Food Research International 139 (2021) 109919

Contents lists available at ScienceDirect

Food Research International

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

Investigation of fennel protein extracts by shot-gun Fourier transform ion cyclotron resonance mass spectrometry

Maria Teresa Melfi [a], Basem Kanawati [b,*], Philippe Schmitt-Kopplin [b,c], Luigi Macchia [d], Diego Centonze [a], Donatella Nardiello [a,*]

[a] Dipartimento di Scienze Agrarie, degli Alimenti e dell'Ambiente, Università degli Studi di Foggia, Via Napoli, 25, 71122 Foggia, Italy
[b] Research Unit Analytical BioGeoChemistry (BGC), Helmholtz Zentrum München, Ingolstaedter Landstrasse, 85764 Neuherberg, Germany
[c] Chair of Analytical Food Chemistry, Technical University of Munich, Alte Akademie 10, D-85354 Freising, Germany
[d] Dipartimento dell'Emergenza e dei Trapianti di Organi, Sezione di Allergologia ed Immunologia Clinica, Università degli Studi di Bari, Piazza G. Cesare, 11, 70124 Bari, Italy

ARTICLE INFO

Keywords:
Direct-infusion FT-ICR-MS
Fennel proteins
Peptide mass fingerprint
Allergenic proteins
Shot-gun analysis

ABSTRACT

A rapid shot-gun method by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) is proposed for the characterization of fennel proteins. After enzymatic digestion with trypsin, few microliters of extract were analyzed by direct infusion in positive ion mode. A custom-made non-redundant fennel-specific proteome database was derived from the well-known NCBI database; additional proteins belonging to recognized allergenic sources (celery, carrot, parsley, birch, and mugwort) were also included in our database, since patients hypersensitive to these plants could also suffer from fennel allergy. The peptide sequence of each protein from that derived list was theoretically sequenced to produce calculated *m/z* lists of possible *m/z* ions after tryptic digestions. Then, by using a home-made Matlab algorithm, those lists were matched with the experimental FT-ICR mass spectrum of the fennel peptide mixture. Finally, Peptide Mass Fingerprint searches confirmed the presence of the matched proteins inside the fennel extract with a total of 70 proteins (61 fennel specific and 9 allergenic proteins).

1. Introduction

Fennel (*Foeniculum vulgare Mill.*) is a member of the *Apiaceae* (formerly called *Umbelliferae*) family, a large group of plants encompassing approximately 300 genera and more than 3000 species, also including some important allergenic plants, such as carrot and celery. Based on scientific evaluation and its use in traditional medicine, fennel emerged as a good source of medicinal products for research, proving noteworthy in the field of pharmaceutical biology, as well as in the research and development for new drugs (Rahimi & Ardekani, 2013). Indeed, several pharmacological properties, both in vivo and in vitro, have been demonstrated such as anti-microbial, anti-viral, anti-inflammatory and anti-mutagenic activities (Badgujar, Patel, & Bandivdekar, 2014). Fennel is usually consumed as seeds in Northern Europe, while in the Mediterranean area the plant is often consumed fresh. Proteins and fat are the less abundant macronutrients; proteins vary between 1.08 g/100 g in stems and 1.37 g/100 g in inflorescences that also reveal the highest fat content (1.28 g/100 g) among all the fennel parts. On the basis of the proximate analysis, it can be calculated that a fresh portion of 100 g of fennel yields, on average, 94 kcal. The highest values are obtained for inflorescences, while leaves and stems give the lowest energetic contribution (Barros, Carvalho, & Ferreira, 2010). The most abundant fatty acid in shoots, stems, and inflorescences is linoleic acid (C18:2), followed by α-linolenic (C18:3) and palmitic (C16:0) acids. Therefore, a diet rich in fennel could bring potential health benefits due to their valuable nutritional composition in essential fatty acids. In spite of the fennel nutritional composition is well known, a few studies have been described dealing with the proteomic analysis and the identification of fennel allergens in the literature (Asero, Mistrello, Roncarolo, & Amato, 2007; Borghesan et al., 2013; Pastorello et al., 2013). It has been reported that allergic patients to *Apiaceae* spices and birch and/or mugwort pollens could also suffer from allergy to fennel, due to the presence of homologues proteins (Borghesan et al., 2013; Jensen-Jarolim et al., 1997). On the other hand, for an accurate understanding of the molecular basis underlying the cross-sensitization of patients with allergy, further and deeper proteomic investigations are

* Corresponding authors.
E-mail addresses: basem.kanawati@helmholtz-muenchen.de (B. Kanawati), donatella.nardiello@unifg.it (D. Nardiello).

https://doi.org/10.1016/j.foodres.2020.109919
Received 19 August 2020; Received in revised form 3 November 2020; Accepted 21 November 2020
Available online 30 November 2020
0963-9969/© 2020 Elsevier Ltd. All rights reserved.

still needed.

In recent years, proteomic analyses have played an important role in analytical and food chemistry; in particular mass spectrometry (MS) based methods have been suggested as confirmatory tools for an accurate protein identification in the field of food quality and safety (Chen, 2008). Emphasis is placed on food processing, in the determination of possible contaminants like bacteria and fungi, and in allergen detection (Piras, Roncada, Rodrigues, Bonizzi, & Soggiu, 2016). The recent development of fast scanning high-resolution hybrid mass spectrometers, such as quadrupole/orbitrap instruments, has opened new avenues to perform targeted proteomic measurements based on LC-MS/MS analysis (Bourmaud, Gallien, & Domon, 2016; Lesur & Domon, 2015). However, ultra-high resolution mass analyzers such as ICR and Orbitrap play a crucial role in deciphering the most peptide components in digested protein samples, due to their high mass accuracy, which enable accurate mass determinations and thus enabling peptide mass fingerprinting as implemented in Mascot. Thus, the bottom-up approach remains the workhorse for protein characterization, among different proteomic strategies. Nevertheless, it is associated to a greatly increased complexity of the generated peptide mixture after protein digestion, requiring highly sensitive and efficient methods that can lead to correct identifications. A prominent technology for high throughput analysis is Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry, providing the highest resolving power and mass measurement accuracy (Marshall, Hendrickson, & Jackson, 1998). Moreover, ICR technology delivers highly sensitive mass spectra due to the ultra-high vacuum conditions, which produce large induced image electrical current on the detector plates of the ICR cell. The large dynamic range and the unmatched sensitivity of FT-ICR-MS currently provides the highest quality data for molecular identifications (Page, Masselon, & Smith, 2004). Nevertheless, to date, the advantages of FT-ICR-MS techniques have not been fully exploited for proteome investigations and compared to the exponentially increasing number of proteomics literature works, performed by TOF or Orbitrap systems, a lower number of applications have been reported (Anderson et al., 2017; Bergquist, Palmblad, Wetterhall, Håkansson, & Markides, 2002; Floris et al., 2018; Liu et al., 2017; Quenzer, Emmett, Hendrickson, Kelly, & Marshall, 2001; Tucholski et al., 2019; van der Burgt et al., 2019; Witt, Fuchser, & Baykut, 2003; Zhang et al., 2019).

The first way of using mass spectrometry data for proteomic determinations is known as peptide mass fingerprint analysis (PMF), based on the mass measurement of enzymatically digested proteins and the comparison of such data with theoretical fingerprints from protein databases (Mann, Højrup, & Roepstorff, 1993; Yates, 1998). During the last decades, PMF database search has become the preferred method of choice for high throughput protein identifications, especially for those protein classes, where MS/MS database data content is not complete or even insufficient for achieving unequivocal peptide characterizations. On the other hand, the risk of false positives is greater with PMF than with other MS approaches such as MS/MS ion search (Dodds, Clowers, Hagerman, & Lebrilla, 2008). When using PMF, false discovery is closely related to mass measurement error; inaccuracies in the measurement of peptide *m/z* values have the unavoidable consequence of leading to errors in sequence associations at both the peptide and protein levels. As for any analytical process, uncertainty in a measurement translates into uncertainty in the derived results from that measurement; Therefore, an effective way to minimize the possibility of incorrect sequence assignments is the use of accurate mass measurement (i.e. with mass error <5 ppm) and setting stringent mass tolerance limits in the database query. Therefore, reliable protein identifications by Peptide Mass Fingerprint can be achieved only if the high measurement accuracy of the FT-ICR mass analyzer is used, allowing to assign detected *m/z*-values to the top-scoring protein sequences and decreasing the score of the other random protein matches (Horn, Peters, Klock, Meyers, & Brock, 2004).

