

Rapid Protein Identification Using MALDI-TOF-MS Biopolymer Mapping: A New Strategy

Larry Engay¹, Isaac Lee¹, Huiyong Ma¹, Quanyin Gao^{1*}, Teddy Collins², Peter Chang³ and Gary Swanson³

¹Herbalife Nutrition, 20481 Crescent Bay Drive, Lake Forest, CA 92630, USA

²Herbalife Nutrition, 3200 Temple School Rd, Winston-Salem, NC 27107, USA

³Herbalife Nutrition, 990 West 190th Street, Torrance, CA 90502, USA

Abstract

A rapid identification method using high mass range matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS) with direct polymer mapping (analyze material as-is without digestion or separation) is developed and demonstrated to be successful in rapid identification of incoming protein raw materials for quality control. This identification method is using an established MALDI-TOF-MS protein library from eight protein standard materials. Future received lots of protein raw materials can be tested for identification based on their MALDI-TOF-MS biopolymer mapping profile compared to those protein mass spectra profiles in the established library. This test method is demonstrated to be specific, rapid, and suitable for its intended use of eight protein materials in our nutrition products with a potential for identification of more raw material types of biopolymers or complicated natural raw materials. To the best of our knowledge this is the first report of such high mass range MALDI-TOF-MS with direct polymer mapping method for rapid protein identification of incoming raw material quality control.

Introduction

Protein material identification is challenging in quality control for raw material testing since the nature of proteins are biopolymers rich in diversity with different sizes and globular structures. Typical quality control identification technique for organic molecules such as FT-IR cannot differentiate proteins such as soy protein isolate vs. whey protein concentrate. FT-IR spectra of these two common proteins give similar infrared spectra with more than 0.95 correlation factor that are usually considered as the same material instead of different raw materials (correlation factor of 0.90 and above matching is typically used for FT-IR confirmation on identification). The reason different proteins exhibit similar FT-IR spectra is due to the facts that they are polymer molecules composed of the same basic amino acids with similar chemical bond vibrational modes. In current Food Chemical Codex for soy protein concentrate and whey protein concentrate [1,2], the identification requirements are not using infrared but based on a sample exhibits the compositional profile with respect to Ash, Fat, Loss on Drying, and Protein content (soy protein). The same compositional profile is required for whey protein concentrate identification using Ash (Total), Fat, Lactose, Loss on Drying, and Protein content [2]. These parameters are not specific enough for identification test from science point of view, particularly, the protein content is measured using elemental nitrogen content and we know nitrogen containing molecules are not limited to proteins. In FCC APPENDIX XVI PROTEIN-BASED INGREDIENTS [3], two new tests are available for protein identification: 1) Peptide Mapping for Identification of Species in Protein Ingredients and 2) Amino Acid Fingerprinting for Bovine Skim Milk Powder and Nonfat Dry Milk. Even though these approaches addressed concern on method specificity, but these methods are tedious and time consuming. These methods usually involve enzymatic digestion of protein to smaller peptide molecules coupled with separation of peptides followed by mass spectrometry peptide mapping. Soy (glycinin G1), Pea (vicilin), Rice (glutelin), Whey (β -lactoglobulin), and Casein (α -S1-casein) peptides were used to facilitate identification of the original protein raw material.

Such peptide mapping for protein identification approach is also reflected in AOAC Official Methods of Analysis (OMA) 2017.11 Identification of Pea, Rice, and Soy Proteins in Raw Materials and

Finished Goods with ESI HPLC-MS/MS [4] and OMA 2017.12 (OMA 2017.12) Identification of Milk Proteins in Raw Materials and Finished Goods [5].

In addition to the above peptide mapping approach, a non-selective full hydrolysis of proteins into individual amino acids was tested for identification purpose (USP FCC APPENDIX XVI Amino Acid Fingerprinting and OMA 2016.15) [6]. The challenge of such approach is that different proteins with different primary, secondary, and tertiary structures may result in similar amino acids proportionality and the proportion difference may not be enough to identify the protein source. Furthermore, this full hydrolysis followed by quantitating individual amino acids results in tedious analytical process in addition to the concern on method specificity with limited capacity for routine protein lot release in a fast-paced production site.

Historically, MALDI-TOF-MS technique has been used for proteomics studies (protein expression and function). Often it is linked to two-dimensional gel electrophoresis that separates different kind of proteins followed by protease digestion breaking the protein polymer into peptide fragments. These fragments can be analyzed using MALDI-TOF-MS with database search to trace back to the type of proteins by these peptide fragments, a peptide mass fingerprint approach [7,8]. Such kind of separation (gel electrophoresis or HPLC) following by enzymatic digestion then analysis with low mass range MALDI-TOF MS peptide mapping is tedious and time consuming.

To further advance the identification test on efficiency and capacity

***Corresponding author:** Quanyin Gao, Herbalife Nutrition, 20481 Crescent Bay Drive, Lake Forest, CA 92630, USA, E-mail: quanying@herbalife.com

Received: 11-Nov-2023, Manuscript No: jabt-23-119872, **Editor assigned:** 13-Nov-2023, PreQC No: jabt-23-119872 (PQ), **Reviewed:** 24-Nov-2023, QC No: jabt-23-119872, **Revised:** 29-Nov-2023, Manuscript No: jabt-23-119872 (R), **Published:** 30-Nov-2023, DOI: 10.4172/2155-9872.1000578

Citation: Engay L, Lee I, Ma H, Gao Q, Collins T, et al. (2023) Rapid Protein Identification Using MALDI-TOF-MS Biopolymer Mapping: A New Strategy. J Anal Bioanal Tech 14: 578.

