UNIVERSITÀ DEGLI STUDI DI FOGGIA

DOTTORATO DI RICERCA IN
"MEDICINA TRASLAZIONALE ED ALIMENTI:
INNOVAZIONE, SICUREZZA E MANAGEMENT"
(XXXII CICLO)

COORDINATORE: PROF. MATTEO ALESSANDRO DEL NOBILE

THE CHALLENGE OF PROTEOMICS IN THE FIELD OF FOOD SAFETY: ALLERGEN DETECTION

Dottorando: Maria Teresa Melfi

Tutor: Donatella Nardiello
Co-tutor: Maria Filomena Caiaffa

ANNO ACCADEMICO 2018/2019

INDEX

ABSTRACT (*English Version*)

Food safety and quality and their associated risks pose a major concern worldwide regarding not only the potential danger to consumer's health but also the relative economic losses. A lack of measures and reliable methods to evaluate and maintain a good control of food characteristics may shatter consumer confidence and affect the food industry economy. In this scenario, it is imperative to develop fast and reliable analytical methods that allow a good and rapid analysis of food products during the whole food chain. Proteomics can represent a powerful tool to address this issue, due to its proven excellent quantitative and qualitative advantages in protein characterization. There are several applications of proteomics to food analysis, aimed at protecting consumer benefits, from the evaluation of the nutritional properties and product traceability (food quality) to allergen detection (food safety) in raw and processed foods. Mass Spectrometry (MS) based methods have been suggested as confirmatory tools for an accurate protein identification in food samples, nevertheless, the characterization of proteins always needs a concerted application of several technologies, where MS measurements represent only the final step. A successful MS identification can be indeed obtained only if supported by proper separation strategies such as two-dimensional liquid chromatography (2D-LC) that has been recently added to the family of LC techniques, providing a remarkable potential increase in peak capacity and resolving power.

In this research project, automated 2D-LC methods based on the use of multi-port switching valves were developed and optimized for the separation of intact proteins from the major food allergen sources. Egg, soy flour and fennel samples were chosen as cases of study: egg and soy flour represent two of the major food allergen sources worldwide (both included in the Big-8 listing), whereas fennel has been recently recognized as an allergenic source, especially in the Mediterranean area. The optimized 2DLC methods are based on an innovative trapping interface made by a reversed-phase column installed in a switching valve, placed between two separation columns. Therefore, the proteins coming from the first column were trapped and focused before injection in the second column, reducing the effects of dead volumes and band broadening. A size-exclusion chromatography (SEC) column was used for the first dimension separation and combined to a reversed-phase (RP) column for the second dimension separation. Protein peaks coming from the second separation column were collected through a fraction collector connected to the UV cell. The optimized separation workflow, as a good compromise between the *comprehensive*

and *multiple heart-cutting* modes, takes on the challenging task of analyzing complex food matrices rich in proteins greatly different in concentration, mass, and polarity, ensuring high resolving power and automation grade. The proposed gel-free method for protein separation allows to perform an automated isolation of individual proteins for their subsequent MS characterization (by both top-down or bottom-up proteomics) or for further bio-analytical investigations. The experimental conditions were optimized by analyses of mixed standard solutions of bovine serum albumin, glucose oxidase, immunoglobulin A, thyroglobulin and myoglobin. Then, the optimized *stop-and-go/active modulation* SEC-RP LC approach was applied to the protein analysis in extracts of egg and soy flour, with the final aim to recovery sufficient protein amounts for the molecular characterization and the assessment of the pattern of allergenic components.

For the protein characterization in fennel extracts, a rapid and sensitive bottom-up method by Fourier transform ion cyclotron resonance (FT-ICR) was optimized. The analyses were performed for the identification of all the proteins present in the whole raw fennel extract (without prior fractionation by 2DLC), taking advantage of the elevated sensitivity and mass accuracy of FTICR-MS. The peptide-level method was previously validated on tryptic digests from ubiquitin standard protein. Then, few microliters of fennel extracts were analyzed by direct infusion after enzymatic digestion with trypsin. The method benefits from the high resolution which allows protein detection in a mass range up to m/z 8000 in a few seconds. The experimental mass spectra peak-lists were compared with theoretical peptide sequences and mass values coming from the in silico-digestion performed on a custom-made proteome database originating from the NCBI fennel protein database. Finally, the matched mass lists were used in the database searching for protein identification by Peptide Mass Fingerprint. A total of 70 proteins were identified, with molecular weights ranging from 4.5 kDa to 250 kDa. In order to clarify the effective allergenic molecules present in the fennel samples, the list of the identified proteins was also matched with data obtained by antibody-based assays for Immunoglobulin E detection, performed by using the same fennel extract. Immunoblotting analysis was performed in sera samples of fennel allergy subjects, and 2 major immunoreactive bands, putatively involved in fennel allergy, at around 33 and 50 kDa were detected, corresponding to the proteins associated to the mugwort-spice-allergy-syndrome.

Keywords: food safety; egg; soy flour; fennel; allergen detection; protein analysis; two-dimensional Liquid Chromatography (2D-LC); Mass Spectrometry (MS).

ABSTRACT (*Italian Version*)

La qualità e la sicurezza alimentare e i rischi a loro associati rappresentano una delle principali preoccupazioni in tutto il mondo, non solo per quanto riguarda il potenziale pericolo per la salute dei consumatori, ma anche per le relative ricadute economiche. La mancanza di misure e metodi affidabili per valutare e mantenere un buon controllo delle caratteristiche degli alimenti può compromettere la fiducia dei consumatori e influire sull'economia dell'industria alimentare. In questo scenario, è indispensabile sviluppare metodi analitici veloci e sicuri che consentano un'analisi rapida ed affidabile dei prodotti alimentari durante l'intera filiera alimentare. La proteomica può rappresentare un potente strumento per affrontare questo problema, grazie ai suoi comprovati eccellenti vantaggi, quantitativi e qualitativi, nella caratterizzazione delle proteine. Esistono diverse applicazioni della proteomica all'analisi degli alimenti, volte a tutelare i benefici dei consumatori, dalla valutazione delle proprietà nutrizionali e della tracciabilità del prodotto (qualità degli alimenti) alla rilevazione degli allergeni (sicurezza alimentare) negli alimenti crudi e trasformati. I metodi basati sulla spettrometria di massa (MS) sono comunemente utilizzati come strumenti di conferma per un'accurata identificazione delle proteine nei campioni alimentari; tuttavia, la caratterizzazione delle proteine necessita sempre di una combinazione di diverse tecnologie e le misure di spettrometria di massa rappresentano solo lo step finale. Una corretta identificazione in spettrometria di massa può essere effettivamente ottenuta solo se supportata da adeguate strategie di separazione come la cromatografia liquida bidimensionale (2D-LC), che è stata recentemente aggiunta alla famiglia delle tecniche di cromatografia liquida, fornendo un notevole aumento della capacità di picco e potere risolvente.

In questo progetto di ricerca, sono stati sviluppati e ottimizzati metodi automatizzati in cromatografia bidimensionale 2D-LC basati sull'uso di valvole di smistamento multi-porta per la separazione delle proteine intatte dalle principali fonti alimentari allergeniche. Campioni di uova, farina di soia e finocchio sono stati scelti come casi di studio: l'uovo e la farina di soia rappresentano due delle principali fonti di allergeni alimentari nel mondo (entrambe incluse nella lista dei "*Big-8*"), mentre il finocchio è stato recentemente riconosciuto come una fonte allergenica, in particolare nell'area mediterranea. I metodi 2DLC ottimizzati si basano su un'interfaccia innovativa di intrappolamento realizzata da una colonna in fase inversa installata su una valvola di smistamento, posizionata tra le due colonne di separazione. Pertanto, le proteine provenienti dalla prima colonna sono

intrappolate e focalizzate prima dell'iniezione nella seconda colonna, riducendo gli effetti dei volumi morti e dell'allargamento di banda. Una colonna cromatografica ad esclusione dimensionale (SEC) è stata utilizzata per la separazione in prima dimensione e combinata ad una colonna a fase inversa (RP) per la separazione in seconda dimensione.

I picchi proteici provenienti dalla seconda colonna di separazione sono stati raccolti attraverso un collettore di frazioni collegato alla cella UV. Il flusso di lavoro di separazione ottimizzato, come buon compromesso tra le modalità *comprehensive* e *multiple heart-cutting*, svolge il difficile compito di analizzare matrici alimentari complesse ricche di proteine molto diverse in concentrazione, massa e polarità, garantendo un elevato potere risolutivo e grado di automazione. Il metodo *gel-free* proposto per la separazione delle proteine consente di eseguire un isolamento automatizzato delle singole proteine per la loro successiva caratterizzazione in spettrometria di massa (mediante approccio proteomico *top-down* o *bottom-up*) o per ulteriori indagini bioanalitiche. Le condizioni sperimentali sono state ottimizzate mediante analisi su miscele di soluzioni standard di albumina sierica bovina, glucosio ossidasi, immunoglobulina A, tireoglobulina e mioglobina. Quindi, l'approccio *stop-and-go/active modulation* SEC-RP LC ottimizzato è stato applicato all'analisi proteica degli estratti di uova e farina di soia, con l'obiettivo finale di recuperare quantità proteiche sufficienti per la caratterizzazione molecolare e per la valutazione di *pattern* allergenici.

Per la caratterizzazione delle proteine negli estratti di finocchio, un metodo *bottom-up* rapido e sensibile è stato ottimizzato mediante l'analizzatore a risonanza ionica ciclotronica a trasformata di Fourier (FT-ICR). Sono state eseguite le analisi per l'identificazione di tutte le proteine presenti nell'intero estratto di finocchio crudo (senza precedente frazionamento mediante 2DLC), sfruttando l'elevata sensibilità ed accuratezza di massa dell'analizzatore FTICR-MS. Il metodo a livello peptidico è stato precedentemente validato sui digeriti triptici di ubiquitina utilizzata come proteina standard di riferimento. Successivamente, pochi microlitri di estratto di finocchio sono stati analizzati mediante infusione diretta in seguito a digestione enzimatica con tripsina. Il metodo beneficia della strumentazione ad risoluzione che consente la rivelazione di proteine in un intervallo di m/z fino a 8000 in pochi secondi. I dati m/z sperimentali degli spettri di massa del campione di finocchio sono stati confrontati con le sequenze teoriche di peptidi provenienti dalla digestione in silico eseguita su un *database* proteomico costruito su misura a partire dal *database* NCBI di tutte le proteine di finocchio. Infine, gli elenchi di massa risultanti dal confronto sono stati utilizzati per la ricerca in banca dati per

l'identificazione delle proteine mediante MASCOT, in modalità *Peptide Mass Fingerprint*. Sono state identificate un totale di 70 proteine, con pesi molecolari che vanno da 4.5 kDa a 250 kDa. Al fine di chiarire le effettive molecole allergeniche presenti nei campioni di finocchio, l'elenco delle proteine identificate è stato inoltre confrontato con i dati ottenuti dai test immunologici per la rilevazione di immunoglobuline E, eseguiti utilizzando lo stesso estratto di finocchio. Sono state eseguite analisi di immunoblotting in campioni di sieri di soggetti allergici al finocchio e sono state rivelate 2 principali bande immunoreattive putativamente coinvolte nell'allergia al finocchio, a circa 33 e a 50 kDa, corrispondenti alle proteine associate alla sindrome allergica da spezia e artemisia.