In this work, a rapid, non-targeted, shot-gun FT-ICR method is described for the protein characterization in fennel samples. Few microliters of the extract have been digested with trypsin and analyzed by direct-infusion FT-ICR-MS, without recurring to any fractionation or purification processes. The instrumental analysis takes a few minutes and most of the time is required for the bioinformatic analysis and database searching. Before MS analysis, a custom-made database was generated including all the fennel proteins reported in the NCBI database under the species *Foeniculum Vulgare*. Moreover, our database was integrated by adding other well-known allergenic proteins not fennel specific, but derived from other plants (celery, carrot, parsley, birch, mugwort), responsible of the so-called spice-mugwort-allergy syndrome. Indeed, the sequence similarity between proteins of different sources could be at the base of the multiple sensitization of allergic patients to several vegetable organisms. In order to obtain confident peptide sequence assignments, a data pre-processing was crucial and was performed by comparing the experimental mass dataset with the theoretical values coming from the in silico digestion of all the proteins included in our custom-made database (both fennel specific and the recognized allergenic proteins from other plants). Then, the protein characterization was performed by database searching in PMF mode of the matched experimental mass peak lists. Although fennel has attracted attention as a medicinal plant with an enormous amount of health benefits, it has been recently identified as an allergenic source, especially in the Mediterranean area. Therefore, this work represents the starting point for allergen characterization in fennel samples, allowing the upgrade of the pattern of allergenic molecules in food products. In addition, even if the proposed strategy was specifically applied to the fennel protein analyses, it should be considered as an innovative and very informative analytical approach that could be also used and transferred to other matrices, such as food products or biological and clinical samples, for a rapid and accurate protein characterization.

2. Experimental section

2.1. Chemicals

Standards of ubiquitin (UBIQ) from bovine erythrocytes (purity grade $\geq$98%, average mass 8565 Da), trypsin from porcine pancreas (proteomics grade, BioReagent, dimethylated) and protease inhibitor cocktail were purchased from Sigma-Aldrich, as well as formic acid ($\geq$98%), sodium dihydrogen phosphate dihydrate, sodium hydroxide, hydrochloric acid, ethylenediaminetetraacetic acid (EDTA), calcium chloride, Trizma® base, Trizma® hydrochloride, 1,4-dithiothreitol, iodoacetamide, and ammonium bicarbonate. Water, methanol, and acetonitrile (LC-MS CHROMASOLV®, $\geq$99.9%) were purchased from Fluka. Ubiquitin standard solutions at a concentration of 1000 mg L^{-1} were prepared in water and stored at -20 °C.

2.2. Protein extraction from fennel

In-house semi-purified 100,000 *xg* supernatant fennel extracts were prepared in duplicate and used for the shot-gun MS analysis. A suitable amount (100 g) of the edible part of fresh *F. Vulgare* (purchased in local supermarkets) was washed properly, minced and homogenated (Heidolph DIAX 900 homogenizer with a Heidolph 10 F probe) for 15 min at 25,000 revolutions per minute (rpm), on ice, in the presence of phosphate buffer solution (PBS) containing Ca^{2+}, Mg^{2+}, EDTA at a final concentration of 2 mM, and 700 μL of plant cell-specific protease inhibitor cocktail composed of 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF), 1,10-phenantroline, pepstatin A, bestatin and *trans*-epoxysuccinyl-Lleucylamido-(4-guanidino)butane (E64), at unknown concentration. The homogenate was then centrifuged at 12,000 *xg*, for 20 min at 4 °C and the supernatant was recovered. This step was twice repeated. Successively, the sample was ultra-centrifuged at 100,000 *xg*, for 2 h, at 4 °C. From an initial amount of 100 g of fresh fennel, 40 mL of 100,000 *xg* supernatant were obtained. The extract was kept at -80 °C, until used. The protein content, determined according to

the colorimetric Bradford method (Bradford, 1976), was 3.5 mg L^{-1}.

2.3. Enzymatic digestion of standard UBIQ and proteins from fennel extract

To a volume of 300 μL of 1000 mg L^{-1} UBIQ standard, 1.1 mg of calcium chloride (as a stabilization agent against thermal and proteolytic degradation, and autolysis phenomena (Papaleo, Fantucci, & De Gioia, 2005)) was added to get a concentration of 20 mM in a final volume of 500 μL. The aqueous substrate protein solution is then buffered at pH 8.5 by adding 100 μL of 200 mM Trizma buffer solution, prepared by mixing Trizma-Base and Trizma-HCl. The in-solution enzymatic digestion was performed by adding 100 μL trypsin (0.1 g L^{-1} in water) at a 1:30 (w:w) ratio. The mixture was incubated at 37 °C for 6 h and then stopped by the addition of 5% formic acid (FA) water solution (5 μL). Final mixtures were stored at −20 °C until the analysis. The in-solution enzymatic digestion of the raw fennel extracts, diluted 1:10 with MilliQ water, was performed in duplicate and adopting for each extract two different protocols:

- Protocol A (6 h); to a volume of 300 μL of extract, calcium chloride (at a final concentration of 20 mM) and 200 mM Trizma buffer solution (100 μL) were added. Then, the enzyme trypsin prepared in water was added at a 1:30 (w:w) ratio. After a 6-h incubation at 37 °C, the digestion was stopped by adding 5 μL of 5% FA.
- Protocol B (18 h), performed following the procedure of Khodadadi et al. with slight modifications (Khodadadi et al., 2017): 400 μL of methanol, 100 μL of chloroform and 300 μL of water were added to 100 μL of fennel protein extract and mixed thoroughly. After centrifugation at 15,000 rpm for 15 min, the upper aqueous phase was discarded, whereas 300 μL of methanol was added slowly to the lower phase. Then, the extract was further centrifuged at 15,000 rpm for 15 min. After drying, the resulting pellet was resuspended in 50 mM NH_4HCO_3 to reach a pH of 8.5. After reduction with 50 mM dithiothreitol for 60 min at 56 °C, and alkylation with 50 mM iodoacetamide for 60 min at 37 °C in the dark, the enzymatic digestion was performed with trypsin at a 1:100 enzyme/protein concentration for 18 h of incubation at 37 °C. The resulting peptides mixtures were acidified with 5% FA (5 μL) and centrifuged at 15,000 rpm for 15 min.

Before ESI-FT-ICR mass spectrometry analyses, the peptide mixtures obtained by both protocols were diluted 1:10 in a mixture of acetonitrile/water (70:30, v/v) containing 0.2% FA, which proved to deliver the highest sensitivity possible, when compared with MS spectra of methanolic peptide samples.

2.4. ESI-FT-ICR mass spectrometry analyses

Ultra-high resolution ESI(+) mass spectra were acquired on a SolariX ion cyclotron resonance Fourier transform mass spectrometer (Bruker Daltonics GmbH, Bremen, DE) equipped with an Apollo II ESI source (Bruker Daltonics GmbH, Bremen, DE) and a 12 T superconducting magnet (Magnex Scientific, Yarnton, UK). The Bruker Solarix contains a linear ion beam guide, which consists of a splitted quadrupole (directly connected to an ion funnel), a quadrupole, which serves as a mass selection filter and a hexapole, which represents a relatively high pressure collision cell, for MS/MS experiments, which are assisted with argon atoms (collision gas). Samples were injected with a flow rate of 2 μL min^{-1}. The MS was calibrated with a 5 mg L^{-1} arginine solution reaching a mass error below 100 ppb and was tuned in order to obtain the highest sensitivity for peptides in the mass/charge (m/z) range of 122–3000 in broadband detection mode. The resolution was on average of R = 400,000 at m/z 400, enabling an excellent signal differentiation on a molecular level. Tryptic peptide mixtures of fennel proteins were analyzed by direct flow injections by the use of Electrospray Ionization double Quadrupole-Fourier Transform-Ion Cyclotron Resonance mass spectrometry (ESI qQ-FT-ICR-MS). A sine mathematic apodization function was applied on the time domain transients and is used to significantly reduce the FT-artifacts side lobes (wiggles), which do not count for the real (chemically relevant) signals in the mass spectra. The sine apodization function was best suited for FT-ICR-MS to date, according to our recent study (Kanawati, Bader, Wanczek, Li, & Schmitt-Kopplin, 2017; Kanawati, 2019). The FT-ICR mass spectra were acquired in the mass range m/z 123–3000. Number of scans: 300 per acquisition. The size of the time domain transient was 1.4 s for the mass spectra of the digested fennel and UBIQ proteins. Technical replicates were performed by multiple analysis (n = 4) of the same sample fennel peptide mixture. MS/MS analysis by Collision-Induced Dissociation (CID) was performed on the most intense ions of the MS-spectrum (scan range 100–3000 m/z; scan number 50; accumulation time 1.8 s; collision energy in the range (10–30 eV). Ion accumulation time in the collision cell was varied from 50 to 300 ms (50, 150, 150 and 300 ms) to examine any possible gain of signal intensity without ion space-charge effect for better mass accuracy. With an accumulation time of 300 ms, decreased signal intensities were observed for all the detected ions of about one order of magnitude. Among the replicate measurements at accumulation times of {50, 100, 150 ms}, a coefficient of variation (CV) percentage on signal intensities lower than 10% was observed.

