative front of each sample as variable (K), and each sample, including replicates, was designated as an observation (N) to construct a data matrix X. The chemometric analysis was performed using Umetrics SIMCA 14.1 software (Umea, Sweden).

## 3. Results and discussion

### 3.1. Qualitative and quantitative electrophoretic 1DE profile

Details of the protein bands and their intensities are shown in Fig. 1.

In this study, beef and chicken displayed four and eight protein bands, respectively. The four protein bands detected in beef were 38, 43, 48, and 59 kDa, all of which were also found in chicken. The four additional protein bands found only in chicken were 62, 131, 142, and 151 kDa. Five different protein bands were detected only in pork: 35, 40, 42, 45, and 49 kDa. These findings were supported by a previous study, in which the extracted proteins for cooked beef, chicken, and pork, as well as lamb, horse, deer, turkey, and duck meats, were in the range of 22–200 kDa [27]. Another study on undigested cooked beef, chicken, pork, and fish also revealed similar findings, wherein the molecular weights of proteins were in the range of 11–223 kDa [28]. According to Xia and coworkers [29], myofibrillar protein such as myosin has prominent protein bands close to 205 kDa. Other myofibrillar proteins such as actin, troponin, and tropomyosin have prominent protein bands at 40–45 kDa, as does light chain myosin at 25 kDa [30]. Other prominent protein bands that have been reported are sarcoplasmic proteins, with a wide range of molecular weights between 19 and 110 kDa; these proteins include enzymes such as M-type creatine-kinase; mitochondrial malate dehydrogenase; l-lactate dehydrogenase A chain isoform 1; and the glycolytic enzymes fructose-bisphosphate aldolase A isoform 1, glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase 1, phosphoglycerate mutase 2, and beta-enolase [15,29]. Based on the 1DE results, the extracted proteins for all samples were most likely actin, troponin, tropomyosin, myoglobin, hemoglobin, cytochrome, and other enzymes that should be confirmed by tandem MS analysis.

Furthermore, the differences in the optical density were significant for certain protein bands (Table 1). This result indicated the presence or absence of that particular protein, which might be used for meat speciation. This technique has previously been demonstrated on mammalian and fish meats for species identification [23]. As shown in Table 1, only the 48-kDa protein band showed significant differences between meat samples, at  $p$ -value  $< 0.05$ . This 48-kDa protein was highly expressed in chicken, but the level was not significantly different from that in pork. This result contrasted with a previous study [28], in which the 48- and 42-kDa proteins were highly expressed in an undigested cooked pork sample and were identified as a beta-enolase and alpha cardiac-actin, respectively, by MALDI-TOF-MS. However, two protein bands had only been detected in the pork sample, namely, 34- and 44-kDa proteins that represent myoglobin and tropomyosin, respectively [31,32]. These

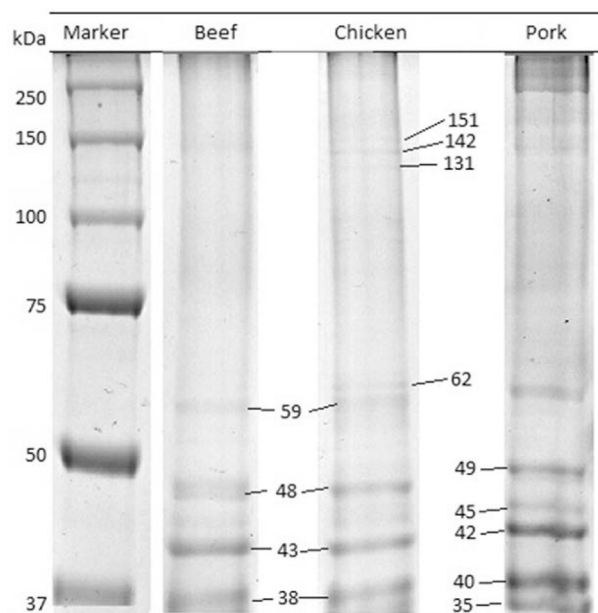


Fig. 1. Electrophoretic 1DE profiles of proteins extracted from beef, chicken, and pork. Each gel was loaded with  $20 \mu\text{g}$  of proteins, based on the total protein concentration determined by the Bradford assay.