Copyright: © 2023 Engay L, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

we have developed a rapid protein identification method using high mass range MALDI-TOF mass spectrometry (up to 20 kilodalton) with direct biopolymer mapping (analyze the sample “as-is” without digestion and separation). Such approach using high mass range MALDI-TOF MS capability to overcome earlier low mass range MALDI-TOF MS limitation (up to 4000 Dalton) with no digestion is required.

Cosima D. Calvano et al. [8] have recently analyzed several sport nutritional supplements of powder products containing important proteins as milk, soy and egg using MALDI-TOF MS coupled with SDS-PAGE using peptide mapping and bioinformatics searching that positively identified proteins. In our study for lab testing efficiency on protein raw material identification high mass range MALDI-TOF-MS with direct biopolymer mapping are used but on a complicated product matrix for protein identification a separation may be needed.

Our newly developed method for protein raw material identification uses a sample in an organic acid solution matrix on a metal target plate, and a laser shot on the metal plate will desorb the proteins from the matrix and ionize the proteins by proton adduct formation from the matrix acid. The detected protein ions by time-of-flight (TOF) mass spectrometry are specific to the protein molecular weight distribution and globular structure. MALDI is a soft ionization and majority of the ions are singly charged. Fast analysis of multiple lots of samples on one target metal plate can be performed with same-day turnaround time in a quality control lab. Direct measurement of these protein polymer molecules mass profiles can differentiate proteins with high specificity. The obtained sample mass spectrum is then compared with an established protein mass spectra profile library for a matching score to determine positive identification (score 2.000 or above for our application using Bruker Biotyper instrument and software).

In comparison to our early studies using differential multiplex PCR assay for the identification of proteins [9] and using microbiome profiling with 16S metagenomics for identification of Milk and Whey proteins [10], this MALDI-TOF-MS polymer mapping method is rapid, simple, and specific for protein material identification. Such identification method can potentially be applicable to more raw material types of biopolymers or complicated natural raw materials.

Experimental Section

Instrument

All MALDI-TOF mass spectrometry experiments were conducted on a Bruker MALDI-TOF Biotyper® Sirius mass spectrometer (Bruker, Billerica MA, USA) equipped with a Bruker MBT Compass HT, flex Analysis, and Compass Explorer Software suite. Protein mass spectra were obtained in the positive-ion mode, and initial laser power set to 35%. The instrument was programmed to automatically adjust laser powder to optimize signal to noise ratio, until an accumulate 240 satisfactory mass spectra were obtained. Voltages were set for ion source one to 20.00 kV, ion source two to 18.15 kV, lens to 6.00 kV, and linear detector to 2.694 kV. Mass to charge ratio (m/z) was range in between 2-20 kilo Dalton (kDa). Detector gain was set to 3.6X (at 2694 V).

Mass calibrations were performed using Bruker Bacterial Test Standard (BTS) prior to each set of MALDI-TOF run.

Materials

(a) **Standard solvent:** 50% ACN+47.5% H₂O+2.5% TFA. (Bruker Scientific, Billerica, Massachusetts, Part# 900666). This solution is used

for preparation of test samples, and Bruker Bacterial Test Standard.

(b) **Bacterial test standard (BTS):** (Bruker Scientific, Billerica, Massachusetts, Part# 8255343) contains a carefully manufactured extract of *Escherichia coli* characteristic peptide and 2 additional spiked proteins for MALDI-TOF mass spectra calibration. The overall mass range covered by Bruker BTS is 3.6 to 17 kDa.

(c) **HCCA (α-Cyano-4-hydroxycinnamic acid) Matrix:** (Bruker Scientific, Billerica, Massachusetts, par# 8255344) HCCA enables easy and convenient preparation of HCCA matrix solutions. The matrix is soluble in standard solvent (acetonitrile 50%, water 47.5% and trifluoroacetic acid 2.5%) that enables highly sensitive MALDI-TOF-MS measurement of peptides and proteins. Table 1 shows the chemical structure and property of HCA.

(d) Soy protein isolate of supplier 1, was obtained from Archer Adam Midland Company, 4666 East Faries Parkway, Decatur, IL62526, protein content ≥90%.

(e) Soy protein isolate of supplier 2, was obtained from Solae, LLC, 5532 Hunt St., Highway 412B, MAIP, Pryor, OK 74361, protein content ≥90%.

(f) Milk protein isolate was obtained from Idaho Milk Products Inc., 2249 South Tiger Drive, Jerome, ID83338, protein content ≥85%.

(g) Whey protein concentrate was obtained from Milk Specialties Global, 7500 Flying Cloud Drive, Suite 500, Eden Prairie, MN 55344, protein content ≥80%.

(h) Flaxseed powder was obtained from Bioriginal Europe/Asia B. V., Bosland 40, 3258 AC Den Bommel, The Netherlands, protein content ≥32%

(i) Pea protein was obtained Farbest Brands, One Maynard Drive, Park Ridge, NJ 07656, protein content ≥83%

(j) Quinoa powder was obtained from Quinoasure, FACTORIA QUINOA SAS, Zona Fr. Pacifico Km 6, Carretera Yumbo Aerop. Bodega 1, PALMIRA, COLOMBIA, protein content ≥14%

(k) Rice protein was obtained from Oryzatein, AxiomFoods, 12100 Wilshire Blvd #800, Los Angeles, CA90025 with protein content ≥80%

(l) Sesame protein concentrate was obtained from Dipasa, 6600 East FM802, Brownsville, TX78521, protein content ≥47%

BTS preparation

(a) Remove Bacteria Test Standard (BTS) from a freezer and equilibrate to room temperature. BTS standard is a mixture of *Escherichia coli* and spiked proteins used for mass calibration shown in Table 2

Table 1: Structure and Property of HCCA (α-Cyano- 4-hydroxycinnamic acid).