2.5. Database searching and sequence analysis

Extraction of mass spectra peak-lists, mass annotation and deconvolution were performed by using Data Analysis 4.4 (Bruker Daltonics). The m/z data were de-noised according to a home-made algorithm (Kanawati et al., 2017). The mass spectra were calibrated by the use of the cluster ions of arginine in positive Electrospray ion mode, which range from m/z 175 of the $[M+H]^+$ monomer ion until reaching m/z 1220 heptamer $[7M+H]^+$ cluster ion. A methanolic 5 mg L^{-1} arginine was utilized for FT-ICR-MS calibration purposes. Re-calibration was performed by the use of traces of arginine cluster positive ions to enhance the mass accuracy. A subset of NCBI protein sequences was derived to represent protein sequences found previously in fennel species. Then, a custom-made protein database was generated, as reported in a public repository (Table 01 of (Nardiello, Kanawati, & Melfi, 2020)). Moreover, ten additional proteins, known as allergens found in other spices (such as celery, carrot or parsley), or belonging to other recognized allergenic organisms (such as birch or mugwort pollen) were also considered and included in the database (Table 02 of (Nardiello et al., 2020)). Then, each fennel protein of that NCBI subset, as well as each allergenic protein from other plants, was subjected to simulated tryptic digestion by the use of the ExPaSy peptide mass calculator tool (ExPaSy Bioinformatics Resource Portal, n.d.). The following parameters were set: enzyme trypsin, two allowed missed cleavages; mass range from 0 to unlimited Dalton; cysteines untreated (for the analysis of UBIQ and fennel extract digested according to the protocol A, 6 h) or treated with iodoacetamide (for the fennel extract digested according to the protocol B, 18 h). For each protein, the in silico enzymatic digestion was performed by selecting consecutively the option to save the theoretical peptide masses in form of $[M+H]^+$, $[M+2H]^{2+}$ and $[M+3H]^{3+}$ in order to have a complete and realistic mass list of all the putative m/z ions, which are possible to exist as a result of the use of ESI to be searched in the experimental mass spectrum.

The m/z list of the possibly generated m/z ions, reflecting the tryptic peptides out of each NCBI fennel and allergenic protein, was used for performing a custom Matlab search inside the experimental high-resolution FT-ICR mass spectrum of the isolated fennel protein mixture, which was digested with trypsin. A 5 ppm mass tolerance search was performed and many relevant fennel proteins could be found in this search step.

Finally, in order to perform the protein characterization, the matched experimental mass datasets were submitted to database

searches by using the MASCOT search engine (Matrix Science, London, UK). The PMF searching was performed under the *Viridiplantae* category of the NCBI database. A maximum number of 2 missed cleavages were allowed and a peptide mass tolerance of 2 ppm was set in the error window. No variable and fixed modifications were set for the fennel extract digested by protocol A, whereas the option carbamidomethylation at cysteine residues was checked for protocol B fennel extract. MS/MS datasets were submitted to database searches by Mascot search engine. The data analysis files were used to search entries under the category other mammalia (for the analysis of standard ubiquitin) and *Viridiplantae* (for the analysis of fennel extracts) of NCBI databases, assuming that peptides were monoisotopic and carbamidomethylated at cysteine residues. A maximum number of 2 missed cleavages were allowed.

3. Results and discussion

3.1. Direct infusion FT-ICR analysis of ubiquitin standard and fennel protein extract

Before analyzing fennel extracts, a standard sample of UBIQ protein was tested to set all the experimental parameters for the direct infusion FT-ICR analysis, as well as to establish the database searching parameters and tune the mass error. Consecutive PMF database searches were performed for UBQ standard by setting an error tolerance of 1, 2 ppm and 5 ppm. While by setting 1 ppm, no protein identification was obtained (although, obviously, with no doubts the identity of ubiquitin is certain), both with a mass tolerance of 2 ppm and 5 ppm, the same results were obtained in terms of score, coverage and number of identified peptides. Therefore a mass tolerance of 2 ppm was fixed for all the analyses. The quality of UBIQ mass spectrum (Fig. 1), dominated by a few high-intensity principal signals, was confirmed by the excellent results obtained by PMF Mascot database search, with a protein score of 152 and a percentage coverage of 80%, associated to the identification of 7 main peptides uniformly distributed along the protein sequence. For selected precursor ions, CID analyses were also performed, therefore the most intense ions were isolated for fragmentation. As an example, Fig. 2 shows the MS/MS spectrum of the triply charged precursor ion at *m/z* 596.6480 corresponding to the 16 residue long peptide TITLEVEPSDTIENVK. CID fragmentation yielded 10 inter-amino acidic residue cleavages with the identification of 13 ion fragments of b/y-type. For all the acquired MS/MS spectra the fragmentation efficiency was calculated, as the number of observed b- and y-type fragment ions divided by the theoretical number of fragment ions, e.g. 2 (N-1), where N is the number of residues for a given sequence (Nardiello, Palermo, Natale, Quinto, & Centonze, 2015). Fragmentation percentages ranging from 25% (for *m/z* at 1081.55359) and 43% (for *m/z* at 596.64796 and 894.46827) were obtained, confirming that triply charged positive ions are favored over doubly charged ions in CID fragmentation, due to their higher fragmentation efficiency. The experimental MS/MS datasets were used in the database searching for protein characterization but nevertheless, a poor peptide match was obtained: only two peptides were successfully obtained by MS/MS ion search against the seven peptides observed by PMF. Therefore, this loss of information, observed for a protein standard whose identity is a priori known, suggested us to adopt, for the protein characterization in all the subsequent analyses, the PMF approach using data from MS spectra of tryptic protein digests, taking full advantage of the high resolving power and mass accuracy of the FT-ICR mass analyzer. Nevertheless, to further confirm our hypothesis, in preliminary experiments, as for ubiquitin, also for the fennel extracts, MS/MS analyses were performed and the most intense ions from the MS spectrum were selected and fragmented by CID. Systematic MS/MS experiments were performed by changing the collision energy and monitoring the intensities of the fragment ions. The collisional energy was optimized for individual molecular species to achieve the maximum S/N ratio. A total mascot generic file, obtained by combining the fragment ion peak lists of each precursor ion, was submitted to database MS/MS ions search. A small number of proteins (less than 20, generally ribosomal proteins, translaction factors and RNA polymerase as partial forms) was observed in the final protein view report, associated to other plants (*Arabidopsis thaliana, Zea mais* or *Oryza sativa*) different from *Foeniculum vulgare*, with low protein scores, in the range 59–83. Moreover, the number of the identified peptides was quite low (not higher than 11), with a protein sequence coverage not exceeding a value of 19%. These results were probably due to the fact that when the acquired MS/MS spectra are searched against a sequence database, unless a peptide is unique to one particular protein, there may be some ambiguity to infer which proteins were effectively present in the sample

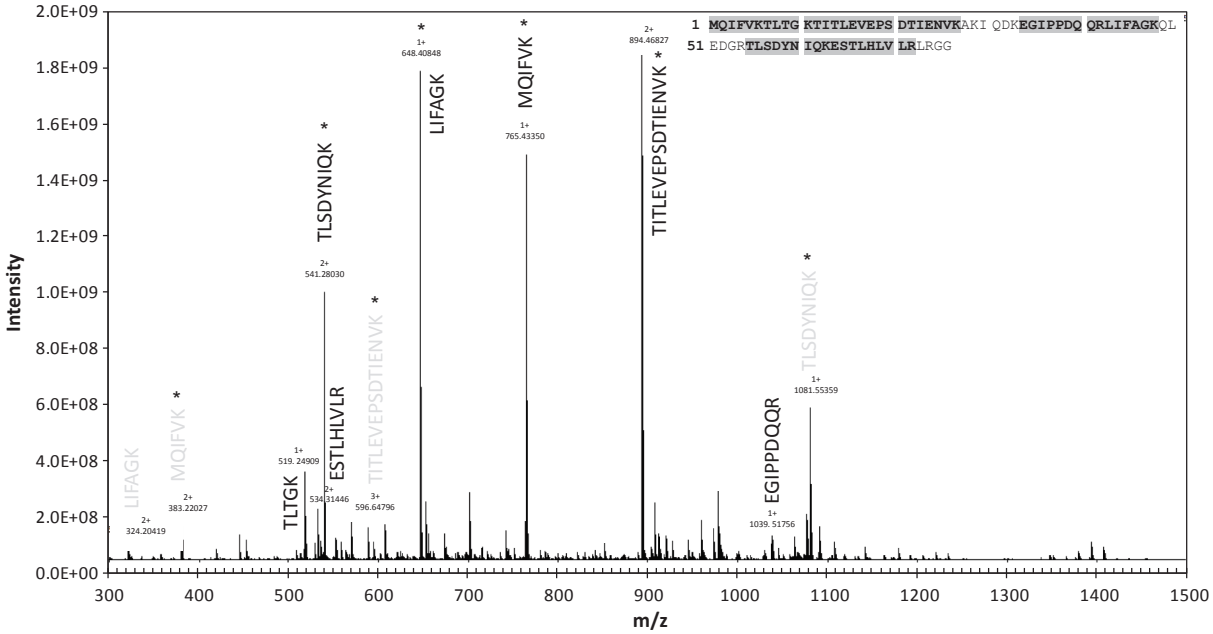

Fig. 1. ESI(+)-FT-ICR mass spectrum of a tryptic mixture (acetonitrile/water (70:30, v/v) + 0.2% FA) of UBIQ standard. Stars indicate the signals of the peptides later subjected to MS/MS analysis.