Parole chiave: sicurezza alimentare; uova; farina di soia; finocchio; rivelazione allergeni; analisi proteine; cromatografia liquida bidimensionale (2D-LC); spettrometria di massa (MS).

1. INTRODUCTION

1.1. Food allergy and the EU Labeling Directives

Food allergy is a matter of public health, affecting 1-3% of adults and 4-6% of children, and in the last 20 years the rate has increased considerably [1,2]. The causes of the increase and spread of allergic reactions are still unclear and may be due to a combination of different factors such as environmental factors (eating habits, introduction of food and breastfeeding), individual factors (genetic background, age and sex, ethnicity), and food processing and preparation [3–7]. Geographical variations in the prevalence of food allergy are driven by genetic factors and further modified by regional or local factors, like pollen exposure or differences in food habits. Inter-country differences in reporting adverse reactions to foods have also been noted and likely attributed to cultural differences.

Food allergy can cause one or more symptoms that can be more or less severe, including angioedema, oral allergic syndrome, urticaria, abdominal pain, diarrhea, nausea, vomiting, itching and even life-threatening reactions, such as anaphylactic shock. There is no cure for allergy since immunotherapy is still far from being routinely available [8]. After an adverse reaction, the allergic population must strictly avoid consuming the offending food. Another important risk for food-allergic consumers, however, is the presence of hidden allergens due to cross-contamination during food processing. The absence of a regulatory framework for managing hidden allergens and a lack of legal action thresholds have prompted the food industry to make excessive use of precautionary allergen labeling (PAL), leading to a loss of consumer trust [9–13].

Recently, various countries have set legal thresholds (e.g., Switzerland, Germany, Belgium, and the Netherlands), but a considerable disparity is observed among these thresholds. In Australia and New Zealand, the Voluntary Incidental Trace Allergen Labeling (VITAL) system establishes eliciting doses (EDs) based on clinical studies for the protection of at least 95% of allergic people (ED05) [9,10]. VITAL thresholds, which have no regulatory status, are set at 0.75 mg per kg for egg proteins, 2.5 mg per kg for milk or tree nut proteins, 5 mg per kg for peanut proteins, 25 mg per kg for soybean proteins, and 50 mg per kg for cashew proteins (portion size: 40 g).

While the Codex Alimentarius Commission Committee on Food Labeling lists the major allergens on a worldwide basis, the foods, which are common causes of allergic reactions, differ between geographical areas, as a result of dietary preferences. Some countries have

chosen to include additional foods on their national list of foods and ingredients that must be declared on food labels. The EU, for example, has chosen to add celery, mustard, sesame seeds, lupin, and molluscs and products thereof to the list of allergens.

Food businesses are obliged to provide safe foods according to Article 14 of the General Food Law (https://ec.europa.eu/food/safety/general_food_law_en). A food is presumed unsafe if it does not comply with the EU (or national) legislation governing its safety. As an example, this could be a food that contains an allergenic ingredient, which should be labeled in accordance with Directive 2000/13/EC, but no declaration is reported.

Even if the food complies with the EU (or national) legislation governing its safety, food business or a competent authority may have reasons to consider the food unsafe. This could occur if a food contains a very high concentration of cross-contact allergen, which could trigger a reaction in a significant number of allergic consumers.

The Labeling Directive (Directive 2000/13/EC) and its later amendments are the only pieces of EU legislation that specifically refer to allergenic foods. The Labeling Directive requires manufacturers to declare all ingredients present in pre-packaged foods sold in the EU with very few exceptions. This directive has been amended a number of times with regard to allergens. The two most important amendments are:

- Directive 2003/89/EC introduced Annex IIIa, which is a list of allergenic foods that must always be labeled when present in a product;
- Directive 2007/68/EC has the most recent amendment of Annex IIIa. It lists all the allergenic foods that must be labeled as well as a few products derived from these foods for which allergen labeling is not required.

The European Food Safety Authority (EFSA) web site also provides information on food allergen labeling in Europe. As a result of the Labeling Directive, it is mandatory to label the allergenic foods listed in Annex IIIa or any product derived from these foods with a few exceptions shown later. The Allergenic foods listed in Annex IIIa are:

- Cereals containing gluten, (i.e. wheat, rye, barley, oats, spelt, kamut or their hybridized strains) and products thereof
- Crustaceans and products thereof
- Eggs and products thereof
- Fish and products thereof
- Peanuts and products thereof
- Soybeans and products thereof

- Milk and products thereof (including lactose)
- Nuts i.e. almonds, hazelnuts, walnuts, cashews, pecan nuts, Brazil nuts, pistachio nuts, macadamia nuts and Queensland nuts and products thereof
- Celery and products thereof
- Mustard and products thereof
- Sesame seeds and products thereof
- Sulfur dioxide and sulfites at concentrations of more than 10 mg/kg or 10 mg/L expressed as SO_2
- Lupin and products thereof
- Molluscs and products thereof

Some products made from the allergenic foods in Annex IIIa have been permanently exempted from allergen labeling based on opinions from EFSA. These products are not likely to cause severe allergic reactions as they only contain trace amounts of protein. The exemptions are reported in the table below (Table 1). The effect of the exemptions is for example that it is possible to label fully refined soybean oil as just "vegetable oil".

Table 1. List of food ingredients and substances permanently excluded from Annex IIIa of the Labeling Directive.

INGREDIENT	PRODUCTS THEREOF PERMANENTLY EXCLUDED
Cereals containing gluten	▪ Wheat-based glucose syrups including dextrose ▪ Wheat-based maltodextrins ▪ Glucose syrups based on barley ▪ Cereals used for making distillates or ethyl alcohol of agricultural origin for spirit drinks and other alcoholic beverages
Fish	▪ Fish gelatine used as a carrier for vitamin or carotenoid preparations ▪ Fish gelatine or Isinglass used as fining agents in beer and wine
Soybean	▪ Fully refined soybean oil and fat ▪ Natural mixed tocopherols (E306), natural D-alpha tocopherol, natural D-alpha tocopherol acetate, natural D-alpha tocopherol succinate from soybean sources ▪ Vegetable oils derived phytosterols and phytosterol esters from soybean sources ▪ Plant stanol ester produced from vegetable oil sterols from soybean sources
Milk	▪ Whey used for making distillates or ethyl alcohol of agricultural origin for spirit drinks and other alcoholic beverages ▪ Lactitol
Nuts	▪ Nuts used for making distillates or ethyl alcohol of agricultural origin for spirit drinks and other alcoholic beverages

The Labeling Directive only covers pre-packaged food. This means that foods sold loose or from catering outlets are not covered by current allergen labeling rules. These foods represent a considerable risk to allergic consumers. That is why EU has proposed a new legislation, which also covers non-pre-packed foods, including the food supplies for restaurants and cafes. The new Regulation (EU) No 1169/2011, dealing with the provision of food information to consumers, entered into application on 13 December 2014. Therefore, from the analytical point of view, it is really important to be able to detect allergens in complex food matrices at low concentrations, since even a small allergen contamination occurred during food processing and not indicated in the food label could represent a serious burden for allergic patients. Hence, allergen determination in food products is essential for the purposes of the legislation, nutrition and public health, and the development of sensitive and selective methods represents an important topic for food assurance and quality control. Although many allergens have been identified in different allergenic sources, as detailed in the Allergome database (www.allergome.org) [13,14], more than 90% of food allergies are caused by cow's milk, egg, fish, crustaceans, peanuts, tree nuts, wheat and soybeans, which are referred as "The Big Eight" (Figure 1) [15].

Figure 1. Icons of the most common allergenic food: "The Big 8".

Any intentional use of the Big 8 foods have to be clearly labeled on food labels since the best treatments for food allergic subjects remains the avoidance of the allergen sources, that means the exclusion of specific foods and/or additives from the diet. These foods must be listed by their common name and may be included in the ingredients list or separately in a "contains" statement. Precautionary labeling, however, such as "*may contain*" or "*processed in a facility with*" is voluntary and is not regulated or required. On the other side, such a precautionary food labeling has the effect of increasing the list of specific

allergenic molecules/sources that each subject has to avoid, thus causing unnecessary dietary restrictions that may negatively affect the quality of life.
Product formulations can change, so labels should be read every time to ensure that packaged foods do not contain the allergenic food to be avoided. Actually, the field of allergenic sources is much more complex, since uncommon food allergies to spices (fennel, celery, etc) have been rapidly expanding in recent years.

1.2. Allergenicity of food proteins

Semantically, the concept of allergenicity is not clearly defined. To a clinical allergist, allergenicity reflects the capacity of an antigen to induce symptoms or a skin reaction, whereas to an immunologist, it reflects either a peculiar type of immunogenicity (i.e., the capacity of a protein to induce IgE antibodies) or simply the capacity to bind IgE antibodies. Similarly, the term allergen is used to describe different molecular properties: the property to sensitize (i.e., induce the immune system to produce high-affinity antibodies, particularly of the IgE class) and the property to elicit an allergic reaction (i.e., to trigger allergic symptoms in a sensitized subject). Moreover, it is also used to indicate the property to bind IgE antibodies. Complete allergens have all these properties. Some proteins, however, are known to elicit allergic symptoms but do not usually sensitize. The distinction between major and minor allergens is relevant for various reasons, but also in relation to the issue of allergenicity. The current definition of major allergen is based on the prevalence of IgE or skin reactivity in subjects that are sensitized (usually very strongly) to the total extract. This definition is unsatisfactory in that it does not reflect the contribution of the allergen to the overall reactivity of the extract. Intuitively, removal of a truly major allergen from an extract is expected to have a noticeable effect on the overall reactivity of that extract: a major allergen should make a difference. Such an interpretation invites a different type of definition. For example, a major allergen is responsible for more than 20% of the allergenic reactivity in more than 20% of the sensitized patients. This requires testing with extracts from which the allergen in question has been selectively removed (e.g., with monospecific antibodies). Alternatively, it could be tested serologically by absorbing out all IgE antibodies to the allergen and then testing the residual activity of the absorbed serum. For most allergens, this aspect has not been

studied. It is likely, however, that the major allergen claim made for many allergens would need to be reconsidered.

Food allergens are almost always proteins, but not all food proteins are allergens. The allergenicity of a protein is due to the IgE-binding epitopes that are widespread within the protein molecule. Epitope mapping is the characterization of all epitopes in an allergen molecule. Not all the epitopes in a protein are recognized by all patients allergic to that protein: some epitopes are immuno-dominant, while others are only recognized by few patients. Two types of epitopes have been described depending on their structure: conformational epitopes, which are associated with the secondary and tertiary structure of the protein, and linear/sequential epitopes, formed by a continuous sequence of amino acid residues in the protein chain. Once the protein is denatured, conformational epitopes are generally modified or destroyed, whereas linear epitopes are maintained. The clinical significance of epitopes may depend on their structure and location within the molecule. For example, short linear IgE-binding epitopes located in hydrophobic parts of allergenic proteins could be used as markers of a persistent food allergy, i.e. to milk and to peanut.