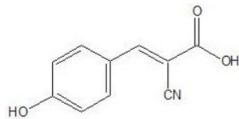
Synonyms	2-Cyano-3-(4-hydroxyphenyl) acrylic acid
Molecular formula	C ₁₀ H ₇ NO ₃
Structural formula	
Molecular weight	189.17g/mol
CAS number	28166-41-8
EC number	248-879-1
Melting point	245 – 250 °C

Table 2: MALDI-TOF-MS molecular weight calibration table.

Protein	Reference mass (average mass)	±300 ppm range
RL29 [M+2H] ²⁺	3637.8 Da	3636.7 Da–3638.8 Da
RS32 [M+H] ⁺	5096.8 Da	5095.3 Da – 5098.3 Da
RS34 [M+H] ⁺	5381.4 Da	5379.8 Da – 5383.0 Da
RS33meth [M+H] ⁺	6255.4 Da	6253.5 Da – 6257.3 Da
RL29 [M+H] ⁺	7274.5 Da	7272.3 Da – 7276.7 Da
RS19 [M+H] ⁺	10300.1 Da	10297.0 Da – 10303.2 Da
RNAse A [M+H] ⁺	13683.2 Da	13679.1 Da – 13687.3 Da
Myoglobin [M+H] ⁺	16952.3 Da	16947.2 Da – 16957.4 Da

The overall mass range calibrated by Bruker BTS is 3.6 to ~17 kDa with mass tolerance limit of ± 300 ppm.

(b) Add 50 µL of Bruker standard solution and dissolve by pipetting up and down (mixing) 20 times limiting bubbles.

(c) Wait 5 min. at room temperature and mix via pipet for an additional 20 times, limiting bubbles.

(d) Centrifuge 10 s at 14,000 rpm.

HCCA matrix preparation

(a) Remove HCCA from refrigeration and wait 5 minutes to bring the matrix tube to room temperature.

(b) Add 250 µl of Bruker standard solution to HCCA (10 mg/mL concentration), and completely dissolve the HCCA by vortex.

(c) Centrifuge HCCA tube at maximum force for 10 s.

Sample preparation

(a) Weigh ~ 2 mg of protein sample in a micro centrifuge tube and add 200 µl of Bruker Standard solvent.

(b) Vortex the tube and Incubate it at 35°C for 35 min and agitate at 900 rpm on a microplate shaker.

(c) Once incubation has completed, centrifuge 10 s at 14,000 rpm

Test procedure

(a) Spot 1 µl of BTS on a MALDI target plate and air dry for mass calibration use.

(b) Spot 1 µl of supernatant of protein sample prepared above on the same plate on a different spot on the MALDI target plate and dry it.

(c) Overlay 1 µl of HCCA matrix over the dried spots of BTS and proteins sample on the MALDI plate and air dry until no liquid remains.

(d) Acquire MALDI-TOF mass spectra

(e) Run identification of samples against customized protein library using Bruker MBT Compass HT or Bruker Compass HD.

Results and Discussion

Mass spectra and peak assignment

In this MALDI-TOF-MS polymer mapping study, eight types of protein raw materials were tested. These mass spectra for each type of protein material are presented. Mass spectrum peak assignments from these proteins are discussed. Where available, proteins and related subunits with molecular weights are presented.

MALDI-TOF-MS mass spectra of these protein materials are shown in Figures 1-9.

(a) **Soy protein:** For soy protein isolate sample mass spectra (Figures 1 and 2) strong molecular ions at about 4.7 kDa were observed for both suppliers’ soy protein isolate materials (Figures 1 and 2). Previously, Cheryl Lynch reported observation of glycinin protein subunits with estimated molecular weight ranging from about 5-7 kDa. We have observed in the mass spectra of soy portions at 4.7 k Da that can be assigned to a sub-unit of glycinin proteins reported earlier [11]. This peak at 4.7 kDa is characteristic for soy proteins. For both suppliers we have obtained similar mass spectra of soy protein isolate.

(b) **Milk protein:** Milk protein concentrate sample mass spectra (Figure 3) indicated peaks at 15,465, 15,127, 14,812, 14,498 Dalton at high molecular range together with peaks at 4,629, 4,305, 3,981, 3,475, 3,118, 2,761, and 2,457 Daltons at lower mass range. These peaks reflected the composition of both whey proteins and casein. Whey protein is observed at high molecular weight with peaks from 14.5 to 15.5 kDa, representing α-Lactalbumin in whey proteins. The low molecular weight peaks, such as strongest peak at about 3475 Dalton, are related to beta-casein [12]. There are multiple peaks observed at 4629, 4305, 3981, 3118, 2761, and 2457 Daltons. Since milk protein has

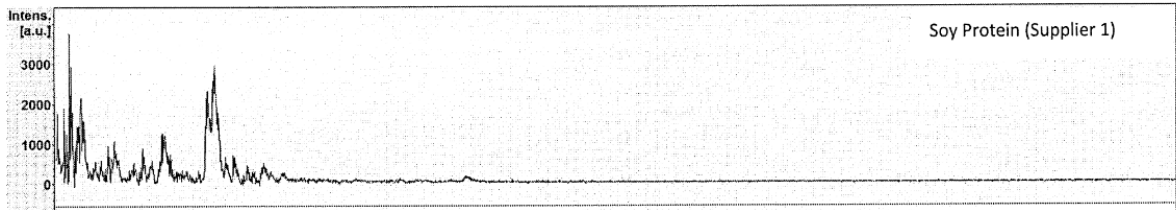


Figure 1: Soy Protein MALDI-TOF Mass Spectrum (Supplier 1), Peaks at 4.7 kDa.