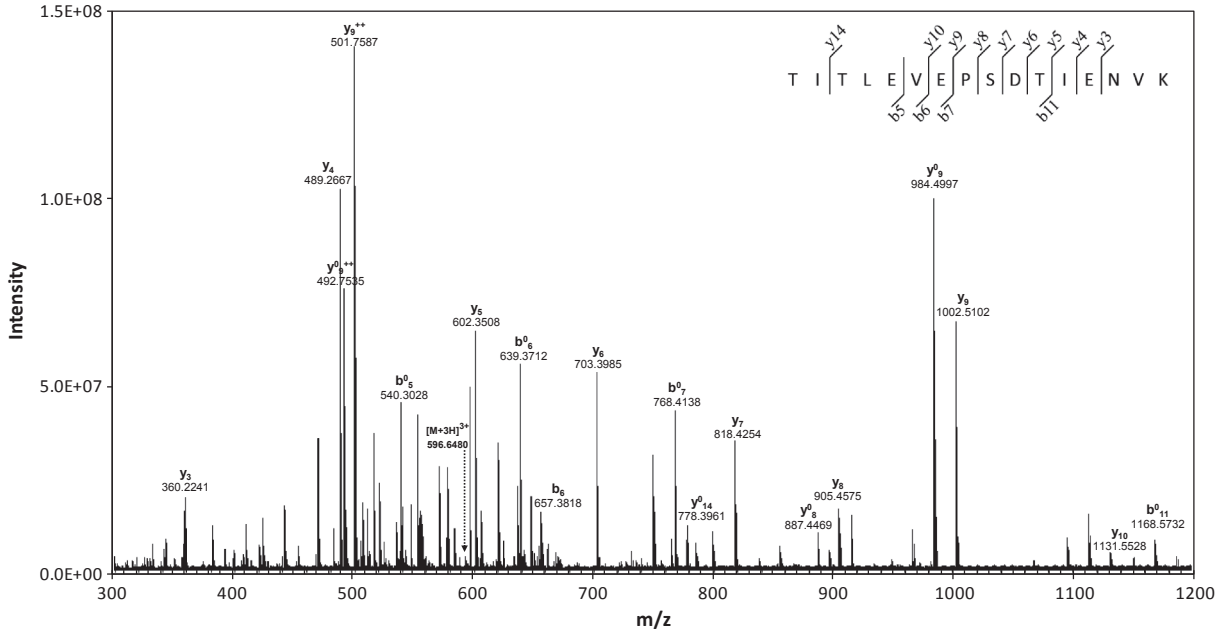

Fig. 2. Collision induced dissociation (CID) mass spectrum of the precursor peptide at *m/z* 596.6480 from UBIQ standard by ESI qQ-FT-ICR-MS. Collisional energy = 10 eV. Ions b^0 and y^0 denote loss of water (−18 Da).

(Cottrell, 2011). As a consequence, the MS/MS database ion searching has the effect of increasing the number of misassigned peptides and decreasing the general benefits of high mass accuracy provided by FT-ICR analyzer. Hence, for a comprehensive investigation of all the proteins present in the whole raw fennel extract, a shot-gun analysis combined with PMF database search was performed on the peptide mixtures derived from the tryptic digestion. Considering the complexity of the fennel extract whose MS spectrum, shown in Fig. 3, is characterized by an elevated number of ion signals exceeding the maximum limit of 1200 entries, set for the free version of the MASCOT database search feature, the MS spectrum was reduced to a peak list. A value of signal-to-noise ratio of 4 combined to an absolute intensity threshold of 10^6 counts was fixed to remove the background noise. This mass list was submitted to PMF MASCOT search, but no protein hit above the acceptance threshold (at a significance level of 0.05) was observed, probably as a result of the information dilution effect in the MASCOT workflow due to database size, complexity, and occurrence of proteins with significant homology. The presence of multiple proteins in the fennel extract and the high number of observed masses in the spectrum significantly increase the likelihood of an incorrect assignment to other proteins in the database (Horn et al., 2004; Karty, Ireland, Brun, & Reilly, 2002). Indeed, in large data sets, there are likely to be several shared mass values, which match to more than one of the proteins in the mixture. In addition, although full details about the Mascot search engine and

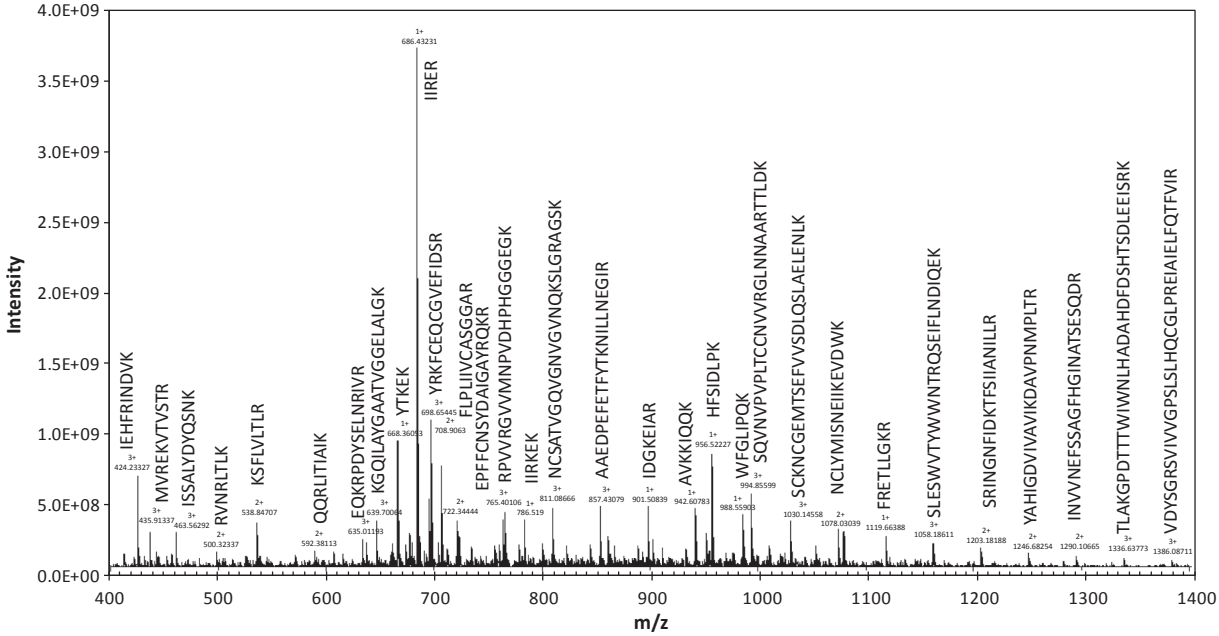

Fig. 3. ESI(+)-FT-ICR mass spectrum of a tryptic mixture (acetonitrile/water (70:30, v/v) + 0.2% FA) of a fennel extract. Sequences matching the in silico peptide digests are shown.

scoring algorithm are not published, the lack of a confident protein identification is presumably due to the fact that the MASCOT works with redundant protein databases (*Viridiplantae* taxonomy against NCBI database contains 9,630,886 sequences - December 2019), thus lowering the probability of obtaining a valid identification above the acceptance threshold. In order to overcome these problems and improve the statistical confidence in the database search results, a data pre-processing was necessary and was hence adopted, as reported below.

3.2. Protein profiling in fennel extracts

Starting from the NCBI protein database under the organism *Foeniculum vulgare*, a custom-made non-redundant proteome database of 92 fennel proteins was generated, as listed in a public repository (Table 01 of (Nardiello et al., 2020)). Indeed, the official, on-line available NCBI database for fennel contains a total of 231 proteins; nevertheless, the most of these proteins (each with its own specific code) are repeated and represent partial forms of the same longer protein, already reported in the list of the protein NCBI database, under a different code. Partial and redundant proteins of the same class were kept out unless single-point variations were observed in the amino acid strings. This is the case for the proteins belonging to the groups of RNA beta polymerases, ribulose-1,5-bisphosphate carboxylase/oxygenases, maturase K and hypothetical chloroplast RF1, present in our subset of fennel database both as entire and partial proteins. In addition to these fennel specific proteins, other allergenic proteins found in other species (such as celery, carrot or parsley), or belonging to other recognized allergenic organisms (such as birch or mugwort pollen) were also included in the final version of our database. Indeed, some publications dealt with the relationship between fennel allergy and birch and mugwort pollen allergy, in the so-called birch-weed or fruit-spice syndrome (Asero et al., 2007; Borghesan et al., 2013; Jensen-Jarolim et al., 1997; Pastorello et al., 2013). Jensen-Jarolim et al. stated that Bet v1 (17 kDa band) and profilin-related allergens (Bet v2, 14 kDa band) could be responsible for allergy to fennel, thus demonstrating the immunological basis of the clinical association between fennel seeds allergy and birch or mugwort pollen allergy (Jensen-Jarolim et al., 1997). Similarly, the major allergens Api g1 and Dau c1, belonging to the *Apiaceae* plant foods celery and carrot, respectively, are Bet v1 homologues (Hoffmann-Sommergruber et al., 1999).