**Table 1**  
Optical density of protein bands on 1DE for beef, chicken, and pork.

MW approx. (kDa) <sup>b</sup>	Sample type <sup>a</sup>		
	Beef	Chicken	Pork
34	ND	ND	0.179 ± 0.001
38	0.191 ± 0.005 <sup>a</sup>	0.205 ± 0.025 <sup>a</sup>	0.226 ± 0.001 <sup>a</sup>
42	0.204 ± 0.011 <sup>a</sup>	0.202 ± 0.015 <sup>a</sup>	0.219 ± 0.013 <sup>a</sup>
44	ND	ND	0.144 ± 0.006
48	0.156 ± 0.002 <sup>a</sup>	0.178 ± 0.007 <sup>b</sup>	0.166 ± 0.003 <sup>ab</sup>
59	0.145 ± 0.003 <sup>a</sup>	0.154 ± 0.004 <sup>a</sup>	0.146 ± 0.001 <sup>a</sup>
62	ND	0.147 ± 0.003	ND
131	ND	0.144 ± 0.002	ND
143	ND	0.149 ± 0.003	ND
150	ND	0.144 ± 0.003	ND

ND stands for not detected.

<sup>a</sup> Values are the means from replicates, with the corresponding standard deviation.

<sup>b</sup> Values are the approximate molecular weight (MW) of proteins on 1DE, based on the bands calibrated to standard protein markers. Means with different superscripts on the same row are significantly different at p-value < 0.05.

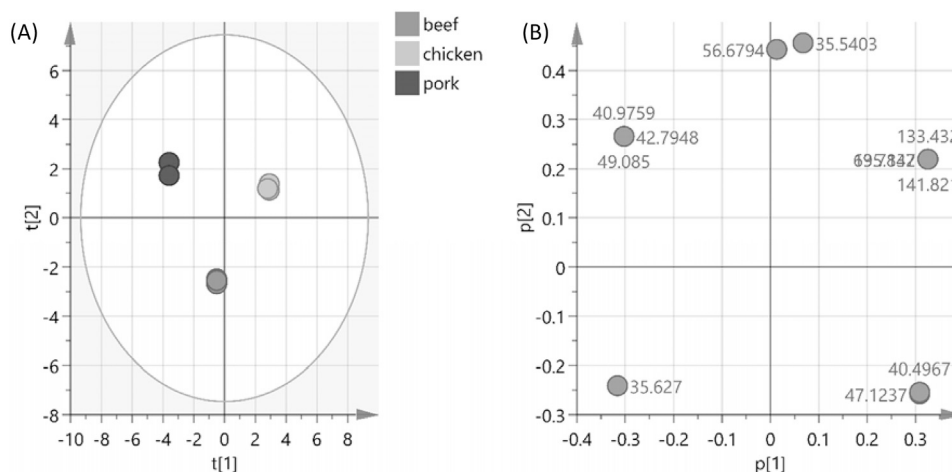
protein bands (34, 42, 44, and 48 kDa) indicated the possibility of potential protein markers for the pork samples. However, qualitative and quantitative analyses of the 1DE profile need to be confirmed by multivariate PCA (instead of univariate analysis) to obtain an overview of the pattern, grouping, similarities, differences, and outliers of the samples prior to protein marker selection.