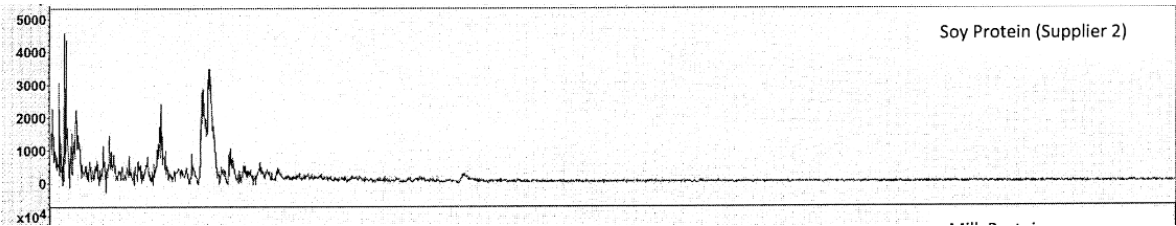


Figure 2: Soy Protein MALDI-TOF Mass Spectrum (Supplier 2), Peaks at 4.7 kDa.

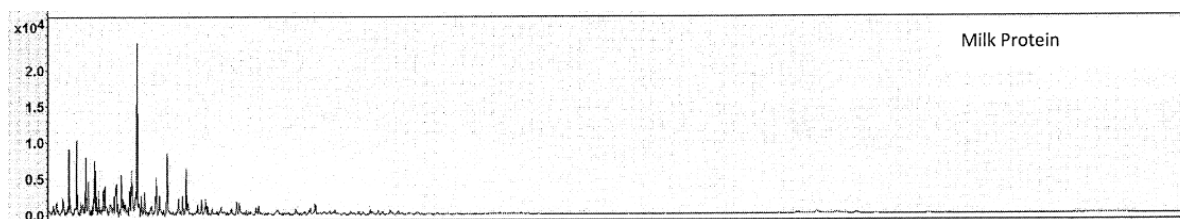


Figure 3: Milk Protein MALDI-TOF Mass Spectrum, Peaks at 14.5 to 15.5 kDa.

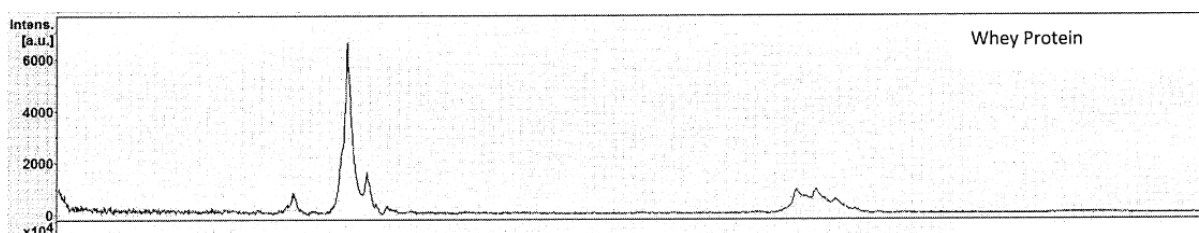


Figure 4: Whey Protein MALDI-TOF Mass Spectrum, Peaks at 14.2 to 14.8 kDa.

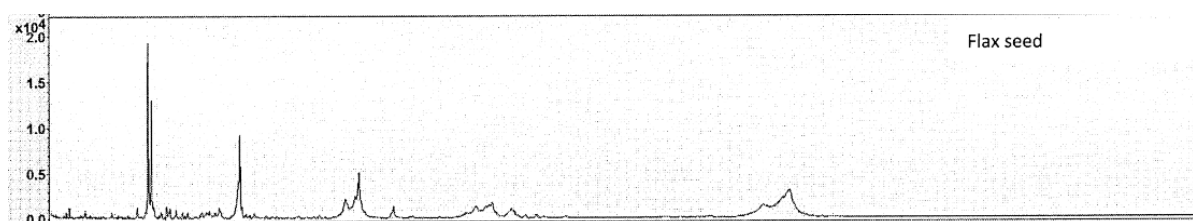


Figure 5: Flaxseed Protein MALDI-TOF Mass Spectrum, Peaks at 15-18 kDa.

many types of proteins and these proteins/peptides below 10 kDa were not further studied.

(c) **Whey protein:** In whey protein concentrate sample analysis, there are three known proteins of α -Lactalbumin, β -Lactoglobulin, and Lactoferrin with molecular weight at about 14.2, 18.3, and 86 kDa, respectively [13]. Figure 4 for whey protein mass spectrum has clearly shown multiple peaks close to 14.2 to 14.8 kDa (14184, 14503, and 14829 Dalton) that are related to α -Lactalbumin. Additional peaks at 7115, 6791 and 5894 Dalton are also observed. These peaks are not identified at present time.

(d) **Flaxseed protein:** Flaxseed powder sample was studied for this protein. We have observed very defined mass spectra peaks at 14,166, 13,699, 9,587, 9,273, 9,020, 7,659, 7,092, 6,871, 5,136, 3,669, and 3,612 Daltons (Figure 5). There is little information in the literature for these peak assignments. In flaxseed protein review [14] lower molecular weight level 2S proteins of flaxseed was proposed to be of 15-18 kDa. 10-15 kDa molecular weight fractions have already been reported by Chung et al. for flaxseed proteins [15].