In the last two decades, great efforts were undertaken to identify the allergenic proteins from plant-derived and animal foods, to study their physicochemical characteristics and their interaction with immune cells. Consecutively, allergen databases were built and maintained to provide extensive information about allergens (e.g. www.allergen.org). It became evident, that only a minority of all known protein families contain food allergens [16]. With regard to plant food allergens, the most relevant protein families are: prolamins, cupins, profilins, and the Bet v 1 superfamily.

The prolamin superfamily contains the largest number of plant food allergens: 2S seed storage albumins, cereal seed storage proteins, cereal α-amylase/trypsin inhibitors and non-specific lipid transfer proteins (nsLTPs). Prolamins were originally defined on the basis of their water/alcohol solubility and of their content of proline and glutamine. Prolamins are characterized by a high content in sulfur-containing amino acid residues and often consist of bundles of four α-helices stabilized by disulfide bonds, involving eight well-conserved cysteine residues. The major role of 2S albumins is to provide proteins to the developing seed. They also have a defensive role against pathogenic fungi. Major allergens in tree nuts, sesame and mustard seeds belong to this family. Cereal α-amylase and protease inhibitors induce a certain resistance of plant tissues to insect pests and include allergens present in wheat, barley, rice and corn. The lipid transfer protein family comprises low molecular weight monomeric proteins (around 7-9 kDa) involved in the synthesis of cutin,

and thus have a protective role in the plant, and particularly in the fruit. They have a very compact and stable tertiary structure constituted by the association of α-helices and loops stabilized by eight disulfide bonds, which define a central cavity containing a lipid-binding site. Binding with hydrophobic ligands also contributes to the stabilization of the molecule. Lipid transfer proteins are frequent and potentially severe allergens: they are one of the numerous defense protein families (also called pathogenesis-related proteins) that are responsible for most of the allergic reactions to fruits from the *Rosaceae* family [17].

The cupin superfamily includes the major globulin storage proteins, which are the cause of most allergic reactions to legumes and nuts. The name comes from their common architecture, consisting of 6-stranded β-sheets associated with α-helices which form a β-barrel cavity (Latin *cupa*, barrel) with a binding site for a hydrophobic ligand. Subgroups in the cupin superfamily have been defined depending on the number of cupin domains present in the protein. Monocupins comprise the majority of cupin proteins, can be monomeric, dimeric or oligomeric, and most are enzymes (e.g. dioxygenases). Germin and germin-like proteins (GLP) are oligomeric monocupins ubiquitarious in plants (e.g. wheat and barley). They have a disc-shape homohexameric structure organized as trimers of dimers. The globulin fractions of seed storage proteins, which can be extracted with saline solutions, are 2-domain cupins. According to their sedimentation coefficient determined by ultracentrifugation, globulins are divided in a smaller fraction, i.e. 7S/8S globulins (called vicilins), and a bigger fraction i.e. 11S globulins (called legumins). 7S/8S globulins are generally trimers with 50-60 kDa molecular weight (MW). Post-translational modifications such as glycosylation often occur. 11S globulins consist of six subunits with a MW around 60 kDa and are rarely glycosylated. Each subunit consists of a non-covalent association of two polypeptide chains. 7S and 11S globulins have a relatively low sequence identity but a common 3D conformation. Globulins are clinically relevant allergens in peanuts, soybean, lentils, walnut, hazelnut and sesame.

Profilins are cytosolic proteins of 12 to 15 kDa exclusively found in flowering plants, such as peanut (Ara h 5), apple (Mal d 4) and celery (Api g 4). They are folded in a compact globular structure of an antiparallel β-sheet enclosed by α-helices on both sides. The high sequence conservation and the even higher 3D structure similarity account for the strong serological cross-reactivity with other plant foods, pollens and *Hevea latex*, which may be of variable clinical significance.

The Bet v 1 superfamily comprises eight families, among which the “pathogenesis-related proteins” (PR 10), major latex proteins. These allergens are homologous of the major birch

pollen allergen Bet v 1 and are present in fruits of the *Rosaceae* family (e.g. apple, cherry, apricot, and pear) and *Apiaceae* vegetables (e.g. celery, carrot). They are polypeptides of 154-160 amino acids with high sequence similarity. The Bet v 1 homologous proteins contain a GXGXXG or a GXG motif, responsible for binding of the phosphate group of oligonucleotides, and share a characteristic fold formed by seven β-sheets surrounding a long C-terminal helix and two additional short helices connecting two β-sheets and forming a large y-shaped hydrophobic cavity able to bind sterols, as observed in structures obtained by X-ray crystallography. Because of their sequence and 3D similarities, the Bet v 1 related proteins cross-react with allergens present in birch pollen, sometimes inducing severe allergic reactions.

Food allergens of animal origin, less numerous than allergens of plant origin, are classified in three main structurally-related families: tropomyosins, parvalbumins and caseins. Tropomyosins are a family of closely related proteins present in muscle and non-muscle cells. Together with actin and myosin, tropomyosins play a key regulatory role in muscle contraction. Tropomyosins form head-to-tail polymers along the length of an actin filament and are the major allergens of two invertebrate groups, Crustacea and Mollusca, that are generally referred to as shellfish. Shrimp, crab, squid, and abalone are assumed to be largely responsible for seafood allergies. Allergenic tropomyosins are heat stable and cross-reactive between the various crustacean and mollusk species. The second largest animal food allergen family are the parvalbumins. Abundant in the white muscle of many fish species, paravalbumins are characterized by the presence of a widely found calcium-binding domain which is known as the “EF-hand”. Parvalbumins with bound calcium ions possess remarkable stability to denaturation by heat. The ability to act as major fish allergens is also linked to the stability of parvalbumins to denaturing chemicals, and proteolytic enzymes. Finally, caseins (in particular αS1, αS2, and β) are major food allergens involved in cow’s milk allergy, which affects predominantly young children. Caseins are structurally mobile proteins present in mammalian milk at a concentration of around 15 mg/mL and are responsible for binding calcium through clusters of phosphoserine and/or phosphothreonine residues.

1.2.1. Egg allergens

Egg is one of the foods whose allergenicity is most altered by cooking or processing. Egg white contains proteins with considerably higher allergenic potential than the egg yolk. The four major allergens in egg white are ovalbumin (OVA; Gal d 2, 54% of the total protein content), ovotransferrin (OVT; conalbumin, Gal d 3, 12%), ovomucoid (OVO; Gal d 1, 11%), and lysozyme (LYS; Gal d 4, 3.5%). Two yolk proteins, α-livetin (chicken serum albumin, Gal d 5) and lipoprotein YGP42 (Gal d 6), have been identified as egg allergens [18,19]. Ovalbumin possesses 4 sulphydryl groups with a single disulfide bridge [20] and is found to be sensitive to heat denaturation. Ovomucoid consists of 3 sub-domains, each being internally linked by disulfide bonds, which makes ovomucoid resistant to heat denaturation and proteolytic digestion [21]. Domain 3 is very stable, has most dominant IgE and IgG-binding epitopes, and is considered the major determinant of the strong allergenicity of the protein (Jiménez-Saiz et al., 2011b; Maeno et al., 2013). Gal d 5 is partially heat-labile and can cause both respiratory and food allergy symptoms in patients with bird-egg syndrome [22], while Gal d 6 is a heat-stable allergen [23]. Eggs are universally used as nutrient and food additive. Common products in which egg is extensively heated are baked products and starches containing egg such as cakes, waffles, muffins, pancakes, egg noodles, egg pasta and bread.

Several human studies have been performed in which egg-allergic patients were challenged with heated and/or unheated eggs. In general, 50–85% of children with egg allergy are able to tolerate baked egg [24] This percentage varies depending on the characteristics of the allergic patients (age, severity of the allergy, etc.), heating procedure, matrix used, etc. A major drawback of these studies is that the individuals are often not challenged to unheated eggs to establish clinical reactivity to egg prior to the oral food challenge (OFC) to the heated egg. This can result in an overestimation of the percentage of allergic subjects tolerating extensively heated egg products.

Urisu et al. [25] compared the allergenicity of heated egg white, freeze-dried egg white, and heated egg white depleted of OVO, recruiting 38 subjects with high levels of IgE antibodies for egg white. Twenty-one subjects (55%) with a positive challenge to freeze-dried egg white had a negative challenge to heated egg white. Sixteen out of 17 (94%) with a positive response to heated egg white did not respond to the heated and OVO-depleted egg white. This might indicate that OVO is the major determinant of causing egg induced allergies and that heating partially reduces the allergenicity of egg white. Escudero et al.

[26] compared the allergenicity of dehydrated egg white (DEW), a product that undergoes a double heat treatment (heating to 59 °C for 6 minutes and spray drying with hot air at 80 °C for 1 minute) with raw egg white (REW). Ten out of 40 egg-allergic patients (25%) had a positive OFC to both DEW and REW. The other 30 patients had a negative OFC to both forms. The allergenicity of commercially available DEW was therefore shown to be equivalent to raw egg whites and the processing of DEW did not affect the allergenicity of the egg proteins. Overall, from these studies, it can be concluded that the majority (50–85%) of children with egg allergy could tolerate extensively heated egg. However, milder forms of treatment (heating <80 °C) might still retain, to a great extent, the allergenic properties of the egg proteins. To look more into the protein allergenicity mechanisms, mouse models have been employed to compare unheated and heated purified egg white allergens. Mice sensitized and challenged with heated OVA (70 °C for 10 minutes) showed decreased clinical symptoms and a shift towards a Th1 response compared to mice sensitized and challenged with unheated OVA [27]. In addition, C3H/HeJ mice orally sensitized with native OVA and OVO were challenged with native and heated (30 minutes in boiling water) OVA or OVO. The native forms did induce symptoms of anaphylaxis; this in contrast to the unheated counterparts [28]. This decrease in allergenicity of heated OVA was shown to be partially the result of an enhanced gastrointestinal digestibility after heating [28,29] and of a reduced intestinal absorption of OVA and OVO molecules that are capable of triggering basophils and T cells [28].

In vitro studies assessing IgE-binding capacity showed that heating of OVA clearly decreased the IgE-binding capacity compared to the unheated OVA [30,31]. Heat treatment of OVO (95 °C, 15 min) lowered the IgE-binding activity of OVO. However, glycation by the Maillard reaction increased the IgE-binding. Ovotransferrin and lysozyme are less well studied, but generally regarded as heat-labile proteins. Jiménez-Saiz et al.[29,32] proved a decrease in IgE-binding after heating for 15 minutes at 95 °C. Limited studies looked at combined processing methods. One study [33] has been reported on the influence of combining various heat treatments with enzymatic hydrolyzes on the structure and allergenicity of pasteurized liquid whole egg. The remaining IgE-binding capacity of the end product, which underwent three heating and two enzymatic treatments, was more than 100-fold reduced compared to the untreated liquid whole egg. UV-C exposure (1.6 to 29.1Wm−2) of an egg white protein solution reduced the IgG-binding capacity of egg white proteins (ELISA), which was attributed to denaturation [34], while no difference in IgE-binding capacity (ELISA) was observed between egg white exposed to UV-C light

(10.6 and 63.7 kJm−2) and untreated egg white [35]. Gamma and electron beam radiation decreased both IgE and IgG-binding capacity of OVO [36]. Therefore, from these studies it emerges that: extensive heating diminishes the allergenicity of egg white proteins and the majority (50–85%) of egg allergic patients are tolerant to heated egg products. In addition, other treatment methods, such as irradiation, might modulate the allergenic properties of eggs; however, more investigation is needed.