### 3.2. Advanced chemometric analysis

Chemometric analysis is a technique that involves a statistical approach to understand the chemical information generated by an analytical instrument. In this study, the relative front of the optical density for each protein band from the 1DE results was analyzed by multivariate PCA to understand the relationship between meat samples

and protein bands (variables), as shown in Fig. 2. The PCA model commonly consists of a score plot (Fig. 2a), showing the sample locations in each principal component (PC), and a loading plot (Fig. 2b), interpreting the relationship between the protein bands. This model reveals the pattern, grouping, and outlier of the PCA model [33]. Samples close to each other on the plot indicate similar properties, whereas those far from each other are dissimilar with respect to the molecular weight of the protein bands [34]. The PCA detected two PCs (Fig. 2a): PC1 and PC2 accounted for 62% and 35%, respectively, thus making 97% for the total variance accounted for in the PCA model. No outlier was observed by ellipse Hotelling's  $T^2$ . According to Fig. 2a, all the meat samples were obviously different from each other. However, this study is in contrast with that by Montowska and coworkers [35], where the cooked beef, pork, and horse meat were grouped together but differed from chicken and turkey. This might be due to the differences in the meat conditions and types.

Fig. 2b displays the loading plot that describes the relationship between all 12 protein bands, which correlate to the raw meats as tabulated in the score plot. Protein bands with similar information (correlated) are clustered together, and the further away they are from the origin, the stronger their contribution to the PCA model is [34]. The 69.71-, 133.43-, 135.84-, and 141.82-kDa proteins located in the upper right-hand corner of Fig. 2b suggested that the proteins were highly expressed in raw chicken. In contrast, only the 35.63-kDa protein was expressed in raw beef. Moreover, the PCA model suggested that raw pork contained proteins of 40.98, 42.79, and 49.09 kDa. Results from this multivariate PCA were different from those from the initial univariate ANOVA (Table 1) for the potential protein markers of pork. This was because the 34- and 44-kDa proteins were not included in the loading plot, with the 40-kDa protein acting as a substitution. This might be due to the characteristics of semi-quantitative 1DE, wherein the molecular weight of each protein band is relative to an estimation of protein standards, prior to univariate ANOVA in which only single variables are assessed per analysis. On the other hand, the multivariate PCA was highly suited for the



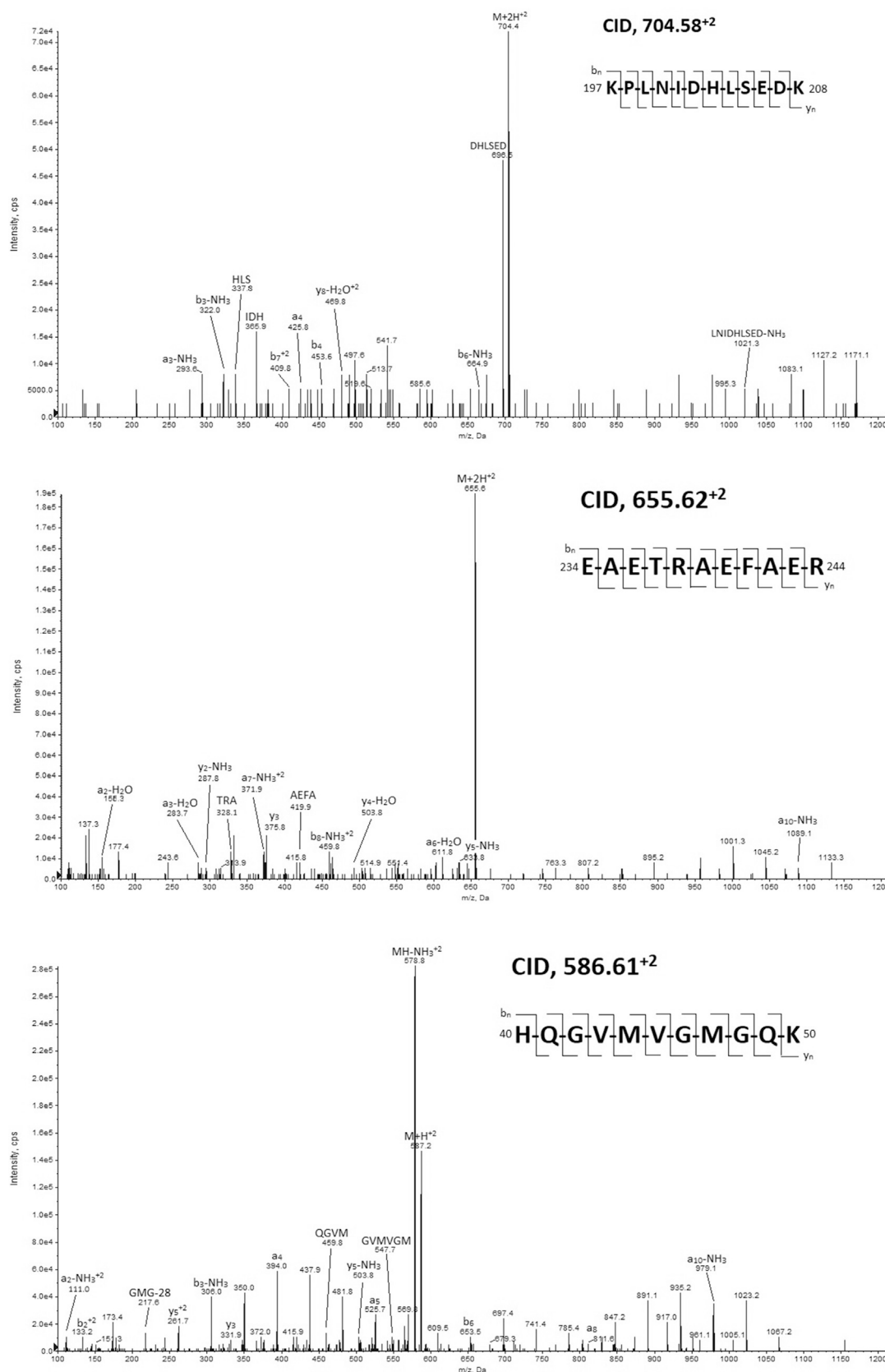
**Fig. 2.** PCA grouping of raw meat samples. (A) Score plot. (B) Loading plot.