Our MALDI-TOF mass spectrum of major peaks around 14,166 and 13,699 Da are likely 2S proteins from Flaxseed. We did not observe higher molecular weight peaks since higher molecular weight molecules are likely to have lower volatility for the laser induced desorption and ionization that lead to low chance for observance. 12S proteins from flaxseed with molecular weights of 11, 18, 29, 42 and 61 kDa were reported previously [16].

(e) **Pea protein:** MALDI-TOF mass spectrum for pea protein indicated peaks are 13,777, 13,615, 13,454, 13,289, 12,080, 6,824, 6,735, 6,653, 5,954, 5086, 3,858, 3,808, 3,757 and 3,007 Daltons (Figure 6). These high molecular weight peaks we observed here were not reported

in early study using LC/MS/MS of quadruple mass spectrometer as typical quadrupole instrument mass range is up to 3000 Dalton [17]. Their reported mass spectra peaks are at m/z 288.35, 294.31, and 295.34, where multiple charging can occur that the molecular ions will be higher in multiples due to high charging state (z value). Munialo et al have reported dominant bands at ~70, 50, 33–14 kDa are vicilin/convicilin fraction subunits of pea proteins [18]. Our observed 13-14 kDa peaks can be attributed to vicilin/convicilin proteins (a legumin-associated storage globulin proteins/a trimeric globulin related to vicilin in pea) as the component of pea proteins.

González-Pérez and Arellano in Handbook of Hydrocolloids [19] reported that Vicilin, the 7S pea protein, is generally found as a trimer, each subunit having a molar mass of about 50 kDa. Vicilin can be cleaved at one or two sites (called the α - β and β - γ processing sites) as specified by the coding sequence of the vicilin genes. Cleavage at the α - β site generates fragments of 19 and 30 kDa, whereas that at the β - γ site generates fragments of 12.5 or 16 and 33 kDa. Cleavage at both sites generates fragments of 12.5, 13.5 and 16 or 19 kDa. Clearly our observed peaks in MALDI-TOF at 13.5 kDa and 12.1 kDa are matching the Vicilin 7S protein with bond cleavage by the laser assisted desorption/ionization process.

(f) **Quinoa protein:** We have studied Quinoa powder sample for this protein and observed mass spectra with peaks at 10,512, 7,839, 6,269, 5,165, 4,391, 4,024, and 2,895 Daltons (Figure 7). There are early works [20,21] that assigned the most abundant SDS-PAGE bands for Quinoa proteins around 10,000, 15,000-35,000 and 50,000 Dalton to 2S albumins, 11S globulins, and 7S globulins, respectively. Our observed peak at 10512 is likely from 2S albumins [21].

(g) **Rice protein:** Rice protein sample MALDI-TOF mass spectra indicated peaks at 5745, 4755, 4657, 4117, 3348, and 2698 Daltons



Figure 6: Pea Protein MALDI-TOF Mass Spectrum, Peaks at 13-14 kDa .

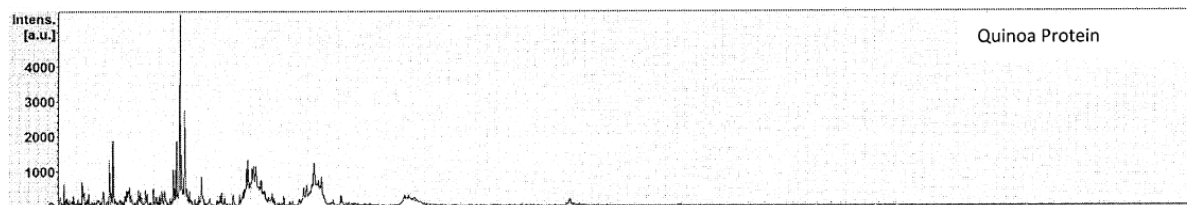


Figure 7: Quinoa Protein MALDI-TOF Mass Spectrum, peaks at 10,512, 7,839, 6,269, 5,165, 4,391, 4,024, and 2,895.

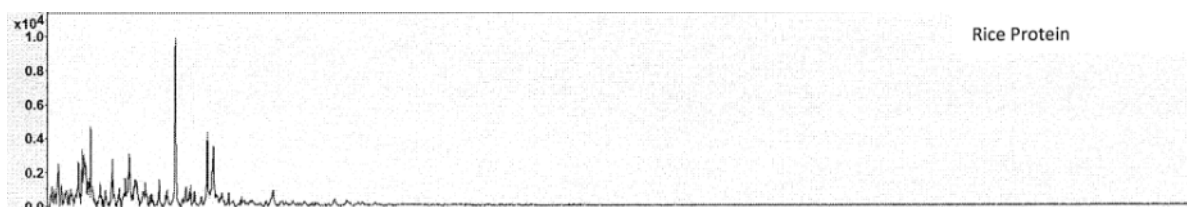


Figure 8: Rice Protein MALDI-TOF Mass Spectrum, Peaks at 1 KDa to 150 KDa.