1.2.2. Soy allergens

Soybean (*Glycine max*) seeds contain approximately 37% of protein, of which eight allergenic proteins (Gly m 1 to Gly m 8) have so far been registered by the International Union of Immunological Societies Allergen Nomenclature Sub-Committee (www.allergome.org). The major storage proteins β-conglycinin (also named Gly m 5 or 7S) and glycinin (also named Gly m 6 or 11S) represent 70% of the whole soybean protein and have been related to severe allergic reactions in European soy allergic subjects [37]. Recently, Gly m 8 (biochemical name 2S-albumin), was reported with high diagnostic value in soy allergic children in Japan [38]. In addition, soybean allergy can result from association to birch pollinosis. Clinical cross-reactivity between the major birch pollen allergen Bet v 1 and the homologous soybean allergen Gly m 4 has been described and sometimes associated with anaphylaxis [39]. The soybean hull allergens Gly m 1 (hydrophobic protein) and Gly m 2 (defensine) have been identified as aeroallergens in isolated asthma outbreaks [40] and are not considered as food allergens [41]. Various additional IgE-binding soy proteins, thus potential soy allergens, have been described. Of these, especially Gly m Bd30k (also named P34), a thiol-protease, might be a major allergen that could affect more than 50% of soy allergic subjects [42]. However, for some of these IgE binding proteins, the clinical relevance is unclear, since the underlying studies were based merely on soy sensitized subjects with unclear clinical reactivity.

Soybeans undergo various processing steps to obtain many different soy products such as soy flour, texturized soy protein (TSP), soy protein concentrates (SPC), soy protein isolates (SPI), protein hydrolysates, as well as fermented products in which the soy proteins and their structures may undergo various modifications. Allergenicity assessment of soy products has been primarily done by using antibody-based in vitro techniques and hardly by means of the more predictive tests such as the DBPCFC (Double-blind, placebo-

controlled food challenge). Comprehensive allergenicity assessment was done with only a few soy products, as for example soybean oil and lecithin. By means of immunoblotting and EAST inhibition experiments, the level of protein found in soy oil was low in comparison to that in soy lecithin.

Extracts from soy lecithin and non-refined oil still contained IgE-binding proteins, while refined soybean oil did not [43]. No allergic reactions were observed with commercially available soybean oils in a double-blind crossover study with seven subjects having a history of soy related systemic allergic reactions [44]. The panel on dietetic products, nutrition and allergies (NDA) of the EFSA considers that it is not very likely that fully refined soybean oil and fat will trigger a severe allergic reaction in susceptible individuals. Using sandwich ELISA (enzyme-linked immunosorbent assay) with plasma from soy allergic patients, soy protein isolate (SPI) and concentrate (SPC) showed less IgE-binding capacity than soy flour. The IgE-binding capacity of tofu was about 20-fold higher than that of soymilk using sandwich ELISA [45].

A few studies have investigated the effect of thermal processing on the IgE and IgG-binding capacity of soy and soy protein fractions. Burks et al. [46] in their experiments heated crude soy and its 7S and 11S protein fractions (80 °C or 120 °C, 60 min). Using sera from children with positive DBPCFC to soy, heating significantly reduced the IgE-binding capacity in ELISA. Some other studies [47,48] have been reported concerning a varying IgG-binding capacity in ELISA or concentration of soy protein in differently processed soy-based foodstuffs or after increased baking time in cookie matrix. Twin-screw extrusion of soybean meal with temperature higher than 66 °C was effective to decrease the binding capacity of specific IgG from calve, as assessed by ELISA analysis, to 0.1% of the original activity [49]. However, the effect was analyzed with animal antibodies and it may not be related to the process of heating only.

Hydrolysis with trypsin, pepsin and chymotrypsin is frequently used to prepare hypoallergenic formulas, but other enzymes of bacterial and fungal origin are also investigated. Artificial digestion of soy protein by pepsin, trypsin, chymotrypsin and intestinal mucosal peptidases was found to reduce the IgE-binding capacity by 10,000-fold using an ELISA inhibition assay [46]. Other studies investigated the degradation of individual soy allergens depending on the selection of enzyme, temperature, and pH. For example, Yamanishi et al. [50] found that the hydrolysis of Gly m Bd 30K was enzyme dependent and most successful for Proleather FG-F (protease from Bacillus subtilis) and Protease N. Tsumura et al. [51] confirmed elimination of Gly m Bd 30K with Pro leather

FG-F using immunoblot. In the same study, it was demonstrated that the presence of β-conglycinin was almost reduced, but no such effect on glycinin was seen, based on SDS-PAGE analysis. The enzymatic hydrolysis of glycinin and β-conglycinin with tryptic and peptic enzymes was also investigated [53–55]. The hydrolysis of both proteins depended on temperature and pH [53]. At low pH, glycinin was denatured and more susceptible to hydrolysis, while β-conglycinin was denatured at higher temperature and became more hydrolyzed in contrast to glycinin which was not affected. The IgG-binding capacity was never completely removed. Van Boxtel et al. [55] have studied the combined effect of enzymatic hydrolysis with heat treatment or high pressure. Fermentation of cracked soybean seeds and soybean flour by various mold strains and bacteria have been shown to reduce the IgE-binding capacity by 65 to 99%, as was investigated using indirect ELISA with human serum [56].

The application of chemical treatments, such as Maillard-type carbohydrate conjugation or transglutaminase treatment, showed some evidence of reducing the IgE production in mice [57] or IgG-binding to soy protein [58]. L'Hocine et al. [59] investigated the effect of ionic strength and pH on the IgG-binding capacity of purified glycinin. Changes in IgG-binding were related to changes in the secondary and tertiary packing of this soy protein. For example, a higher IgG-binding capacity was shown at low pH (2.2) and at neutral pH (7.2). For the purpose of preservation and structural modification, foods can be treated with high hydrostatic pressure (HHP). During HHP treatment, non-covalent bonds (hydrogen, ionic and hydrophobic bonds) are broken. A slight reduction in IgG-binding was observed after 15 min HPP treatment at 300 MPa [60]. The IgE-binding was also reduced by 44% using ELISA. Another non-thermal processing step is controlled pressure drop (DIC), in which food is subjected to a short (1–3 min) drop in pressure [61,62]. DIC treatment at 6 bar for 3 min almost abolished the IgE-binding capacity of soybean proteins according to immunoblot analysis [61]. However, aqueous extracts were investigated, but potentially low extractability of soy proteins after DIC treatment was not controlled. To sum up, the prediction of allergenicity of soybean and products thereof is limited because of a very limited number of high-quality studies performed in soy allergic humans or done with sera from clinically confirmed soy allergic donors. Although evidence exists that the allergenicity of soy may be reduced or retained by food processing, there has been no indication for increased allergenicity due to food processing. Apart from highly refined soybean oil and other soybean products in which the level of soybean proteins are reduced

below clinically relevant levels, one-step processing may not fully abolish soy allergenicity.

1.2.3. Fennel allergens

Fennel (*Foeniculum vulgare*) is a member of the *Apiaceae* (formerly called Umbrelliferae) family, a large group of plants encompassing approximately 300 genera and more than 3000 species. These species include some important allergenic plants, such as carrot and celery. Based on scientific evaluation and its use in traditional medicine, *Foeniculum vulgare* 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. Indeed, several pharmacological properties, both in vivo and in vitro, have been demonstrated including anti-microbial, anti-viral, anti-inflammatory, anti-mutagenic activities, etc. [63]. Fennel is usually consumed as seeds in Northern Europe, while in the Mediterranean area the plant is often consumed fresh. Because of the low consumption, a few studies have been reported dealing with the identification of fennel allergens in the literature. Some publications deal with the relationship between fennel allergy and birch and mugwort pollen allergy, in the so-called birch-weed or fruit-spice syndrome [64,65]. Jensen-Jarolim E. et al. [66] stated that Bet v 1 (17 kDa band) and profilin-related allergens (Bet v 2, 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. Similarly, the major allergens Api g 1 and Dau c 1, belonging to the *Apiaceae* plant foods, celery and carrot respectively, are Bet v 1 homologues [67].
In another study, Pastorello and co-workers studied the association between fennel and peach allergy [68]. The study aimed at investigating the clinical and immunological relationship between peach and fennel allergy and, therefore, at detecting and characterizing the putative allergens responsible for this relationship, since a high number of individuals with fennel allergy symptoms had been observed in a group of severe peach-allergic patients. In particular, sera from 25 fennel allergy patients were used for IgE immunoblotting analysis and IgE-binding proteins were detected by incubation with an ^{125}I-labeled anti-human IgE antibody. Sixty percent (60%) of patients' sera reacted toward an approximately 9 kDa band, whereas 44% of the sera recognized a protein of approximately 15 kDa and 96% reacted with bands in the range of 65-75 kDa. An

immunoblotting inhibition experiment was also performed to evaluate the cross-reactivity between fennel and peach extracts, using a pool of sera from patients selected on the basis of their immunoblotting response pattern to fennel proteins. This experiment revealed that pre-incubation of the pool with a peach extract at different dilutions completely inhibited IgE binding to fennel proteins, suggesting a high cross-reactivity between these two plant foods. Moreover, the paper demonstrated that the 9 kDa band could be identified, by Liquid Chromatography-tandem Mass Spectrometry (LC-MS/MS), as a lipid-transfer protein (LTP), characterized by a high homology with Pru p3 protein in peach, suggesting its importance in the cross-reactivity evaluation. For this reason, fennel allergy is considered an LTP-related food hypersensitivity and therefore fennel has been included in the list of foods that cause "LTP syndrome" [69]. More recently, Borghesan et al. [70] demonstrate for the first time the existence of a 60 kDa allergen in fennel, also recognized by the serum of patients with the mugwort-celery-spice syndrome. In this study, patient's IgE reactivity against mugwort and fennel extracts was investigated by immunoblot analysis under reducing conditions. They hypothesized the presence of two distinct proteins in a single fennel band at 60 kDa, presumably corresponding to an homologous protein of Api g5 and to the cofactor-independent phosphoglyceromutase.

1.3. Analytical methods for the determination of allergenic proteins

1.3.1. Traditional methods for food allergen analysis

The description of the pattern of individual allergens associated to each food is of crucial importance both for the diagnostic process and the setting of a safe diet for each allergic subject. Although some advances have been made during the last few years leading to the identification of new allergens in many allergenic sources and to their characterization, this knowledge is still fragmentary and does not allow the definition of a comprehensive pattern.