The list of the common allergens, considered in the present work (coming from *Apiaceae* spices or associated with mugwort-birch-pollen-allergy-syndrome), is shown in Table 1. Other details are reported in a public repository (Table 02 of (Nardiello et al., 2020)). For each protein of our custom-made database (both fennel specific and homologues allergens), a complete list of theoretical peptide sequences originating from in silico digestion with trypsin was generated by using the PeptideMass on-line tool from UniProtKB (https://www.uniprot.org/uniprot/) and compared with the experimental molecular masses of the fennel mass spectrum. A filtering criterion, based on a custom-designed MATLAB algorithm, was applied to match the experimental *m/z* values with the in silico enzymatic digestion data. For the mass range from 400 to 1400 *m/z*, the peptide mass error tolerance (Δm) was calculated based on the instrumental resolution values (R = m/Δm) observed in the fennel MS spectrum at different *m/z* values. In the MATLAB algorithm for the mass comparison between experimental and theoretical data, a peptide mass tolerance of 5 ppm was set. The advantage of setting a constant specific mass tolerance error in ppm is the fact that this error will be updated to lower mDa errors when moving to lower *m/z* search values. The reduction of mDa search tolerance mass error is important, especially for low *m/z* values <800, since the mass resolving power of FT-ICR technique increases in a mathematical power function when moving to lower *m/z* values, thus enabling better matching results even with lower mDa errors. This, of course, helps in reducing the number of false positives out of the database search. The setup of a 5 ppm mass tolerance error allows striking a balance between greater information capture and a reduced number of incorrect sequence assignments, minimizing the risk of false-positive results.

Table 1
List of common allergens in *Apiaceae* spices and proteins associated to mugwort-birch-pollen-allergy-syndrome, included in the custom-made protein database.

Protein Name	Organism (vulgar name)	Protein Code	MW (Da)	Length (No. residues)
Allergen Api g 5	*Apium graveolens* (celery)	P81943	9441	86
Profilin	*Apium graveolens* (celery)	Q9XF37	14,276	134
Major allergen Api g 1, isoallergen 1	*Apium graveolens* (celery)	P49372	16,321	154
Major allergen Api g 1, isoallergen 2	*Apium graveolens* (celery)	P92918	17,090	159
Cofactor-independent phosphoglyceromutase	*Apium graveolens* (celery)	Q9SDL3	61,125	559
Pathogenesis-related protein 1	*Petroselinum crispum* (parsley)	Q40795	16,356	155
Non-specific lipid-transfer protein	*Dacus carota* (wild carrot)	P27631	12,504	120
Major pollen allergen Bet v 1-A	*Betula pendula* (silver birch)	P15494	17,571	160
Major pollen allergen Bet v 1-B	*Betula pendula* (silver birch)	P45431	17,537	160
Major pollen allergen Art v 1	*Artemisia vulgaris* (mugwort)	Q84ZX5	13,404	132

In the public repository (Figure 02 of (Nardiello et al., 2020)), the expanded views of the fennel MS spectrum in the mass segments 600–700 and 700–800 amu, also showing the amino acidic sequences matching the in silico peptide digests, are reported. Other mass segments: 800–1000 and 1000–1200 amu are shown in Figure 03 of (Nardiello et al., 2020).

Finally, in order to corroborate the method reliability and perform the protein characterization in the fennel extracts, the matched experimental mass datasets with the theoretically digested NCBI fennel proteins, obtained by the Matlab algorithm, were submitted to PMF Mascot search against the full redundant NCBI database. Before database searching, the matched multi-charged *m/z* values deriving from the Matlab processing results have been converted to $[M+H]^+$ ions (considering the charge state of the corresponding in silico peptides from Expasy Peptide Mass tool, according to the following equation: $[M+H]^+ = z*m/z + (1-z)$ proton mass, where z is the charge state), since this is the query format required for PMF search. A schematic step-by-step workflow for database searching and bioinformatic analysis (totally performed by free-access tools without recurring to dedicated software) is reported in Fig. 4. The submission of the mass data to MASCOT search engine returned fennel proteins as top-scoring hits; other potential matching proteins belonging to a different class were totally absent in most of the cases; only in 8 protein view results on a total of 61 proteins, incorrectly assigned proteins were also observed as a second marginal protein hit, associated to very low scoring levels. Since the Mascot score reflects the probability that the match between the observed molecular masses and the digested database entry is a random event (score = -LogP; Mascot also reports a score threshold based on the selected significance level, by default 0.05, then a protein hit is statistically significant if its score is above the threshold), it is possible to calculate the score difference between the highest and the second-highest protein hit (Δscore) to estimate the accuracy in protein identifications (Dodds et al., 2008). The higher Δscore is, the more accurate becomes the identification associated with the first protein hit

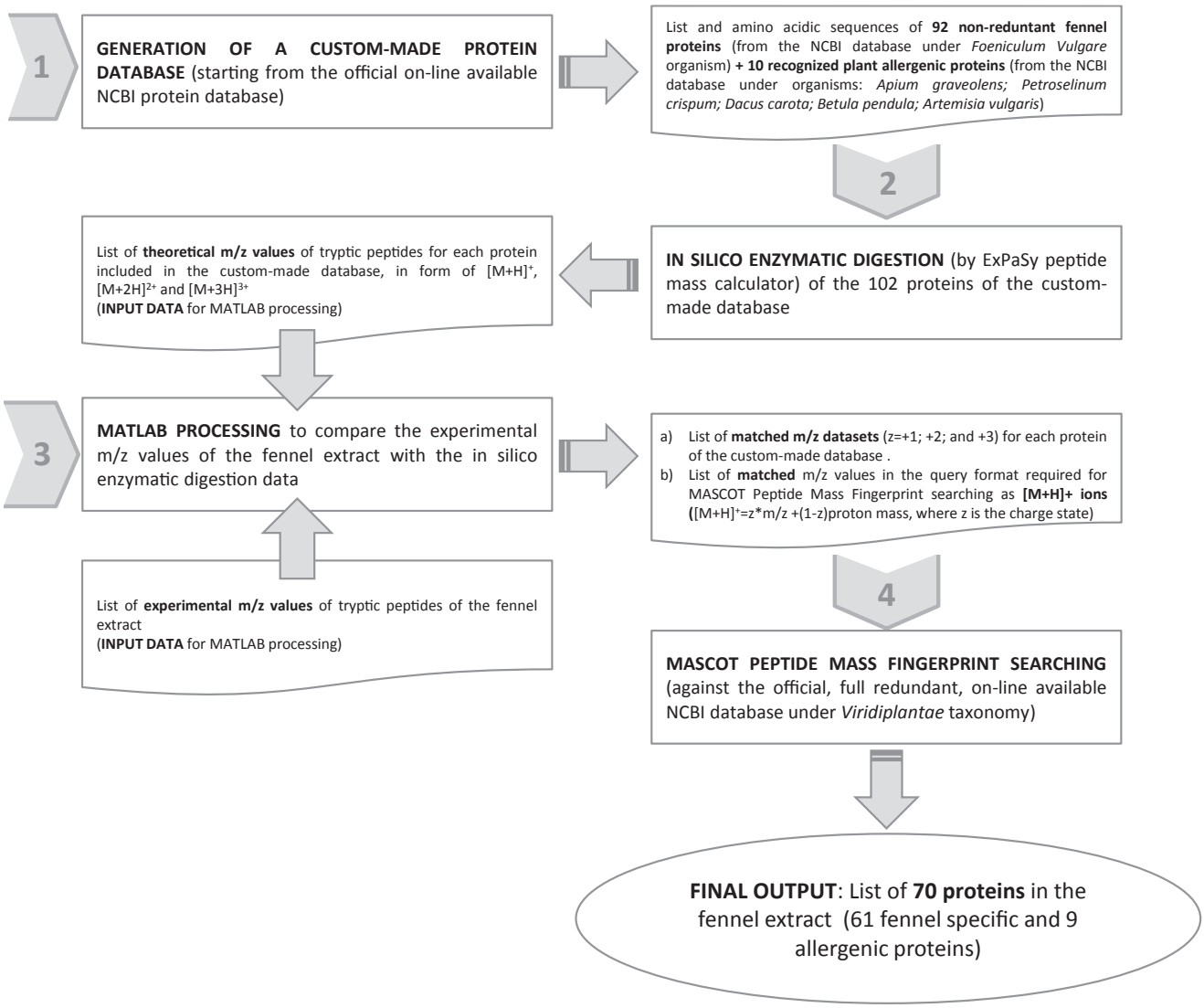

Fig. 4. Schematic step-by-step workflow for database searching and bioinformatic analysis for fennel protein characterization.