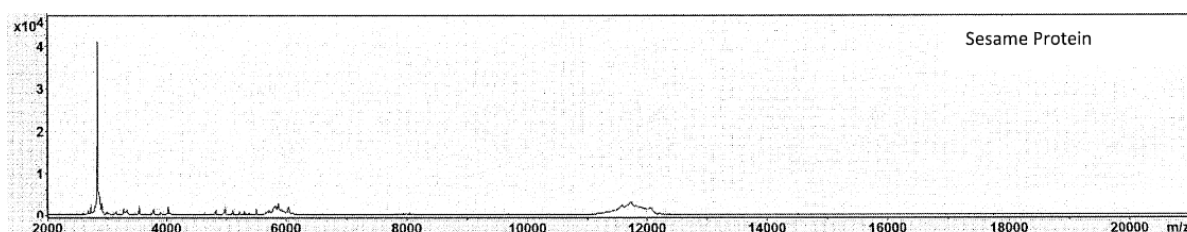


Figure 9: Sesame protein MALDI-TOF Mass Spectrum, Peak at 12,058, 11,719, 11,576, and 7,955.

(Figure 8). Generally, the major fractions of rice endosperm protein are glutelin (66-78%), followed by globulin (9.6-10.8%), albumin (3.8-8.8%), and prolamin (2.6-3.3%) [22].

Typical intact rice protein, with major bands at the molecular weights of 11, 12, 19, 20, and 31 kDa [23]. These high molecular weight proteins were not observed for the rice protein material in our MALDI-TOF-MS study.

J.S. HAMAD has provided a table for rice bran proteins hydrolysate from molecular weight of 1 KDa to 150 KDa [24]. The proteins hydrolysates ranging from 3-10 KDa have about 15% (by area) of the total proteins in size exclusion chromatograph isolation. Our MALDI-TOF observed peaks of 5.7, 4.7, 4.1, and 3.3 KDa are matching these protein hydrolysates mass ranges.

(h) Sesame protein: Sesame protein concentrate sample mass spectrum indicated a broad molecular weight of 12,058, 11,719, 11,576, and 7,955 Daltons in addition to peaks at 5,866, 4,975, 3,527, and 2,825 Daltons (Figure 9).

Immunoglobulin E (IgE)-binding proteins were identified at 78, 52, 45, 34, 32, 29, 25, 20, 9, and 7 kDa. Analyzing internal sequences, the protein at 45 kDa was found to be a 7S vicilin-type globulin, a seed storage protein of sesame. The peak at 7 kDa was reported to

be a 2S albumin [25]. The molecular weights of 7S globulins (36-66 kDa), acidic 11S aggregates (29-32 kDa), basic 11S aggregates (20-22 kDa), lower polypeptides (14-20 kDa) and 2S albumins (6-8 kDa) were reported [26].

Our MALDI-TOF observed mass at 7,955 Da can potentially be attributed to 2S albumin of sesame proteins. The rest of the peaks observed in our MALDI-TOF-MS are not assigned. Molecular weight less than 7 kDa are likely peptides from the processing of the Sesame proteins in manufacturing.

Table 3 is a summary of the signature peaks in each protein observed from our MALDI-TOF-MS study. The peak m/z values are correlated with literature molecular weight information in above discussion for each protein.

Specificity of the MALDI-TOF-MS method

The mass spectra analysis indicates that each type of protein has its own specific characteristic that is different in mass spectrum profile compared to any other type of proteins as observed in the overlay mass spectra of all eight protein materials studied (Figure 10). The soy protein isolates from two different suppliers have similar mass spectra (Figures 1 and 2). Using these protein mass spectra, we have built an in-house protein library. The library is created by gathering at least 20 different

Table 3: MALDI-TOF-MS Protein Peak Assignment.

Category	Characteristic m/z peaks observed	Peak assignment
Soy protein	4.7 kDa	sub-unit of glycinin proteins
Milk protein	14.5 to 15.5 kDa	α-Lactoalbumin (whey proteins)
	3475 Da	beta-casein
Whey protein	14.2 – 14.8 kDa	α-Lactoalbumin
Flaxseed protein	14.2 and 13.7 kDa	2S proteins
Pea Protein	12.1 and 13.5 kDa	7S protein
Quinoa Protein	10.5 KDa	2S albumins
Rice Protein	5.7, 4.7, 4.1, and 3.3 kDa	Protein hydrolysates
Sesame protein	7,955 Da	2S albumin