In this perspective, reliable analytical methodologies are required to assess the pattern of allergenic components really contained in the food products. The characterization of the allergen profile of a food implies the identification of all the potentially allergenic molecules contained in it. Since the allergenic source may contain more than one allergen, an in-depth characterization is required to describe the complete allergome by classifying

as “allergenic” or “not allergenic” the proteome components. Several clinical observations and literature reports suggest that the allergenic molecules identified so far are much fewer than those actually contained in the allergenic sources [71–74]. In addition, the listing of allergens currently known may be affected by some factors linked to the process of allergen identification; indeed, a high concentration in the natural source, and/or a high structural stability to proteolysis and high temperature, are two features frequently found in the best characterized food allergens.

The traditional, and still most common, skin prick testing methods are based on the use of commercially available raw protein extracts derived from allergy sources. However, it is well known that the allergen composition of extracts is very variable and their standardization appears impossible [75–81]. Ideally, a reagent used for the diagnosis of an allergy to a specific food should exactly contain all the potential allergens (all together in the raw food and extracts, or separated into individual purified allergens) of that food, and nothing more. To achieve this aim, two conditions should be fulfilled: (i) the entire profile of allergenic molecules contained in the allergenic sources should be known, and (ii) reliable protocols and methodologies, useful to assess the pattern of allergenic components really contained in the reagents used by the allergy test systems, should be available.

Current methods, typically utilized by food industries in allergen monitoring plans, employ an antibody-based recognition in the format of enzyme-linked immunosorbent assay (ELISA) and immunoblotting, because of the advantage of specificity of IgE against its specific allergen. Although this approach offers several advantages that contributed to the wide commercialization of different kits for the detection of single allergens, they could be affected by cross-reactivity phenomena, leading to false-positive results due to a possible interaction of the antibodies with the food matrix; that decrease the confidence in the results obtained [82–84]. In addition, food processing or sample preparation can also generate false negatives consequent to allergen modification that might mask the allergen itself from being recognized by the target antibody [85,86]. In view of the numerous drawbacks of the currently established methods for allergen analysis and in order to overcome such limitations, alternative non-immunological methods have been investigated in the last decade. Thanks to their high specificity, sensitivity, and accuracy, MS-based methods have been suggested as confirmatory tools for unambiguous identification and characterization of proteins and peptides [87,88]. Nevertheless, the isolation and identification of a new protein/allergen molecule need a concerted application of several technologies and the mass spectrometry measurements represent only the final step.

Therefore, it would seem to be important to improve the technologies supporting the sample preparation and analysis, including extraction and clean-up processes of the protein molecules from the allergenic source and their chromatographic separation.

Protein separation represents a key issue in proteomic analysis and two-dimensional gel electrophoresis (2D-E) [89,90] and liquid chromatography (LC) [91–93] are the most used techniques.

Two-dimensional gel electrophoresis is still the most widely used method in quantitative and qualitative proteomic studies and is the only technique that can resolve up to 10,000 protein species from large sets of complex protein mixtures. This technology separates the samples by two consecutive techniques: isoelectric focusing, which discriminates proteins based on their isoelectric point, followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), which discriminates proteins based on their molecular weight. Despite the high resolution power, 2D-E is extremely laborious, time-consuming, and more sensitive to technical reproducibility error, since large sets of gel repetition and sample are usually needed. Moreover, 2-DE technique also fails to resolve low abundant and hydrophobic proteins as well as those with molecular size out of the range of 5–150 kDa or with extreme pH range (< 3.5 and > 10) [94].

Some examples of studies have been recently reported for the detection of allergens in several foods, such as beer [95], beef [96], milk [98,[98], rice [99], and fish [100], by gel-based and gel-free approaches, coupled with LC-MS/MS. However, despite proteomic analysis of food allergens has become a key issue in the food safety field at present only fragmentary information on allergenic molecules contained in allergenic sources is available.

1.3.2. New trends in food allergen detection

Over the years, two-dimensional Liquid Chromatography (2D-LC) has played a key role in the field of proteomics for the analysis and isolation of proteins from complex food matrices, before their characterization by MS.

Since 90’s, two-dimensional liquid chromatography (2D-LC) has been added to the family of LC techniques [101,102]. The remarkable potential increase in peak capacity and resolving power represent a huge advantage of 2D-LC respect to one-dimensional (1D) methods [103,104]. Significant improvements in LC column technology and

instrumentation have been made in the last decade and the introduction of superficially porous particles has opened new possibilities in terms of speed and resolution. Orthogonality and coverage are the main issues to consider when choosing a useful pair of stationary phases and eluent conditions for successful 2D-LC separations [105]. Furthermore, respect to 1D-LC method development, an additional number of experimental factors have to be carefully evaluated, such as coupling systems between the two dimension columns, configuration mode and design of instrumental devices [106,107]. The interface between the columns typically consists of a multiport switching valve connected to sample loops or trapping columns, which allow to pre-concentrate the sample before being re-introduced in the second dimension column. This approach, called *active modulation*, is widely used in liquid chromatography for the separation of tryptic peptides [108,109] or different metabolites [106,110–113]. In nano-LC-MS/MS configurations, the use of a trap column (reversed phase C18 or strong cation exchanger) in an automated sample injection system is commonly applied to two- as well as one-dimensional peptide separation, allowing the injection of large volumes of sample in a short time [114–116]. Recently, multidimensional systems have been developed for the online peak fractionation and direct MS characterization of monoclonal antibody variants [117], allowing to collect up to six [118] or nine peak fractions in trap cartridges [119] from the first dimension column. Reduction and alkylation processes of the trapped protein peaks were on-column performed, before reversed-phase separations of the reduced proteins and entering the MS system, providing intact and chain-specific information for a rapid characterization of size and charge variants of biotherapeutics [118]. Very recently, a comprehensive two-dimensional liquid chromatography system consisting of twelve capillary monolithic columns in the second dimension was developed for a comprehensive protein analysis in biological samples [120].

In omic-type 2D-LC applications, the *comprehensive* mode is mainly used to analyze unknown mixtures of samples at high complexity, in order to obtain detailed information on untargeted components [121–125]. Each peak eluting from the first column is transferred to the second column through a collection loop device placed between the two columns. Then, for a sequential collection of aliquots from the first dimension and the subsequent reinjection onto a second column, the sampling process should be performed at high frequency to avoid remixing of analytes, just successfully separated in the first column. For these reasons, fast gradients, short columns and high flow rates in the second column have to be combined to reduced flows in the first dimension column. Indeed, the

duration of the ^{2}D cycle corresponds to the sampling time but, generally, such an operative 2D-LC approach determines the necessity of working under elution conditions that are sub-optimal in terms of flow-rates and gradient profiles for the selected first and second dimension columns.

Heart-cutting 2D-LC differs from comprehensive chromatography, because only one or a few target fractions are taken from the collection loop, before switching the valve for the second dimension separation [110,126]. In this way, the problems associated to the link between sampling time and ^{2}D cycle are solved and both dimensions can operate under optimal conditions, with longer ^{2}D gradients leading to a higher separation efficiencies. On the other side, such a gain in chromatographic quality causes a loss of information, since not all the ^{1}D peaks are reanalyzed in the second dimension. For these reasons, despite its use is known from more than two decades, 2D-LC technique can be considered an emerging technology, and instrumental and methodological implementations are still needed in particular in proteomic field on preparative scale.

The 2DLC, as a gel-free proteomic strategy, can be easily combined both to a bottom-up approach (after protein enzymatic digestion) and top-down MS, allowing to overcome the intrinsic limits of 2D-PAGE in recovering the proteins embedded in the polyacrylamide media as intact species [127].

Mass spectrometry represents the technique of election in proteomic studies mainly aimed at protein characterization and quantification. Thanks to the performance offered by the latest generation of mass analyzers, new efforts have been placed on the development of MS methods able to deliver both qualitative and quantitative information about allergenic proteins in food. Despite the need for expensive equipment and trained personnel, the chance to provide multiplexing and unequivocal allergen identification accounts for the overall strength of the MS-based approaches compared to previously established methods. Noteworthy, current knowledge in the allergen detection field suggests that the challenge to design a unique protocol feasible for different food matrices potentially contaminated by several allergens it still far to be accomplished. A more realistic objective would be to develop tailored approaches based on matrix similarity, whether rich in carbohydrate (such as bread, cookies, etc) or rich in fat (e.g. chocolate bar, chocolate dessert, etc). As far as MS-based allergen detection is concerning, two methodological options are to date available: i) detection of the intact protein representative of the allergenic ingredient, that is usually the most abundant in the proteomic profile; ii) detection of the target analytes, namely markers, that are signature peptides, properly selected, resulting from the

enzymatic digestion of the whole allergenic ingredient. In both cases, the sampling is a crucial step to provide a proper representativeness of the analysis, and for an accurate absolute quantification, the protein/peptide content should refer to that of a suitable standard either the whole protein or a derived peptide likely isotopically labeled. The availability and eventual costs of such standards, together with their ionization efficiency are often the main drivers guiding the choice between the two aforementioned approaches.

Fundamentally, MS measures the mass-to-charge ratio (*m*/*z*) of gas-phase ions. Mass spectrometers consist of an ion source that converts analyte molecules into gas-phase ions, a mass analyzer that separates ionized analytes based on *m*/*z* ratio, and a detector that records the number of ions at each *m*/*z* value. The development of electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), the two soft ionization techniques capable of ionizing peptides or proteins, revolutionized protein analysis using MS. The role of the mass analyzer is central to MS technology. For proteomics research, four types of mass analyzers are commonly used: ion trap (quadrupole ion trap, QIT; linear ion trap, LIT or LTQ), time-of-flight (TOF) mass analyzer, Orbitrap and Fourier-transform ion cyclotron resonance (FTICR) mass analyzers. Different coupling systems with various mass analyzers are to date available such as ESI-qTOF, ESI-IT, or MALDI-TOF more addressed to qualitative investigations and protein/peptide characterization [128–130]. Quantitative triple quadrupole and ESI-IT systems have the advantage of identification and quantification through fragmentation settings in the MS collision cell [131]. Few data are reported in literature about the use of FTICR mass spectrometry in proteomic analysis. It is quite expensive, but represents a prominent technology for high throughput analysis, providing the highest resolving power and mass measurement accuracy. Moreover, the large dynamic range and unmatched sensitivity of FTICR-MS currently provides the highest quality data for protein identification. In FT-ICR mass spectrometry, the determination of the mass-to-charge ratio (m/z) of ions is based on their cyclotron frequency in a fixed magnetic field. The cell is the heart of the system, where ions are stored, mass analyzed and detected. The general steps of an FT-ICR MS experiment are: (1) ion formation outside of the detector; (2) ion focusing and accumulation; (3) transportation of ions into a Penning trap; (4) selection of ions based on mass-to-charge ratio and ejection of these ions from the Penning trap; (5) excitation; (6) detection; (7) fast Fourier transform of the digital time-domain signal; (8) conversion of frequency to mass-to-charge ratio. One of the biggest advantages of FT-ICR-MS is its unparalleled mass resolution. In addition, mass measurements can be made on highly complex chemical or