compared to the second-highest score protein. Then, the ratio of the probabilities that the observed match is a random event can be calculated by the following equation (Dodds et al., 2008): $P_1/P_2 = 10^{-(\Delta score/10)}$. For the fennel extract MS analysis, the difference between the first ranking protein match and the second-highest-ranking match ranged from 50 to 700, hence the corresponding ratio of the probabilities ranged from 5 to 70 orders of magnitude, confirming the accuracy of the first ranking identification and advantages of the high mass accuracy measurements in proteomic determinations. Moreover, an automatic Decoy search was also performed for an additional evaluation of accuracy in sequence assignments. Then, the database search was repeated using identical search parameters, against a randomized and reversed sequence database. Although this approach cannot be used to get a false positive rate when the number of matches is too small to give an accurate estimate, it can be informative to see the result of repeating a PMF search against a decoy database to validate the peptide identification results and filter out false positive proteins (Huang, Wang, Yu, & He, 2012). For all the proteins found in the fennel extracts, the number of Decoy matches was always null, and the highest scoring protein hit associated with the Decoy search was always well below the significance threshold, with protein scores not exceeding a value of 37. Therefore, the absence of random matches further confirmed the high accuracy in the dataset of the peptide profiling by PMF.

Our results in fennel protein characterization showed that the 18-h protocol B led us to a higher number of proteins (70 against 53 obtained by the 6-h protocol A), demonstrating its major efficiency in the enzymatic process. By comparing the replicate analyses obtained by each digestion protocol, the same number of proteins was obtained (i.e. 53 for protocol A, replicate 1 and 2, and 70 for protocol B, replicate 1 and 2). For some proteins, only slight differences were observed in terms of number of identified peptides between the digestion replicates, anyway with no effects to score and sequence coverage. Indeed, the peptide lists also include overlapping peptides throughout the protein amino acidic sequence, that is peptides that partially or totally cover the same amino acidic sequence already present in the other identified peptides of the same protein. Therefore, in our replicate analyses, an observed reduced number of peptides (sometimes found for a few fennel proteins), did not determine a lowering in the protein coverage. A complete list of all the characterized proteins by multiple analyses of the same fennel sample (according to the digestion protocol B) is displayed in Table 2. A protein profile with molecular weights that range between 4.5 (ribosomal protein L36) and 250 kDa (hypothetical protein RF2) was observed. Good results were observed in terms of MASCOT scores in the range 84–847, with a sequence coverage higher than 14% and a number of peptides ranging from 3 to 99 for individual proteins. For all the observed proteins, the mass errors evaluated as root mean square (rms)

Table 2
List of proteins in the fennel (*Foeniculum vulgare*) extract by direct injection ESI-FT-ICR-MS analysis and Peptide Mass Fingerprint Mascot database searching.

Entry Name	Protein Code	Sequence Coverage %	MASCOT Score	No. Matched Peptides	MW (Da)	Length(No. residues)
Ribosomal protein L36 (plastid)	YP_009235913.1	89	222	11	4460	37
Ribosomal protein L32 (chloroplast)	YP_009235928.1	92	235	9	5997	53
Ribosomal protein L33 (chloroplast)	YP_009235900.1	87	146	7	7687	66
Photosystem I	YP_009235931	70	118	5	8939	81
Ribosomal protein S16 (chloroplast)	YP_009235861.1	71	163	8	9124	78
Translation initiation factor 1 (chloroplast)	YP_009235914.1	64	190	8	9181	77
Ribosomal protein S19 (chloroplast)	YP_009235920.1	61	172	7	10,545	92
Ribosomal protein L23 (chloroplast)	YP_009235922.1	67	195	6	10,724	93
Ribosomal protein S15 (chloroplast)	YP_009235937.1	72	181	10	10,796	90
Ribosomal protein S14 (chloroplast)	YP_009235877.1	80	215	10	11,750	100
Ribosomal protein S18 (chloroplast)	YP_009235901.1	82	218	13	11,986	101
RNA polymerase beta subunit, partial (chloroplast)	AFK09970.1	31	134	5	12,859	117
Ribosomal protein L14 (chloroplast)	YP_009235916.1	85	203	9	13,511	122
Ribosomal protein S12 (chloroplast)	YP_009235858.1	55	139	6	13,738	123
RNA polymerase beta' subunit, partial (chloroplast)	AFK10055.1	75	200	10	13,824	123
Ribosomal protein S11 (chloroplast)	YP_009235912.1	49	147	8	14,972	138
Ribosomal protein L16 (chloroplast)	YP_009235917.1	60	194	10	15,289	135
ATP synthase CF1 epsilon subunit (chloroplast)	YP_009235885.1	21	86	4	15,292	140
Ribosomal protein L20 (chloroplast)	YP_009235902.1	70	200	11	15,295	128
Ribosomal protein S8 (chloroplast)	YP_009235915.1	56	172	9	15,829	134
Ribosomal protein S7 (chloroplast)	YP_009235925.1	59	239	13	17,357	155
Cytochrome b6/f complex subunit IV (chloroplast)	YP_009235910.1	84	90	3	17,459	160
NADH-plastoquinone oxidoreductase subunit J (chloroplast)	YP_009235882.1	41	165	8	18,605	158
NADH-plastoquinone oxidoreductase subunit I (chloroplast)	YP_009235934.1	40	113	6	19,567	167
Photosystem I assembly protein Ycf3 (chloroplast)	YP_009235880.1	42	102	6	19,600	168
RNA polymerase C, partial (chloroplast)	ACB88304.1	54	214	12	19,693	176
ATP synthase CF0 subunit I (chloroplast)	YP_009235865.1	59	212	11	20,527	181
Homogentisate geranylgeranyl transferase, partial	ADG26669.1	50	106	5	20,569	186
Photosystem I assembly protein ycf4 (chloroplast)	YP_009235890.1	54	137	7	21,194	184
CLP protease proteolytic subunit (chloroplast)	YP_009235904.1	29	84	4	22,188	197
Ribosomal protein S4 (chloroplast)	YP_009235881.1	55	290	14	23,341	201
Ribosomal protein S3 (chloroplast)	YP_009235918.1	70	269	16	25,095	218
NADH-plastoquinone oxidoreductase subunit K (chloroplast)	YP_009235883.1	53	227	13	25,448	225
Phenylalanine ammonia-lyase, partial	QDG10107.1	43	270	14	25,740	232
Ribosomal protein S2 (chloroplast)	YP_009235868.1	38	149	10	26,874	236
Chloroplast envelope membrane protein (chloroplast)	YP_009235891.1	37	124	6	26,989	229
Maturase K, partial (chloroplast)	ACB88392.1	63	180	11	27,760	238
Ribosomal protein L2 (chloroplast)	YP_009235921.1	42	157	10	30,100	274
Cytochrome *f* (chloroplast)	YP_009235892.1	48	264	14	35,170	320
Delta12-fatty acid acetylenase, partial	AAO38034.1	20	94	6	37,861	324
RNA polymerase alpha subunit (chloroplast)	YP_009235911.1	50	236	15	38,479	336
Photosystem II protein D2 (chloroplast)	YP_009235874.1	36	140	7	39,550	353
NADH-plastoquinone oxidoreductase subunit 1 (chloroplast)	YP_009235935.1	29	150	8	40,317	363
NADH-plastoquinone oxidoreductase subunit 7 (chloroplast)	YP_009235936.1	69	362	18	45,734	395
Photosystem II CP43 chlorophyll apoprotein (chloroplast)	YP_009235875.1	26	158	9	51,879	473
Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit, partial (chloroplast)	YP_009235887.1	40	224	17	52,596	475
ATP synthase CF1 beta subunit (chloroplast)	YP_009235886.1	64	333	19	53,589	498
Acetyl-CoA carboxylase carboxyltransferase beta subunit (chloroplast)	YP_009235888.1	43	163	11	55,333	489
ATP synthase CF1 alpha subunit (chloroplast)	YP_009235864.1	49	269	15	55,426	507
Photosystem II CP47 chlorophyll apoprotein (chloroplast)	YP_009235905.1	32	228	14	56,004	508
Maturase K (chloroplast)	YP_009235860.1	52	328	25	61,157	518
Hypothetical chloroplast RF1 (chloroplast)	YP_009235926.1	49	384	27	74,606	635
RNA polymerase beta (chloroplast)	YP_009235870.1	62	666	48	78,281	678
Photosystem I P700 apoprotein A2 (chloroplast)	YP_009235878.1	24	171	11	82,345	734
Photosystem I P700 apoprotein A1 (chloroplast)	YP_009235879.1	44	270	17	83,158	750
NADH-plastoquinone oxidoreductase subunit 5 (chloroplast)	YP_009235927.1	14	107	9	85,465	750
RNA polymerase beta subunit (chloroplast)	YP_009235871.1	52	516	44	120,866	1070
Nuclear matrix constituent protein 1-like	BAF64423.1	53	737	79	128,731	1119
RNA polymerase beta' subunit (chloroplast)	YP_009235869.1	51	693	65	157,541	1386
Hypothetical chloroplast RF1 (chloroplast)	YP_009235938.1	54	847	99	217,228	1817
Hypothetical chloroplast RF2 (chloroplast)	YP_009235923.1	47	717	95	247,783	2119

did not exceed 3 ppm and for the corresponding matching peptides, a mean mass error of 2.4 ± 1.4 ppm was observed, demonstrating a high accuracy grade and confidence in sequence assignments. In a public repository, the primary structures of all the proteins found in the fennel extracts are shown (Figure 01 of (Nardiello et al., 2020)). Amino acid sequences which correspond to the identified peptides are highlighted in bold red. At least a sequence tag of four amino acid lengths was considered to enable unambiguous protein identification from a sequence database. Details about all annotated ions with their corresponding peptide sequences are provided in the Table 03 of (Nardiello et al., 2020).