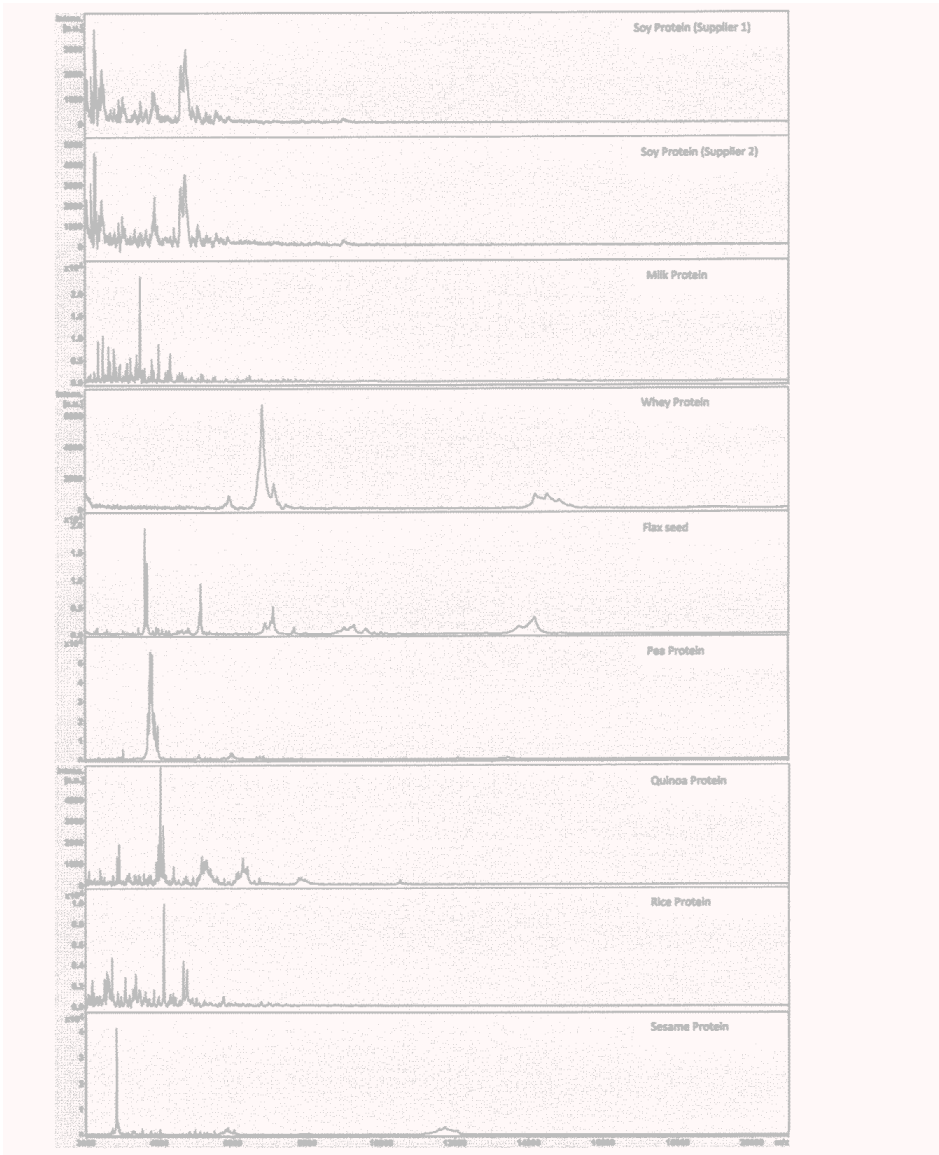


Figure 10: Overlaid MALDI-TOF mass spectra of different protein materials.

reference standard spectra to generate a representative mass spectrum profile. Accurate identification by the mass spectra matching using the library search function generated matching score was achieved for positive identification (log score 2.000 or above).

New lots of protein raw materials using high mass range MALDI-TOF MS identification test method with library matching score are

successful for multiple lots of soy protein isolates and milk proteins. Examples of the method applications with new lots received are shown in Figures 11 for soy protein isolate materials and Figure 12 for milk protein concentrate with corresponding scores shown in Tables 4 and 5. In Table 4, the only matching score higher than 2.000 is Soy Protein (score of 2.100) with all the rest of other proteins matching score of 0.7100 or lower, indicating a positive identification of Soy Protein for

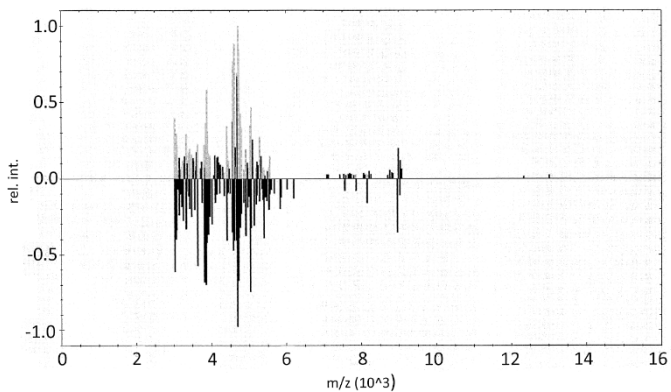


Figure 11: Comparison of Soy protein sample and the library mass spectra matching score (Sample spectrum is the upper half of the plot and the library spectrum is the lower half of the plot).

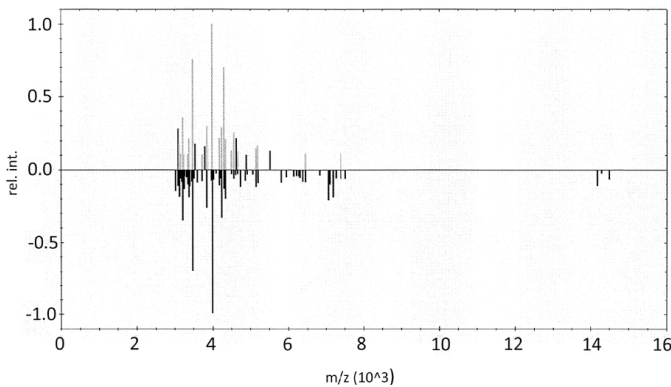


Figure 12: Milk protein sample and the library mass spectra comparison (Sample spectrum is the upper half of the plot and the library spectrum is the lower half of the plot).

Table 4: An example of a new lot soy protein isolate library matching scores.

Protein Library Name	Log(Score)
Soy	2.100
Flax seed	0.710
Milk	0.580
Pea	0.490
Quinoa	0.420
Whey	0.370
Rice	0.080
Sesame	<0

Table 5: An example of a new lot milk protein concentrate library matching scores.

Protein Library Name	Log(Score)
Soy	2.300
Flax seed	1.140
Milk	0.700
Pea	0.170
Quinoa	0.160
Whey	<0
Rice	<0
Sesame	<0

this sample. In Table 5, one can see the only matching score higher than 2.000 is milk protein (score of 2.300) and the rest of the protein scores are 1.14 or lower indicating a positive identification of Milk Protein for this sample.

From these study results and discussion above, our high mass range MALDI-TOF MS with direct polymer mapping identification method is demonstrated to be a successful rapid identification method. Multiple lots of received protein raw materials can be tested on same day turnaround time for Quality Control.