**Table 2**  
Details of the potential protein markers for pork meat samples from tandem LCMS analysis.

MW (kDa)	Observed $m/z$	Theoretical $m/z$	Error (Da) <sup>a</sup>	Peptide sequence	Protein name	Matched <sup>b</sup>	Accession no.
40	704.40 <sup>2+</sup>	704.37 <sup>2+</sup>	0.03	(R)KPLNIDHLSSEDK(L)	Troponin T	13/26 (50.0)	Q75NG9
	655.62 <sup>2+</sup>	654.31 <sup>2+</sup>	1.31	(K)EAETRAEFAER(S)	Tropomyosin alpha-1 chain	19/40 (47.5)	P42639
	586.61 <sup>2+</sup>	585.79 <sup>2+</sup>	0.82	(R)HQGVVMVGMGQK(D)	Actin cytoplasmic 1	24/66 (36.4)	Q6QAQ1
42	411.02 <sup>2+</sup>	410.73 <sup>2+</sup>	0.29	(R)VLDYRR(K)	COP9 signalosome complex subunit 4	3/5 (60.0)	A7Y521
	611.47 <sup>3+</sup>	610.31 <sup>3+</sup>	1.16	(R)VLGQGLADSACQLETLR(L)	Ribonuclease inhibitor	5/5 (100.0)	P10775

<sup>a</sup> Deviation of peptide mass between observed and theoretical mass, based on the MS-Taq database.

<sup>b</sup> Number of matched fragment ions in the database search (matched fragment ions over submitted fragment ions). Number in parentheses shows percentage of matching.



**Fig. 3.** Targeted tandem LCMS spectrum for the 40-kDa protein band. The protein band contains (a) troponin T (KPLNIDHLSSEK), (b) tropomyosin alpha-1 chain (EAETRAEFAER), and (c) actin cytoplasmic 1 (HQGVMVGMGQK).

analysis of high-dimensional datasets characterized by few samples and many variables, without losing the information carried by each variable;

in this case, the technique simultaneously analyzed all the variable protein bands [34,36]. Therefore, only proteins of 40, 42, and 49 kDa

would be excised for in-gel tryptic digestion followed by tandem LCMS for protein marker annotation.