As far as the allergenic profile is concerned, nine homologues proteins among the total ten investigated allergens were found in the fennel extract (see Table 3), suggesting their active role in the spice-mugwort-birch-pollen-allergy-syndrome. Presumably, these proteins are involved in the allergic reaction to plant-derived foodstuff, due to their structural

Table 3
List of the allergenic homologues proteins found in the fennel extract by direct injection ESI-FT-ICR-MS analysis and Peptide Mass Fingerprint Mascot database searching.

Entry Name (*Organism*)	Protein Code	Sequence Coverage %	MASCOT Score	No. Identified Peptides
Allergen Api g 5 (*Apium graveolens*)	P81943	51	88	4
Major allergen Api g 1, isoallergen 1 (*Apium graveolens*)	P49372	39	136	6
Major allergen Api g 1, isoallergen 2 (*Apium graveolens*)	P92918	49	133	6
Cofactor-independent phosphoglyceromutase (*Apium graveolens*)	Q9SDL3	43	291	19
Pathogenesis-related protein 1 (*Petroselinum crispum*)	Q40795	35	103	5
Non-specific lipid-transfer protein (*Dacus carota*)	P27631	49	84	4
Major pollen allergen Bet v 1-A (*Betula pendula*)	P15494	69	158	7
Major pollen allergen Bet v 1-B (*Betula pendula*)	P45431	62	163	8
Major pollen allergen Art v 1 (*Artemisia vulgaris*)	Q84ZX5	41	126	7

similarities to the recognized allergenic proteins from different sources. Thus, our outcomes allow contributing to provide an explanation for a number of clinically observed cross-reactivities in type I allergy consequent to the presence of similar and homologues proteins between mugwort pollen, birch pollen, celeriac and spices of the *Apiaceae* family. Moreover, the coverage percentages observed for the allergenic proteins in our fennel extract can be also used to give an estimation of the protein homology grade of different organisms. Indeed, the structural similarity among proteins of fennel and different species will be at least equal to the sequence coverage obtained by PMF database search. As can be noted from Table 3, for some of these allergenic proteins, the homology grade seems to be very high (this is the case for the major pollen allergen Bet v 1-A, from *Betula pendula,* see also the public repository, Figure 04 of (Nardiello et al., 2020)), supporting the hypothesis of a common molecular basis at the origin of the cross-sensitization of patients with spice-pollen allergy.

4. Conclusion

An untargeted shot-gun approach by FT-ICR-MS is presented for the ultra-high-resolution analysis of fennel proteins. A direct infusion electrospray FT-ICR-MS analysis of the fennel extract samples allows obtaining unambiguous peptide sequence assignments. Successful identifications were obtained by using ultra-high-resolution techniques and proper algorithms capable of handling the thousands of signals generated by such analytical platforms. The development of a custom-made fennel protein database allowed to overcome the limits of the official NCBI protein database in terms of redundancy and incompleteness, that could represent a serious problem in protein identification process by database searching. A data pre-processing is proposed to compare the experimental mass peak list with theoretical data derived from the in silico enzymatic digestion. Peptide mass fingerprinting could be further improved by moving from a redundant to a non-redundant sub-database, which specifically addresses fennel specific and other allergenic proteins. Furthermore, the Mascot PMF search improved the chance of achieving unambiguous protein identifications, especially taking advantage of the very high mass accuracy provided by FT-ICR-MS technique. We find that the described strategy (direct-infusion FT-ICR-MS, peak list extraction, production of a sub-database of non-redundant protein entries specific for fennel proteins and recognized allergenic proteins, calculation of tryptic digestions and matching the calculated post-digestion peptides with the experimental high-resolution FT-ICR mass spectrum of tryptic fennel protein digestion mixture, and finally, the subsequent Mascot database searching in PMF mode) represents a very informative approach for a rapid and accurate protein characterization in fennel extracts.

CRediT authorship contribution statement

M.T. Melfi and B. Kanawati are the co - first authors of the study. Study concept and design: B. Kanawati and D. Nardiello. Acquisition of data: M.T. Melfi and B. Kanawati. Analysis and interpretation of data: B. Kanawati, D. Nardiello, and M.T. Melfi. Drafting of the manuscript: D. Nardiello, B. Kanawati and M.T. Melfi. Clinical studies on hypersensitive patients with fennel allergy: L. Macchia. Critical revision of the manuscript for important intellectual content: P. Schmitt-Kopplin and D. Centonze.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is part of the Ph.D. research project of Maria Teresa Melfi titled “The Challenge of Proteomics in the field of Food Safety: Allergen detection”, which is (partly) financed by the Italian Ministry of Education, University and Research (MIUR). M.T. Melfi and B. Kanawati are the co-first authors of the study. All the authors read and approved the final manuscript. Study concept and design: B. Kanawati and D. Nardiello. Acquisition of data: M.T. Melfi and B. Kanawati. Analysis and interpretation of data: B. Kanawati, D. Nardiello, and M.T. Melfi. Drafting of the manuscript: D. Nardiello, B. Kanawati and M.T. Melfi. Clinical studies on hypersensitive patients with fennel allergy: L. Macchia. Critical revision of the manuscript for important intellectual content: P. Schmitt-Kopplin and D. Centonze.

References

Anderson, L. C., Dehart, C. J., Kaiser, N. K., Fellers, R. T., Smith, D. F., Greer, J. B., Leduc, R. D., Blakney, G. T., Thomas, P. M., Kelleher, N. L., & Hendrickson, C. L. (2017). Identification and characterization of human proteoforms by top-down LC-21 Tesla FT-ICR mass spectrometry. *Journal of Proteome Research, 16*(2), 1087–1096. https://doi.org/10.1021/acs.jproteome.6b00696. Scopus.

Asero, R., Mistrello, G., Roncarolo, D., & Amato, S. (2007). Detection of some safe plant-derived foods for LTP-allergic patients. *International Archives of Allergy and Immunology, 144*(1), 57–63. https://doi.org/10.1159/000102615. Scopus.

Badgujar, S. B., Patel, V. V., & Bandivdekar, A. H. (2014). Foeniculum vulgare Mill: a review of its botany, phytochemistry, pharmacology, contemporary application, and toxicology. *Biomed Research International, 2014*. https://doi.org/10.1155/2014/842674. Scopus.

Barros, L., Carvalho, A. M., & Ferreira, I. C. F. R. (2010). The nutritional composition of fennel (Foeniculum vulgare): Shoots, leaves, stems and inflorescences. *LWT - Food Science and Technology, 43*(5), 814–818. https://doi.org/10.1016/j.lwt.2010.01.010. Scopus.

Bergquist, J., Palmblad, M., Wetterhall, M., Håkansson, P., & Markides, K. E. (2002). Peptide mapping of proteins in human body fluids using electrospray ionization fourier transform ion cyclotron resonance mass spectrometry. *Mass Spectrometry Reviews, 21*(1), 2–15. https://doi.org/10.1002/mas.10016. Scopus.

Borghesan, F., Mistrello, G., Amato, S., Giuffrida, M. G., Villalta, D., & Asero, R. (2013). Mugwort-fennel-allergy-syndrome associated with sensitization to an allergen homologous to Api g 5. *European Annals of Allergy and Clinical Immunology, 45*(4), 130–137. Scopus.

Bourmaud, A., Gallien, S., & Domon, B. (2016). Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: Principle and applications. *Proteomics, 16*(15–16), 2146–2159. https://doi.org/10.1002/pmic.201500543.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry, 72*(1–2), 248–254. https://doi.org/10.1016/0003-2697(76)90527-3. Scopus.