Conclusion

High mass range MALDI-TOF MS identification of proteins based on direct protein polymer mass mapping is demonstrated to be rapid in analysis and specific for its selectivity. This analytical method is proven to be suitable for rapid identification test in a quality control laboratory.

Acknowledgments

The authors thank Bruker on their recommendation for the MALDI-TOF MS instrument type purchased for this project and their recommendation of their cataloged reagents and matrix for the instruments.

Conflict of Interest

During rapid protein identification method development using MALDI-TOF mass spectrometry, all authors were employed by Herbalife Nutrition.

References

1. USP-FCC Food Chemical Codex (2023) Monograph on Soy Protein Concentrate.
2. USP-FCC, Food Chemical Codex (2023) Monograph on Whey Protein Isolate.
3. USP-FCC (2023) Food Chemical Codex APPENDIX XVI.
4. AOAC Official Methods of Analysis 21st Edition (2019) AOAC INTERNATIONAL: 2017, 11.
5. AOAC Official Method of Analysis 21st Edition (2019) Identification of Milk Proteins in Raw Materials and Finished Goods, ESI HPLC-MS/MS. AOAC INTERNATIONAL
6. AOAC Official Method of Analysis 21st Edition (2019) Quantification of Whey Protein Content in Milk-Based Infant Formula Powders, Sodium Dodecyl Sulfate-Capillary Gel Electrophoresis (SDS-CGE).
7. D'auria E, Agostoni C, Giovannini M, Riva E, Zetterstrom R, et al. (2005) Proteomic evaluation of milk from different mammalian species as a substitute for breast milk. 1708-1713.
8. De Ceglie C, Calvano CD, Zamboni CG. MALDI-TOF MS for quality control of high protein content sport supplements. Food Chemistry 176: 396-402.
9. Thompson C, Gao Q, Lu Z, Babajanian S, Chang P, et al. (2020) Development of a differential multiplex PCR assay for the supplemental identification of different sources of proteins. J. AOAC Int. 103: 205-209.
10. Thompson C, Gao Q, Chang P, & Swanson G (2021) Differentiation of Milk and Whey Protein Concentrates by Microbiome Profiling Using 16S Metagenomics. J. AOAC Int. 104: 757-764.
11. Lynch CJ, Rha CK, & Catsimpoalas N (2017) Tryptic hydrolysis of glycine and its subunits. J. Sci. Food Agric 28: 971-979.
12. Cunsolo V, Muccilli V, Saletti R, & Foti S (2011) Applications of mass spectrometry techniques in the investigation of milk proteome. Eur. J. Mass Spectrom 17: 305-32.
13. Gulzar M, Bouhallab S, Jardin J, Briard-Bion V, & Croguennec T (2013) Structural consequences of dry heating on alpha-lactalbumin and beta-lactoglobulin at pH 6.5. Food Research International 51: 899.
14. Oomah BD, & Mazza G, (1993) Flaxseed proteins—a review. Food Chem 48: 109-114.
15. Chung MWY, Lei B, & Li-Chan ECY (2005) Food Chem 90: 271-279. doi: 10.1016/j.foodchem.2003.07.0381
16. Madhusudhan KT, & Singh N (1985) Isolation and characterization of the major

- fraction (12 S) of linseed proteins. *J. Agric. Food Chem* 33: 673-677.
17. Li H & Aluko RE (2010) Identification and inhibitory properties of multifunctional peptides from pea protein hydrolysate. *J. Agric. Food Chem* 58: 11471-11476.
18. Munialo CD, Martin AH, Van Der Linden E, & De Jongh HHJ (2014) Fibril formation from pea protein and subsequent gel formation. *J. Agric. Food Chem* 62: 2418-2427.
19. González-Pérez S & Arellano JB (2009) *Handbook of Hydrocolloids*, Wood head Publishing Series in Food Science, Technology and Nutrition: 383-419, ISBN 978-1-84569-587-3
20. Galindo-Luján R, Pont L, Minic Z, Berezovski MV, Sanz-Nebot V, et al. (2021) Characterization and differentiation of quinoa seed proteomes by label-free mass spectrometry-based shotgun proteomics. *Food Chem* 363: 130250.
21. Brinegar C, Sine B, & Nwokocha L (1996) High-Cysteine 2S Seed Storage Proteins from Quinoa (*Chenopodium quinoa*). *J. Agric. Food Chem*: 1621-1623.
22. Zhao Q, Selomulya C, Xiong H, Chen XD, Ruan X, et al. (2012) Comparison of functional and structural properties of native and industrial process-modified proteins from long-grain indica rice. *J. Cereal Sci* 56: 568-575.
23. Shih, FF & Daigle, & KW (2000) Comparison of functional and structural properties of native and industrial process-modified proteins from long-grain indica rice. *J. Am. Oil Chem. Soc* 77: 885-889.
24. Hamada JS (2000) Characterization and functional properties of rice bran proteins modified by commercial exoproteases and endoproteases. *J. Food Sci* 65: 305-310.
25. Beyer K, Bardina L, Grishina G, & Sampson HA (2002) Identification of sesame seed allergens by 2-dimensional proteomics and Edman sequencing: seed storage proteins as common food allergens. *J. Allergy Clin Immunol* 110: 154-159.
26. Ahmadian-Kouchaksaraei Z, Varidi M, Varidi MJ, & Pourazarang H (2014) Influence of processing conditions on the physicochemical and sensory properties of sesame milk: A novel nutritional beverage. *LWT-Food Sci. Technol* 57: 299-305.