### 3.3. Protein and peptide annotation

The decoy, randomized, and concatenated SwissProt database contained 551,705 entries, but only 1416 were related to pigs (*Sus scrofa*). Using data-dependent MS/MS analysis, the most intense precursor ions were selected and then subjected to CID to produce fragmented ions, which consisted of a, b, and y ions as well as an immonium ion and an ion from loss of small neutral molecules [37,38]. The peak lists from each spectrum were deposited into the MS-Taq search engine to get the highest MOWSE score for peptides and their protein annotation. As shown in Table 2, the precursor ions resulting from in-gel tryptic digestion of the 40-kDa protein band were 704.40<sup>2+</sup>, 655.62<sup>2+</sup>, and 586.61<sup>2+</sup> *m/z*, whereas only 411.02<sup>2+</sup> and 611.47<sup>3+</sup> *m/z* were detected for the 42- and 49-kDa protein bands, respectively. The 40-kDa protein band contained proteins such as troponin T, tropomyosin alpha-1 chain, and actin cytoplasmic 1. Other protein bands such as the 42- and 49-kDa ones only consisted of COP9 signalosome complex subunit 4 and ribonuclease inhibitor, respectively. This result was supported by previous studies where the thin filament proteins such as actin, troponin, and tropomyosin had molecular weights ranging from 40 to 45 kDa [30,39]. In this study, the criterion for a good protein marker for pork was that it must be a myofibrillar protein, because myofibrillar proteins are involved in maintaining and organizing meat cells and make up 50% of total meat protein [40].

Fig. 3 shows the MS/MS spectra for a digested 40-kDa protein band. The peptide sequence for the precursor ion of 704.40<sup>2+</sup> *m/z* was KPLNIDHLSSEDK (Fig. 3a). The fragmented ions were b<sub>3</sub>-NH<sub>3</sub> (322.0<sup>+</sup> *m/z*), b<sub>7</sub><sup>2+</sup> (409.8<sup>+</sup> *m/z*), b<sub>4</sub> (453.6<sup>+</sup> *m/z*), b<sub>6</sub>-NH<sub>3</sub> (664.9<sup>+</sup> *m/z*), and y<sub>8</sub>-H<sub>2</sub>O<sup>+2</sup> (469.8<sup>+</sup> *m/z*). The immonium ions are clearly indicated by the one-letter code of the amino acids included, e.g., HLS (337.8<sup>+</sup> *m/z*), IDH (365.9<sup>+</sup> *m/z*), DHLSED (696.5<sup>+</sup> *m/z*), and LNIDHLSSE-NH<sub>3</sub> (1021.3<sup>+</sup> *m/z*). The peptide sequence EAETRAEFAER represented the precursor ion of 655.62<sup>2+</sup> *m/z* (Fig. 3b). Only one type of b ion (b<sub>8</sub>-NH<sub>3</sub><sup>2+</sup>, 459.8 *m/z*) was observed in the MS/MS spectrum. Nevertheless, the a ions were often used as diagnostics for b ions, because their pairings could be observed in low energy CID MS/MS spectra in which a-b ion pairs are separated by 28 Da in correspondence to a carbonyl group [38]. For the precursor ion of 586.61<sup>2+</sup> *m/z*, the peptide sequence suggested by the database after depositing all the fragmented ions was HQGVMVGMGQK (Fig. 3c.) Apart from the listed ions, there were few additional unassigned peaks observed on the MS/MS spectra. This was in agreement with a previous study that claimed it was normal to observe unassigned peaks in MALDI spectra of peptide masses from gel [21].