biological samples, without the need for any separation method, such as gas or liquid chromatography. It is often combined to an ESI technique to produce ions using a high voltage applied to a liquid to create an aerosol. It is especially useful in producing ions from macromolecules (such as proteins) because it overcomes the propensity of these molecules to fragment when ionized. ESI is different from other ionization processes (e.g. MALDI) since it may produce multiple-charged ions, effectively extending the mass range of the analyzer to accommodate the kDa-MDa orders of magnitude observed in proteins and their associated polypeptide fragments. Protein identification via MS is usually carried out in the form of whole-protein analysis ("top-down" proteomics) or analysis of enzymatically or chemically produced peptides ("bottom-up" proteomics). The most widely applied method for protein digestion involves the use of enzymes. Many proteases are available for this purpose, each having their own characteristics in terms of specificity, efficiency and optimum digestion conditions. Trypsin is most widely applied in bottom-up proteomics and can be considered as the gold standard in proteomics, cleaving the peptide bonds C-terminal to the basic residues Lysine (Lys) and Arginine (Arg), except when followed by Proline (Pro). The advantageous properties of tryptic peptides lead to high quality MS/MS fragmentation spectra and confident peptide identification in protein database searches. This, in turn, increases the accuracy of inference of protein identity. Standardized protocols have been described for in-solution and in-gel protein digestion. A typical protocol involves denaturation of the protein using chaotropic agents like urea or guanidine, reduction of disulfide bridges using dithiothreitol (DTT), and subsequent alkylation of the cysteines by iodoacetic acid or iodoacetamide. After reagent removal and buffer exchange, the trypsin digestion is typically performed at neutral pH in an ammonium bicarbonate buffer at 37 °C. Depending on the way the digestion is performed, it may take up to 18 h (overnight digestion). The experimental conditions for trypsin digestion can be optimized for a specific application, for instance using a design of experiments approach. The digestion is stopped by the addition of (formic) acid. Despite the many advantages of trypsin, it may be necessary to use other proteases in specific cases, such as a lack or an over-abundance of Lys and Arg in the protein sequence or pH incompatibility. A wide range of alternative proteases is available with different cleavage specificities [132].

1.4. Aim of the research project

Considering the increased attention on the detection of allergenic food ingredients over the recent years, reliable analytical methodologies are required to assess the pattern of allergenic components really contained in the food products. The exact determination of the real allergens is not an easy task and the risk of identifying as allergenic molecules the most abundant and/or the most stable proteins in the food extract is very high. Indeed, the list of allergens, identified so far and reported in the allergome database, includes the most abundant compounds present in the natural source, which could mask the effective under-expressed proteins, really responsible for the allergenic response.

Aim of the present Ph.D. project is the development of analytical methods by Two-Dimensional Liquid Chromatography (2D-LC) and Mass Spectrometry for the identification of proteins as putative allergens in food samples of animal and plant origin. To this purpose, automated two-dimensional liquid chromatography (2D-LC) methods were developed for the separation of intact proteins from food allergenic sources such as eggs and soy flour, included in the Big-8 listing. Since the interface between the two separation dimensions in a 2D-LC configuration is the heart of the system, the trapping devices, the design of the switching valve and its connection tubing have been tested to get the best compromise between resolution and analysis time. Then, the optimization of the experimental conditions, associated to each one –dimensional (1D) separation mode, have been carefully evaluated in order to take full advantage of each separation mechanism and improve the resolving power. An *active modulation* 2D-LC approach was proposed based on the use, as an interfacing system, of a reversed phase guard column installed on switching valve placed between the two dimensions. The ideal match for protein determination by separation modes at high orthogonality was chosen through the combination of size exclusion (SEC) and reversed phase (RP) liquid chromatography. Therefore, SEC was selected as a first dimension for separating proteins on the basis of their size in solution (directly correlated to their molecular weight), while RP, for protein separation by the length of the hydrophobic tail, was placed in the second dimension (because of the compatibility with Mass Spectrometry). An automated fraction collector was connected to the UV detection cell in order to isolate the protein peaks coming from the effluent of the second column. The optimized separation workflow, as a good compromise between the comprehensive and multiple heart-cutting modes, takes on the

challenging task of analyzing complex food matrices rich in proteins greatly different in concentration, mass and polarity, ensuring high resolving power and automation grade.

Finally, a rapid, sensitive and shot-gun method by Fourier transform ion cyclotron resonance mass spectrometry was developed for the protein identification in fennel samples, recently recognized as allergenic source in the Mediterranean area. The protein profile of the fennel extract was also examined by immunoblotting analysis in order to identify fennel protein bands that react with IgE from sera of *Foeniculum vulgare* allergic patients.

2. MATERIALS AND METHODS

2.1. Chemicals

Standard of proteins were supplied by Sigma Aldrich (Steinheim, Germany): Bovine Serum Albumin (98%), β-lactoglobulin from bovine milk (≥90%), glucose oxidase from *Aspergillus Niger* (≥65%), IgA from human serum (≥95%), thyroglobulin from bovine thyroid (≥90%), myoglobin from equine heart (≥90%), ubiquitin from bovine erythrocytes (purity grade ≥ 98%, average mass 8565 Da) and trypsin from porcine pancreas (proteomics grade, BioReagent, dimethylated). Solvents and reagents used for the preparation of mobile phases and for protein extraction were also purchased from Sigma Aldrich: water (Chromasolv®, for HPLC), hexane (Chromasolv® plus, for HPLC, ≥ 98,5%), acetonitrile (Chromasolv® plus, for HPLC, ≥99,9%), sodium dihydrogen phosphate dihydrate, sodium chloride, sodium hydroxide, hydrochloric acid, ethylenediaminetetraacetic acid (EDTA), trifluoroacetic acid (TFA, ReagentPlus®, 99%), polyethylene glycol 200 (PEG 200, average molecular weight 200), calcium chloride, hydrochloric acid, Trizma® base, Trizma® hydrochloride, formic acid, 1,4-dithiothreitol, iodoacetamide and ammonium bicarbonate. Water, methanol, and acetonitrile for MS analysis (LC-MS CHROMASOLV®, ≥99.9%) were from Fluka. Modified trypsin (porcine) and chymotrypsin (bovine) were purchased from Princeton Separations (Adelphia, NJ, USA).

For each protein, individual stock solutions at a concentration of 1000 mg L^{-1} were prepared in water and stored at -18 °C. Working standard solutions and multi-component standards were prepared by dilution with water and stored at 5 °C between injections.

2.2. Liquid chromatography system

Chromatographic separations were performed on an Ultimate 3000 LC system, (Thermo Fisher Scientific, Waltham, MA, USA), which consisted of two ternary pumps for gradient elution equipped with micro vacuum degassers, a thermostated autosampler connected to an injection valve with a loop of 100 µL, a column compartment, two 10-port switching valves, a multiple wavelength UV-Vis detector and an automated fraction collector (AFC). Acquisition and data processing were performed by software Chromeleon™, version 6.8

(Thermo Fisher). For the first dimension separation, two size-exclusion columns coupled with the Security Guard Cartridges were used: Yarra™ SEC-2000 (300×7.8 mm, Phenomenex, Torrance, CA, USA) and MAbPac SEC-1 Analytical Column (300×4 mm, Thermo Fisher Scientific). A reversed phase *widepore* column (Aeris™ widepore XB-C18 150×4.6 mm, 3.6 μm, Phenomenex) was used for the second dimension separation, based on core-shell particle technology. A security guard column ULTRA Cartridges UHPLC WIDEPORE C18 for 4.6mm (AJ0-8769, 2 × 4.6 mm I.D. with sub-2 μm particles, Phenomenex) and a C4 column (Eurosil Bioselect 300-5 Vertex Plus Column, 50 x 4.6 mm, Knauer, Berlin, Germany) were used as a collection trap system to focus the proteins coming from the first dimension column, before their selective elution in the second dimension. The connection tubings (Viper™ Capillary Stainless Steel Fingertight Fittings, Thermo Fisher Scientific) were 0.180 mm I.D. and kept at the shortest length possible.

2.3. Protein analysis in egg samples

2.3.1. Protein extraction from lyophilized egg samples

Proteins from lyophilized egg, purchased in local supermarkets, were extracted as already reported [133,134], with slight modifications. Briefly, 60 mL of n-hexane (15 mL/g) were added to 4 g of lyophilized sample. After sonication for 30 min and centrifugation at 8000 rpm (room temperature) for 10 min, the supernatant was discharged, while the pellet (defatted sample) was air-dried, until constant weight. Then, 2 g of defatted sample were suspended in 40 ml of distilled water and incubated at 4 °C for 16 hours (water extraction). After centrifugation at 9000 rpm for 20 min at 4 °C, the supernatant was brought to room temperature, while 30 mL of 0.1 M HCl (solvent ratio of 1:15, w:v) was added to the pellet in an ultrasound bath for 30 min, followed by stirring at 4 °C for 1 hour (acidic extraction). After centrifugation at 9000 rpm at 4 °C for 20 min, the supernatant was brought to room temperature and the pellet was subsequently treated with 30 mL of 0.1 M NaOH (solvent ratio of 1:15, w:v). After sonication for 30 min, stirring at 4 °C for 1 hour (alkali extraction) and centrifugation at 9000 rpm at 4 °C for 20 min, the pellet was kept and stored at room temperature. The supernatants obtained after water, acid and alkali extractions were combined and subjected to isoelectric point precipitation by adjusting the pH at 4.5 with 1.0 M HCl. The mixture was stirred for 5 min at 4 °C and the separation of

the precipitate was carried out in a centrifuge (8000 rpm, 10 min at 4 °C). Then, the supernatant (final protein extract) was filtered on 0.2 μm regenerated cellulose membrane (Phenex™-RC syringe filters, Phenomenex, Torrance, CA, USA) before chromatographic injection. The pellets isolated from alkali extraction and after isoelectric point precipitation were air-dried, re-suspended in 0.1 M phosphate buffer solution and sonicated for 10 min. Insoluble particles were removed by centrifugation at 11000 *rpm* for 10 min at 4 °C, while the supernatant solutions were filtered on cellulose membrane before injection. Therefore, for each sample, the three protein extracts (the final supernatant after isoelectric point precipitation and the solutions following the re-dissolution in phosphate buffer of pellets from alkali and isoelectric point precipitation) were analyzed separately.