Chen, C.-H. W. (2008). Review of a current role of mass spectrometry for proteome research. *Analytica Chimica Acta, 624*(1), 16–36. https://doi.org/10.1016/j.aca.2008.06.017. Scopus.

Cottrell, J. S. (2011). Protein identification using MS/MS data. *Journal of Proteomics, 74*(10), 1842–1851. https://doi.org/10.1016/j.jprot.2011.05.014. Scopus.

Dodds, E. D., Clowers, B. H., Hagerman, P. J., & Lebrilla, C. B. (2008). Systematic characterization of high mass accuracy influence on false discovery and probability scoring in peptide mass fingerprinting. *Analytical Biochemistry, 372*(2), 156–166. https://doi.org/10.1016/j.ab.2007.10.009. Scopus.

ExPaSy Bioinformatics Resource Portal. (n.d.). https://web.expasy.org/peptide_mass.

Floris, F., Chiron, L., Lynch, A. M., Barrow, M. P., Delsuc, M.-A., & O'Connor, P. B. (2018). Application of tandem two-dimensional mass spectrometry for top-down deep sequencing of calmodulin. *Journal of the American Society for Mass Spectrometry, 29*(8), 1700–1705. https://doi.org/10.1007/s13361-018-1978-y. Scopus.

Hoffmann-Sommergruber, K., Riordain, G., Ahorn, H., Ebner Laimer Da Camara, C., Machado, M., Pühringer, H., Scheiner, O., & Breiteneder, H. (1999). Molecular characterization of Dau c 1, the Bet v 1 homologous protein from carrot and its cross-reactivity with Bet v 1 and Api g 1. *Clinical and Experimental Allergy, 29*(6), 840–847. https://doi.org/10.1046/j.1365-2222.1999.00529.x. Scopus.

Horn, D. M., Peters, E. C., Klock, H., Meyers, A., & Brock, A. (2004). Improved protein identification using automated high mass measurement accuracy MALDI FT-ICR MS peptide mass fingerprinting. *International Journal of Mass Spectrometry, 238*(2), 189–196. https://doi.org/10.1016/j.ijms.2004.03.016.

Huang, T., Wang, J., Yu, W., & He, Z. (2012). Protein inference: A review. *Briefings in Bioinformatics, 13*(5), 586–614. https://doi.org/10.1093/bib/bbs004. Scopus.

Jensen-Jarolim, E., Leitner, A., Hirschwehr, R., Kraet, D., Wuthrich, B., Scheiner, O., … Ebner, C. (1997). Characterization of allergens in Apiaceae spices: Anise, fennel, coriander and cumin. *Clinical Experimental Allergy, 27*(11), 1299–1306. https://doi.org/10.1111/j.1365-2222.1997.tb01175.x.

Kanawati, B. (2019). Chapter 6—Data processing and automation in Fourier transform mass spectrometry. 53.

Kanawati, B., Bader, T. M., Wanczek, K.-P., Li, Y., & Schmitt-Kopplin, P. (2017). Fourier transform (FT)-artifacts and power-function resolution filter in Fourier transform mass spectrometry. *Rapid Communications in Mass Spectrometry, 31*(19), 1607–1615. https://doi.org/10.1002/rcm.7940.

Karty, J. A., Ireland, M. M. E., Brun, Y. V., & Reilly, J. P. (2002). Defining absolute confidence limits in the identification of *Caulobacter* proteins by peptide mass mapping. *Journal of Proteome Research, 1*(4), 325–335. https://doi.org/10.1021/pr025518b.

Khodadadi, E., Fakheri, B. A., Aharizad, S., Emamjomeh, A., Norouzi, M., & Komatsu, S. (2017). Leaf proteomics of drought-sensitive and -tolerant genotypes of fennel. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 1865*(11), 1433–1444. https://doi.org/10.1016/j.bbapap.2017.08.012.

Lesur, A., & Domon, B. (2015). Advances in high-resolution accurate mass spectrometry application to targeted proteomics. *Proteomics, 15*(5–6), 880–890. https://doi.org/10.1002/pmic.201400450.

Liu, G., Tu, Z., Wang, H., Zhang, L., Huang, T., & Ma, D. (2017). Monitoring of the functional properties and unfolding change of Ovalbumin after DHPM treatment by HDX and FTICR MS. *Food Chemistry, 227*, 413–421. https://doi.org/10.1016/j.foodchem.2017.01.109.

Mann, M., Højrup, P., & Roepstorff, P. (1993). Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biological Mass Spectrometry, 22*(6), 338–345. https://doi.org/10.1002/bms.1200220605.

Marshall, A. G., Hendrickson, C. L., & Jackson, G. S. (1998). Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrometry Reviews, 17*(1), 1–35. https://doi.org/10.1002/(SICI)1098-2787(1998)17:1<1::AID-MAS1>3.0.CO;2-K. Scopus.

Nardiello, D., Kanawati, B., & Melfi, M. T. (2020). Data processing for fennel protein characterization by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS). 1. https://doi.org/10.17632/s354fkt6fp.1.

Nardiello, D., Palermo, C., Natale, A., Quinto, M., & Centonze, D. (2015). Strategies in protein sequencing and characterization: Multi-enzyme digestion coupled with alternate CID/ETD tandem mass spectrometry. *Analytica Chimica Acta, 854*, 106–117. https://doi.org/10.1016/j.aca.2014.10.053.

Page, J. S., Masselon, C. D., & Smith, R. D. (2004). FTICR mass spectrometry for qualitative and quantitative bioanalyses. *Current Opinion in Biotechnology, 15*(1), 3–11. https://doi.org/10.1016/j.copbio.2004.01.002.

Papaleo, E., Fantucci, P., & De Gioia, L. (2005). Effects of calcium binding on structure and autolysis regulation in trypsins. A molecular dynamics investigation. *Journal of Chemical Theory and Computation, 1*(6), 1286–1297. https://doi.org/10.1021/ct050092o.

Pastorello, E. A., Farioli, L., Stafylaraki, C., Scibilia, J., Giuffrida, M. G., Mascheri, A., … Pravettoni, V. (2013). Fennel allergy is a lipid-transfer protein (LTP)-related food hypersensitivity associated with peach allergy. *Journal of Agricultural and Food Chemistry, 61*(3), 740–746. https://doi.org/10.1021/jf303291k.

Piras, C., Roncada, P., Rodrigues, P. M., Bonizzi, L., & Soggiu, A. (2016). Proteomics in food: Quality, safety, microbes, and allergens. *Proteomics, 16*(5), 799–815. https://doi.org/10.1002/pmic.201500369.

Quenzer, T. L., Emmett, M. R., Hendrickson, C. L., Kelly, P. H., & Marshall, A. G. (2001). High sensitivity fourier transform ion cyclotron resonance mass spectrometry for biological analysis with Nano-LC and microelectrospray ionization. *Analytical Chemistry, 73*(8), 1721–1725. https://doi.org/10.1021/ac001095q.

Rahimi, R., & Ardekani, M. R. S. (2013). Medicinal properties of Foeniculum vulgare Mill. In traditional Iranian medicine and modern phytotherapy. *Chinese Journal of Integrative Medicine, 19*(1), 73–79. https://doi.org/10.1007/s11655-013-1327-0.

Tucholski, T., Knott, S. J., Chen, B., Pistono, P., Lin, Z., & Ge, Y. (2019). A top-down proteomics platform coupling serial size exclusion chromatography and Fourier transform ion cyclotron resonance mass spectrometry. *Analytical Chemistry, 91*(6), 3835–3844. https://doi.org/10.1021/acs.analchem.8b04082.

van der Burgt, Y. E. M., Kilgour, D. P. A., Tsybin, Y. O., Srzentić, K., Fornelli, L., Beck, A., … Nicolardi, S. (2019). Structural analysis of monoclonal antibodies by ultrahigh resolution MALDI in-source decay FT-ICR mass spectrometry. *Analytical Chemistry, 91*(3), 2079–2085. https://doi.org/10.1021/acs.analchem.8b04515.

Witt, M., Fuchser, J., & Baykut, G. (2003). Fourier transform ion cyclotron resonance mass spectrometry with NanoLC/microelectrospray ionization and matrix-assisted laser desorption/ionization: Analytical performance in peptide mass fingerprint analysis. *Journal of the American Society for Mass Spectrometry, 14*(6), 553–561. https://doi.org/10.1016/S1044-0305(03)00138-7.

Yates, J. R. (1998). Database searching using mass spectrometry data. *Electrophoresis, 19*(6), 893–900. https://doi.org/10.1002/elps.1150190604.

Zhang, X., Jiang, X., Wang, X., Zhao, Y., Jia, L., Chen, F., … Han, F. (2019). A metabolomic study based on accurate mass and isotopic fine structures by dual mode combined-FT-ICR-MS to explore the effects of Rhodiola crenulata extract on Alzheimer disease in rats. *Journal of Pharmaceutical and Biomedical Analysis, 166*, 347–356. https://doi.org/10.1016/j.jpba.2019.01.021.