## 4. Conclusion

In this study, an approach to identify potential protein markers from raw pork was developed by qualitative and quantitative 1DE in combination with multivariate PCA followed by tandem LCMS. By comparing the protein bands on the 1DE, we were able to screen the level of protein expression in all samples differentiated using ANOVA, and a pair-wise multiple comparison was used to identify the potential protein markers. However, these results were limited to a single variable of protein band (univariate) for every analysis and did not project the latent variable (protein bands) with respect to the meat samples. Therefore, we further analyzed the data, using PCA, to enable the interpretation of the complex dataset by dimensional reduction and project the latent variables (protein bands) as well as discriminate all the meat samples to the response variables (protein bands). This improved technique also enabled the confident selection of protein bands for excising prior to in-gel tryptic digestion, instead of solely relying on their optical density. Using the database search program, the suggested protein identifications as well as their peptides were troponin T (KPLNIDHLSSEDK), tropomyosin

alpha-1 chain (EAETRAEFAER), and actin cytoplasmic 1 (HQGVMVGMGQK). However, validation studies are highly recommended for both the protein and peptide markers to confirm marker reliability. This chemometrics-mediated GeLCMS could potentially be used as a guideline to assist chemists in analysis of any gel-based separation of biomolecules, regardless of the field of study.

## Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

## Informed consent

Not applicable.

## Significance of study

Gel-enhanced liquid chromatography-mass spectrometry (GeLCMS) is a study of peptides originating from in-gel digestion of excising protein bands from one-dimensional electrophoresis (1DE) prior to tandem liquid chromatography-mass spectrometry (LC-MS/MS). Typically, the protein bands are selected by comparing the optical density with other protein bands. However, in this study, we propose an improvement in protein bands selection. After protein separation by 1DE, the potential protein bands have been identified by a multivariate principal component analysis (PCA). The identified protein bands from the PCA have been subjected to tandem LC-MS/MS. This qualitative statistical approach provides a clear discrimination of protein bands according to the samples, in this case comparing different meat types (beef, chicken and pork). Whereas in conventional 1DE, the potential protein bands were identified solely based on the optical density. This technique will improve and adds another dimension of selecting the protein bands from the 1DE which can be applied in various research fields.

## Acknowledgements

The study was funded by the Universiti Putra Malaysia (Serdang, Selangor, Malaysia) under Putra Young Initiative Grant (VOT 9466300).

## Conflict of interest

Mohd Hafis Yuswan declares that he has no conflict of interest. Wan Mohd Aizat declares that he has no conflict of interest. Mohd Nasir Mohd Desa declares that he has no conflict of interest. Shuhaimi Mustafa declares that he has no conflict of interest. Amalia Mohd Hashim declares that she has no conflict of interest. Rozi Mohamed declares that she has no conflict of interest. Nur Ajreena Rahim declares that she has no conflict of interest. Dhilia Udie Lamasudin declares that she has no conflict of interest.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.chemolab.2019.103825>.

## References

- [1] J.M. Regenstein, M.M. Chaudry, C.E. Regenstein, The kosher and halal food laws, *J. Food Sci. Technol.* 3 (2003) 111–127.
- [2] J.Z.K. Khattak, et al., Concept of halal food and biotechnology, *Adv. J. Food Sci. Technol.* 3 (5) (2011) 385–389.
- [3] K. Nakyinsige, Y. Che Man, A.Q. Sazili, Halal authenticity issues in meat and meat products, *Meat Sci.* 91 (3) (Jul. 2012) 207–214.
- [4] O. Djurković-Djaković, B. Bobić, A. Nikolić, I. Klun, J. Dupouy-Camet, Pork as a source of human parasitic infection, *J. Clin. Microbiol. Infect.* 19 (7) (2013) 586–594.