2.3.2. Instrumental set-up for chromatographic separation of egg proteins

A reversed phase guard column placed on a single switching valve was chosen as an interfacing system. For the first dimension separation, the size-exclusion column Yarra™ SEC-2000, packed with 3 μm ultra-pure silica particles densely bonded with a proprietary hydrophilic surface chemistry, was used, coupled with the Security Guard Cartridge GFC-2000. The reversed phase *widepore* column Aeris™ XB-C18 was used for the second dimension separation, based on core-shell particle technology. A security guard column ULTRA Cartridges UHPLC WIDEPORE C18 for 4.6mm was placed on a 10-port switching valve and used as a collection trap system to focus the proteins coming from the first dimension column, before their selective elution in the second dimension. The mobile phases consisted of 0.1 M phosphate buffer (pH 6.8) containing 0.1 M sodium chloride (eluent A, for the first dimension separation), 0.1% TFA in water and 0.1% TFA in acetonitrile (eluent B and C, for the second dimension separation). An optimized ternary gradient elution program was developed by analyses of protein standard mixtures of bovine serum albumin (BSA), β-lactoglobulin (β-LG) and glucose oxidase (GOx). Protein detection was performed at a wavelength of 214 nm with a data collection rate of 100 Hz. One-dimensional SEC separations (1D-SEC) were performed on the Yarra column coupled with the guard cartridge, under isocratic conditions with 0.1M phosphate buffer + 0.1 M NaCl at pH 6.8 as the eluent at a flow rate of 0.7 mL min^{-1}. For the quantitative analyses in protein extracts, calibration data were obtained by three series of 1D-SEC analyses on

three different days, by injecting five working standard solutions of BSA, GOx and β-LG, each at concentrations of 10, 50, 100, 300 and 500 mg L^{-1}. A CHRIST rotational vacuum concentrator (model RVC 2-18 CD plus) was used for solvent evaporation of the protein fractions isolated by 2D-LC separations.

2.3.3. Shotgun analysis of egg powder by nano-LC-ESI-IT-MS/MS

In-solution enzymatic digestion of lyophilized egg sample. 2.1 mg of lyophilized egg were mixed with 100 µL of 50 mM ammonium bicarbonate. Then, to an aliquot of 10 µL of the obtained solution, 50 µL of 50 mM ammonium bicarbonate were added. After addition of 2 µL of 0.1 M dithiothreitol, the solution was incubated for 30 min at 50 °C and then cooled down at room temperature. A volume of 4 µL of 0.1 M iodoacetamide was added and the solution was incubated in the dark at room temperature for 40 min. Afterward, the mix trypsin/chymotrypsin prepared in 25 mM ammonium bicarbonate was added at a 1:50 (w:w) ratio. After an overnight incubation at 37 °C, the digestion was stopped by adding 5 µL of a 5% FA water solution.

LC-MS/MS analysis. The analyses for the LC-MS/MS characterization of the egg peptide mixture coming from the enzymatic digestion were performed by a nanoLC apparatus, Ultimate 3000 (Dionex LC-Packings, Amsterdam, The Netherlands), which consisted of an autosampler, a low pressure gradient micro-pump series equipped with flow managers, a column thermostat and an UV detector set at 214 nm. The UV flow cell was connected to an ESI-Ion Trap HCT ultra ETD II Basic System (Bruker Daltonics Srl, Bremen, Germany). The nanoLC-ESI-IT-MS/MS system was controlled by software Chromeleon CHM-1 (Dionex) and Hystar 2.3 (Bruker Daltonics). A PepMap C18 nano trap column (300 µm i.d. x 5 mm, 5-µm particle size, LC Packings) was used for concentrating and desalting the injected sample. Chromatographic separations were carried out by a PepMap C18 analytical column (15 cm length x 75 µm i.d., 3-µm particle size, 100-Å pore diameter; LC Packings). The mobile phase consisted of 0.1% FA in water (eluent A) and ACN/0.1% FA in water (80:20 v/v) (eluent B). Sample elution through the analytical column was obtained at a flow rate of 0.300 µL min^{-1}. An optimized gradient program was applied as follows: 6 min isocratic step at 96% A and 4% B; 120 min linear gradient to 10% A and 90% B; 10 min isocratic step at 10% A and 90% B; 1 min with the initial

mobile phase composition (96% A and 4% B), where the system was re-equilibrated for 43 minutes, with a total run-time of 180 minutes. The microliter pick-up injection mode was selected, and 4 µL of sample were mixed to the eluent A to overfill the 10 µL sample loop. A 10 port switching valve was used to combine sampling and switching functions. Then, the gradient elution was coupled with a single injection mode; after 6 minutes from the sample injection, the 10-port valve was switched in order to connect online the trap column with the analytical column.

Mass spectrometry analyses were performed by using a high capacity ion trap, coupled to a nano-ESI source, operating in the positive ion mode. The following parameters were set: spray voltage 4.5 kV; sheath gas (nitrogen) flow rate 10 L/min; capillary voltage 1.5 V; heated capillary temperature 160 °C. A full scan MS acquisition in the 300-1500 m/z range was performed with the acquisition of Base Peak Chromatogram (BPC) and Total Ion Current (TIC) profile. MS/MS analysis by Collision-Induced Dissociation (CID) was performed using unattended data-dependent acquisition mode and the auto-MS/MS event (scan range 100-2400 m/z) was carried out with a number of precursor ions of 3. The minimal signal required for precursor ion selection was set to an absolute threshold of 10000. The Ion Charge Control (ICC) was on and set at a target of 200,000 and a maximum accumulation time of 100 ms. The duty cycle time was 180 ms and 220 ms for the MS and MS/MS mode, respectively. The fragmentation was performed by activating the options MS/MS fragmentation amplitude (MS/MS FragAmplTM 1.00 V) and smart fragmentation (SmartFragTM 30-200%).

Data analysis and database search. Extraction of mass spectra peak-lists from chromatograms, mass annotation and deconvolution were performed by using Data Analysis 4.0 (Bruker Daltonics). The acquired MS and MS/MS datasets were submitted to database searches by using Biotools 3.2 (Bruker Daltonics) and MASCOT search engine (Matrix Science, London, UK). The data analysis files were used to search entries under the Chordata category of the Swiss-Prot database, assuming that peptides were monoisotopic and carbamidomethylated at cysteine residues. A maximum number of 2 missed cleavages were allowed and, for both precursor peptide ion and MS/MS tolerance, a peptide tolerance of 0.3 Da was set in the error window to match the peptide mass values. The option "*automatic error tolerant*" search was checked. Protein identification was accepted when MASCOT search results delivered scores higher than the identity threshold ($p<0.05$). Peptide matches above the identity threshold were submitted to a post-database

search validation by a manual inspection of the corresponding MS/MS spectra, in terms of peptide ion score (PIS), rank and normalized delta score (nΔs, i.e. the difference between the best and the second best ion score, divided by the best score) [135].

2.4. Protein analysis of soy flour samples

2.4.1. Protein extraction from soy flour samples

Hydrophilic proteins from soy flour samples were performed following the procedure of Ge et al. [133] that has been slightly modified. Briefly, proteins were extracted by adding 22 mL of 25 % (w/w) of PEG 200 aqueous solution to 1 g of sample. After sonication for 60 min at 45 °C and centrifugation at 8000 rpm at 4 °C for 10 min, the supernatant was kept and stored at room temperature. A solid-phase extraction (SPE) procedure on TELOS C18 (EC, 6 mL/500 mg of sorbent, Kinesis) was applied for the sample clean-up. A Visiprep^{TM} DL SPE Vacuum Manifold (Supelco) equipped with 24 flow control valves was used to provide a flow rate of 1-2 drops/second during the steps of sample loading and elution, to ensure optimal retention and compound desorption, respectively. An aliquot of 3 mL of the protein extract was loaded on the cartridge previously activated by sequential treatment with 3 mL of MeOH and 3 mL of water. After washing with 3 mL of water, the elution was performed with 3 mL of 80:20 MeOH/0.1% TFA. The eluate was evaporated to dryness at 30 °C by a CHRIST rotational vacuum concentrator (model RVC 2-18 CD plus). Finally, the residue was solubilized in water, filtered on 0.2 µm regenerated cellulose membrane (Phenex™-RC syringe filters, Phenomenex, Torrance, CA, USA) and then injected.

2.4.2. Instrumental set-up for chromatographic analysis of soy flour extracts

The two-dimensional liquid chromatography was coupled with an active modulation interface, based on the use of a double switching valve system (left valve, LV; right valve, RV) combined to a short C4 analytical trapping column. For the first dimension separation, a size-exclusion column MAbPac SEC-1 Analytical Column coupled with the Security Guard Cartridge GFC-2000 was placed on the right switching valve. The reversed-phase

core-shell column Aeris™ widepore XB-, coupled with a security guard column ULTRA Cartridges UHPLC *widepore* C18 was used for the second dimension separation. The C4 column (Eurosil Bioselect 300-5 Vertex Plus Column) was placed on the left switching valve and used as a collection trap system to focus the proteins coming from the first dimension column, before their selective elution in the second dimension. The mobile phases consisted of 50 mM phosphate buffer (pH 6.8) containing 0.3 M sodium chloride (eluent A, for the first dimension separation), 0.1% TFA in water (eluent B, for the trap washing step and the second dimension separation) and 0.1% TFA in acetonitrile (eluent C, for the second dimension separation). An optimized ternary gradient elution program was developed by analyses of protein standard mixtures. Protein detection was performed at a wavelength of 214 nm with a data collection rate of 100 Hz. For the quantitative analyses in protein extracts, one-dimensional SEC separation (1D-SEC) were performed in the MabPac column coupled with the guard cartridge, under isocratic conditions with 50 mM phosphate buffer + 0.3 M NaCl at pH 6.8 as an eluent at a flow rate of 0.25 mL min^{-1}. Calibration data were obtained by three series of 1D-SEC analyses on three different days, by injecting seven working standard solutions of BSA (injection volume 15 µL) at concentrations of 5, 10, 50, 100, 300, 500 and 1000 mg L^{-1}. Analogously, at the same concentration levels, calibration lines in three different working sessions were also recorded for BSA by 1D-RP separation on Aeris™ column. The optimized 1D-RP elution gradient was the following: 2 min isocratic step at 100% B; a convex gradient of 18 min by curve 3 of PeakNet software to 40% (B) and 60% (C) coupled with an isocratic step of 2 min; 5 min to 20% (B) and 80% (C) by the concave curve number 9 followed by an isocratic step of 2 min; 1 min to the initial mobile composition phase composition, at which the system was re-equilibrated for 8 min. The injection volume was 35 µL.

2.5. Protein analysis of fennel samples

2.5.1. Protein extraction from fennel samples

An in-house semi-purified 100,000 x g supernatant fennel extract was produced 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, on ice, in the presence of phosphate buffer solution (PBS) containing ions 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 hours, 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 was 3.5 mg/mL.

2.5.2. Enzymatic digestion

For the bottom-up MS analysis, the in-solution enzymatic digestion of the ubiquitin standard (UBIQ) solution and of fennel extracts was performed as follows. To a volume of 300 µL of 1000 mg L^{-1} UBIQ standard, calcium chloride, as a stabilization agent against thermal and proteolytic degradation, and autolysis phenomena [136] was added to get a final concentration of 20 mM. 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 hours and then stopped by the addition of 5% FA water solution (5µL). Final mixtures were stored at −20 °C until the analysis.