- [5] M.E. Ali, M.A. Razzak, S.B.A. Hamid, Multiplex PCR in species authentication: probability and prospects - a review, *Nov, Food Anal. Methods* 7 (10) (2014) 1933–1949.
- [6] V. Fajardo, I. González Isabel, M. Rojas, T. García, R. Martín, A review of current PCR-based methodologies for the authentication of meats from game animal species, *Trends Food Sci. Technol.* 21 (8) (2010) 408–421.
- [7] A.T. Perestam, K.K. Fujisaki, O. Nava, R.S. Hellberg, Comparison of real-time PCR and ELISA-based methods for the detection of beef and pork in processed meat products, *Jan, Food Control* 71 (2017) 346–352.
- [8] M.E. Ali, M.M. Rahman, S.B.A. Hamid, S. Mustafa, S. Bhassu, U. Hashim, Canine-specific PCR assay targeting cytochrome b gene for the detection of dog meat adulteration in commercial frankfurters, *Jan, Food Anal. Methods* 7 (1) (2014) 234–241.
- [9] S. Soares, J.S. Amaral, M.B.P.P. Oliveira, I. Mafrá, A SYBR Green real-time PCR assay to detect and quantify pork meat in processed poultry meat products, *Meat Sci.* 94 (1) (2013) 115–120.
- [10] E.A. Zvereva, et al., Enzyme immunoassay and proteomic characterization of troponin I as a marker of mammalian muscle compounds in raw meat and some meat products, *Jul, Meat Sci.* 105 (2015) 46–52.
- [11] R. Korte, J. Brockmeyer, Novel mass spectrometry approaches in food proteomics, *Nov, Trends Anal. Chem.* 96 (2017) 99–106.
- [12] R. Pedreschi, M. Hertog, K.S. Lilley, B. Nicolai, Proteomics for the food industry: opportunities and challenges, *Jul, Crit. Rev. Food Sci. Nutr.* 50 (7) (2010) 680–692.
- [13] M. Montowska, E. Pospiech, Species-specific expression of various proteins in meat tissue: proteomic analysis of raw and cooked meat and meat products made from beef, pork and selected poultry species, *Feb, Food Chem.* 136 (3–4) (2013) 1461–1469.
- [14] A. Ruiz Orduna, E. Husby, C.T. Yang, D. Ghosh, F. Beaudry, Assessment of meat authenticity using bioinformatics, targeted peptide biomarkers and high-resolution mass spectrometry, *Oct, Food Addit. Contam. A* 32 (10) (2015) 1709–1717.
- [15] T.-Y. Yu, J.D. Morton, S. Clerens, J.M. Dyer, In-depth characterisation of the lamb meat proteome from longissimus lumborum, *Mar, EuPA Open Proteom.* 6 (1) (2015) 28–41.
- [16] C. von Barga, J. Dojahn, D. Waidelech, H.-U. Humpf, J. Brockmeyer, “New sensitive high-performance liquid chromatography–tandem mass spectrometry method for the detection of horse and pork in halal beef, *Dec, J. Agric. Food Chem.* 61 (49) (2013) 11986–11994.
- [17] J.S. Cottrell, Protein identification using MS/MS data, *Sep, J. Proteom.* 74 (10) (2011) 1842–1851.
- [18] J.A. Henkin, M.E. Jennings, D.E. Matthews, J.O. Vigoreaux, Mass processing - an improved technique for protein identification with mass spectrometry data, *J. Biomol. Tech.* 15 (4) (2004) 230–237.
- [19] A. Lundby, J.V. Olsen, GeLCMS for in-depth protein characterization and advanced analysis of proteomes, in: K. Gevaert, J. Vandekerckhove (Eds.), *Gel-Free Proteomics*, vol 753, Humana Press, Totowa, NJ, 2011, pp. 143–155, no. 2.
- [20] Y. Zhang, B.R. Fonslow, B. Shan, M.-C. Baek, J.R. Yates, Protein analysis by shotgun/bottom-up proteomics, *Apr, Chem. Rev.* 113 (4) (2013) 2343–2394.
- [21] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels, *Jan, Anal. Chem.* 68 (5) (1996) 850–858.
- [22] Y.-C. Huang, T.-Y. Chen, S.-S. Jeng, H.-M. Chen, D.-F. Hwang, Species identification of a causative moray eel meat by SDS-PAGE, *J. Food Drug Anal.* 18 (4) (2010) 232–238.
- [23] M. Montowska, E. Pospiech, Species identification of meat by electrophoretic methods, *Acta Sci. Pol. Technol. Aliment* 6 (1) (2007) 5–16.
- [24] T. Nur Azira, Y.B. Che Man, R.N. Raja Mohd Hafidz, M.A. Aina, I. Amin, Use of principal component analysis for differentiation of gelatine sources based on polypeptide molecular weights, *May, Food Chem.* 151 (May) (2014) 286–292.
- [25] T. Nur Azira, I. Amin, Y.B. Che Man, Differentiation of bovine and porcine gelatins in processed products via sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and principal component analysis (PCA) techniques, *Int. Food Res. J.* 19 (3) (2012) 1175–1180.
- [26] H.-R. Aerni, D.S. Cornett, R.M. Caprioli, Automated acoustic matrix deposition for MALDI sample preparation, *Anal. Chem.* 78 (3) (2006) 827–834.
- [27] F.-C. Chen, Y.-H.P. Hsieh, R.C. Bridgman, Monoclonal antibodies to porcine thermal-stable muscle protein for detection of pork in raw and cooked meats, *Jun, J. Food Sci.* 63 (2) (1998) 201–205.
- [28] S. Wen, et al., Discrimination of in vitro and in vivo digestion products of meat proteins from pork, beef, chicken, and fish, *Nov, Proteomics* 15 (21) (2015) 3688–3698.
- [29] X. Xia, B. Kong, J. Liu, X. Diao, Q. Liu, Influence of different thawing methods on physicochemical changes and protein oxidation of porcine longissimus muscle, *Apr, LWT - Food Sci. Technol. (Lebensmittel-Wissenschaft -Technol.)* 46 (1) (2012) 280–286.
- [30] S. Ali, W. Zhang, N. Rajput, M.A. Khan, C. Li, G. Zhou, Effect of multiple freeze-thaw cycles on the quality of chicken breast meat, *Apr, Food Chem.* 173 (2015) 808–814.
- [31] A. Martín, M.A. Asensio, M.E. Bermúdez, M.G. Córdoba, E. Aranda, J.J. Córdoba, Proteolytic activity of *Penicillium chrysogenum* and *Debaryomyces hansenii* during controlled ripening of pork loins, *Sep, Meat Sci.* 62 (1) (2002) 129–137.
- [32] F. Huang, M. Huang, X. Xu, G. Zhou, Influence of heat on protein degradation, ultrastructure and eating quality indicators of pork, *Feb, J. Sci. Food Agric.* 91 (3) (2011) 443–448.
- [33] E. Bäckryd, B. Ghafouri, A.K. Carlsson, P. Olausson, B. Gerde, Multivariate proteomic analysis of the cerebrospinal fluid of patients with peripheral neuropathic pain and healthy controls - a hypothesis-generating pilot study, *Jul, J. Pain Res.* 8 (8) (2015) 321–333.
- [34] S. Wold, M. Sjöström, Chemometrics, present and future success, *Dec, Chemometr. Intell. Lab. Syst.* 44 (1–2) (1998) 3–14.
- [35] M. Montowska, M.R. Alexander, G.A. Tucker, D.A. Barrett, Rapid detection of peptide markers for authentication purposes in raw and cooked meat using ambient liquid extraction surface analysis mass spectrometry, *Oct, Anal. Chem.* 86 (20) (2014) 10257–10265.
- [36] S. Mazzara, et al., Application of multivariate data analysis for the classification of two dimensional gel images in neuroproteomics, *J. Proteom. Bioinform.* 4 (1) (2011) 16–21.
- [37] K.F. Medzihradsky, Peptide sequence analysis, *Methods Enzymol.* 402 (2000) 209–244.
- [38] H. Steen, M. Mann, “The abc’s (and xyz’s) of peptide sequencing, *Sep, Nat. Rev. Mol. Cell Biol.* 5 (9) (2004) 699–711.
- [39] Y.L. Xiong, Myofibrillar protein from different muscle fiber types: implications of biochemical and functional properties in meat processing, *Jan, Crit. Rev. Food Sci. Nutr.* 34 (3) (1994) 293–320.
- [40] E. Huff-Lonergan, Chemistry and biochemistry of meat, in: F. Toldrá (Ed.), *Handbook of Meat Processing*, first ed., Wiley-Blackwell, Oxford, UK, 2010, pp. 5–24.