For the in-solution enzymatic digestion of the raw fennel extract, two different protocols were adopted:

- Protocol A (6hh): to a volume of 300 µL of extract, calcium chloride and Trizma buffer solution were added. Then, the enzyme trypsin prepared in water was added at a 1:30 (w:w) ratio. After a 6-hour incubation at 37 °C, the digestion was stopped by adding 5 µL of 5% FA.
- Protocol B (18hh), performed following the procedure of Khodadadi et al. with slight modifications [137]: 400 µL of methanol, 100µL of chloroform and 300 µL of water were added to 100 µl of protein extract and mixed thoroughly. After centrifugation at

15000 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 15000 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 hours of incubation at 37 °C. The resulting peptides mixtures were acidified with 5% formic acid (pH < 3) and centrifuged at 15000 rpm for 15 min.

Before ESI-FTICR mass spectrometry analyses, the peptide mixtures obtained by both protocols were diluted 1:10 in a mixture acetonitrile/water (70:30, v/v) containing 0.2% FA.

2.5.3. 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). Samples were injected with a flow rate of 2 µL min^{-1}. The MS was calibrated with a 5 mg/L-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 approximately 500–3000 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. The detection range was 500–3000 Da. Tryptic digests of fennel proteins were analyzed by direct flow injections by the use of Electrospray Ionization (operating in the positive ion mode) double Quadrupole-Fourier Transformation-Ion Cyclotron Resonance/mass spectrometry (ESI qQ-FT-ICR-MS).

2.5.4. 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). Fennel FT-ICR mass spectra were exported

to peak lists at a signal to noise ratio (S/N) of 4 and with an intensity threshold of 10^6 arbitrary units. 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 from the monomer until reaching m/z 1220 heptamer cluster ion. A methanolic 5 ppm arginine was utilized for FT-ICR-MS calibration purposes. A subset of NCBI protein sequences was derived to represent protein sequences found previously in fennel species and a custom-made protein database was generated. 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. 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 [i]. The following parameters were set: two allowed missed cleavages; mass range from 0 to unlimited Dalton; cysteines treated with nothing (for the analysis of UBIQ and fennel extract digested according to the protocol A, 6 hh) or with iodoacetamide (for the fennel extract digested according to the protocol B, 18 hh). 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 mass list of all the putative m/z ions 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 FT-ICR-MS of the isolated fennel protein mixture, which was digested with trypsin. A 5 ppm mass tolerance search was used and many relevant fennel proteins could be identified. Finally, in order to confirm the protein identity, the matched experimental mass datasets were submitted to database searches by using the MASCOT search engine (Matrix Science, London, UK). The Peptide Mass Fingerprint 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 0.005 Da and 5 ppm was set in the error window; no variable and fixed modifications or carbamidomethylation at cysteine residues were selected for the fennel extract digested by protocol A and B, respectively.

3. RESULTS AND DISCUSSION

3.1. Development and optimization of analytical methods for protein analysis in egg samples

In the present Ph.D. project, an *active modulation* 2D-LC approach was proposed and optimized by using, as an interfacing system, a reversed phase guard column placed on a single switching valve. Indeed, to the best of our knowledge, no data concerning the use of the active modulation interface for the isolation of intact proteins in food samples on preparative scale by 2D-LC are reported, representing an area of research that could deserve more attention in the future. The design of the trapping device between the two dimensions was tested to get the best compromise in terms of resolution and analysis time. Proteins coming from the first column were focused at specific time frames, before their elution in the second dimension column. A number of practical factors have been considered during 2D-LC method development, first of all the selection of two separation modes. The separation mechanism in the first dimension should be orthogonal to the mechanism in the second dimension in order to obtain the best resolving power. The most commonly applied chromatographic mode for the first dimension protein separation is the size exclusion chromatography [138], also used in the present project. Reversed phase liquid chromatography was selected as a second dimension mechanism, since the use of a solvent system of acetonitrile/water allows to desalt the eluates from the first dimension, making the sample suitable for direct MS analyses. Therefore, the ideal match for protein determination by separation modes at high orthogonality was chosen through the combination of size exclusion (SEC) and reversed phase (RP) liquid chromatography. An automated fraction collector was connected to the UV detection cell in order to isolate the protein peaks coming from the effluent of the second column.

3.1.1. Optimization of chromatographic separation conditions of one dimensional (1D) methods based on 1D-SEC and 1D-RP chromatography

The optimization of the experimental conditions associated to each one–dimensional (1D) separation mode was performed with standard BSA, β-LG and GOx, as a protein mix

model. The chosen proteins are structurally different, so displaying secondary interactions over a wide range of the physicochemical space. In fact, they cover a pI range between 4.8 and 5.8, a molecular weight from 20 to 140 kDa and a hydrophobicity grade (estimated as the number of the hydrophobic residues divided by the total number of the amino acids in the mature protein sequence) ranging from 36% to 52%. Then, the development of two different one-dimensional methods (1D-SEC and 1D-RP) was planned. For each 1D approach, the key factors affecting the protein separation, including column features, particle size and technology, mobile phase composition, gradient elution and flow-rates, were evaluated to achieve the best results in terms of peak efficiency, resolution and analysis times.

First dimension separation based on size-exclusion chromatography. For the first dimension separation, a size-exclusion column, packed with 3 µm ultra-pure silica particles, was selected with a pore size of 300 Å that allows smaller species to enter the silica beads and, depending on their apparent size, to be eluted in order of decreasing size. Despite being a seemingly simple isocratic method for separating biomolecules according to their hydrodynamic radius [138], secondary ionic interactions between the stationary phase and proteins could occur [138,139]. Therefore, although SEC bonded stationary phases are designed to minimize interactions between negatively-charged silanol groups on the silica surface and basic proteins, the presence of ionic secondary interactions can dramatically impact chromatographic performances [140]. Then, the use of buffer salts in the mobile phase is necessary in order to influence positively the retention of proteins by modulating and reducing ionic interactions. In addition, both recovery and peak shape improve as salt concentration increases [139]. Unfortunately, the ionic interactions are not the only kind of secondary interactions that occur during size exclusion chromatography by gel filtration. Bonded phases typically have some diol ligands that cover the silica surface and create a weak hydrophobic interaction media. At increased salt concentrations in the mobile phase, hydrophobic interactions occurring between the bonded phase and hydrophobic proteins can be detrimental to separation. As a result, increasing salt concentration in the mobile phase, on the one hand, the undesired ionic interactions are reduced, on the other hand hydrophobic interactions increase, then making method development for SEC separation a balance between the two mechanisms.

Efforts were undertaken to estimate the influence that phosphate buffer mobile phases at neutral pH (6.8) and different concentrations (50-100 mM) can have on SEC separations.

The best results in terms of resolution between the protein couples GOx/BSA and BSA/β-LG were obtained with phosphate buffer at a concentration of 100 mM. The addition of sodium chloride (0-100 mM) or sulfate (100 mM) was tested to evaluate the impact of secondary interactions on the retention behavior. As the salt concentration in the mobile phase increases, the retention times of three proteins slightly increase (~ 0.3 min for GOx and β-LG, 0.8 min for BSA). Moreover, both peak shape and protein recovery improve, confirming that ionic interactions between the stationary phase and proteins did not take place at a higher ionic strength. This hypothesis was also corroborated by the behavior of the retention time versus the log of the molecular weight for the proteins in the mixture, obtained under different isocratic elution conditions. Indeed, a possible source of error in the calibration curve is the non-ideal adsorption that may alter the retention volume, therefore the slope of the calibration curve in the linear portion is a measure of the stationary phase selectivity [141]. The regression parameters of data acquired by using phosphate buffer in absence and in presence of 100 mM sodium chloride were then calculated. The determination coefficients of 0.9957 and 0.9324 observed when NaCl is present or absent in the mobile phase, respectively, confirmed the capability of buffer salts to minimize the secondary interactions. The log plot of the molecular weights of standard proteins against the retention time, as shown in Figure 2, is also a useful tool for the estimation of the molecular weight of unknown proteins in food samples.

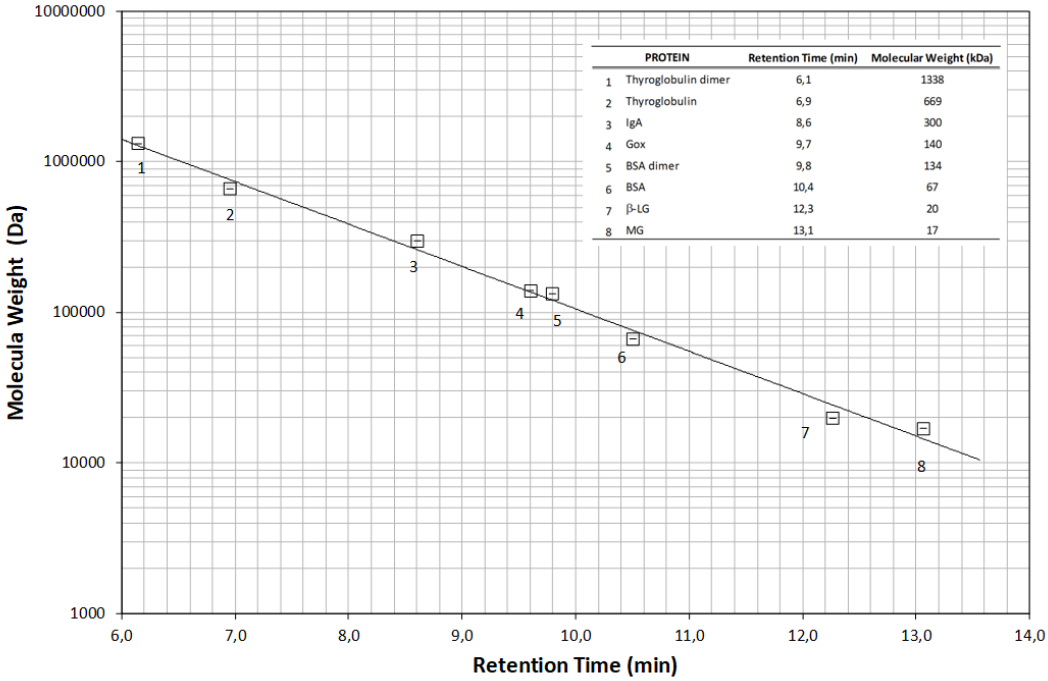


Figure 2. Protein mixture calibration by One-Dimensional SEC separation (1D-SEC). Column: Yarra™ SEC-2000 300×7.8 mm + Security Guard Cartridge GFC-2000 4 x 3.0mm ID. Eluent: 0.1M phosphate buffer + 0.1 M NaCl at pH 6.8, isocratic mode. Flow rate: 0.7 mL min^{-1}. Injection volume: 100 µL.

Finally, the separation conditions were optimized in terms of flow rate, explored in the range 0.5-1.0 mL min^{-1}, under step gradient profiles; a good compromise between resolution and analysis time was obtained at a constant flow of 0.7 mL min^{-1}. As an example, Figure 3 shows the separation by size exclusion of the standard protein mix at a concentration of 500 mg L^{-1}, under the optimized mobile phase composition. Despite of the low purity grade of GOx (as reported on the product label), a good chromatographic separation was observed with high efficiency for BSA and β-LG and a resolution of 3.3.

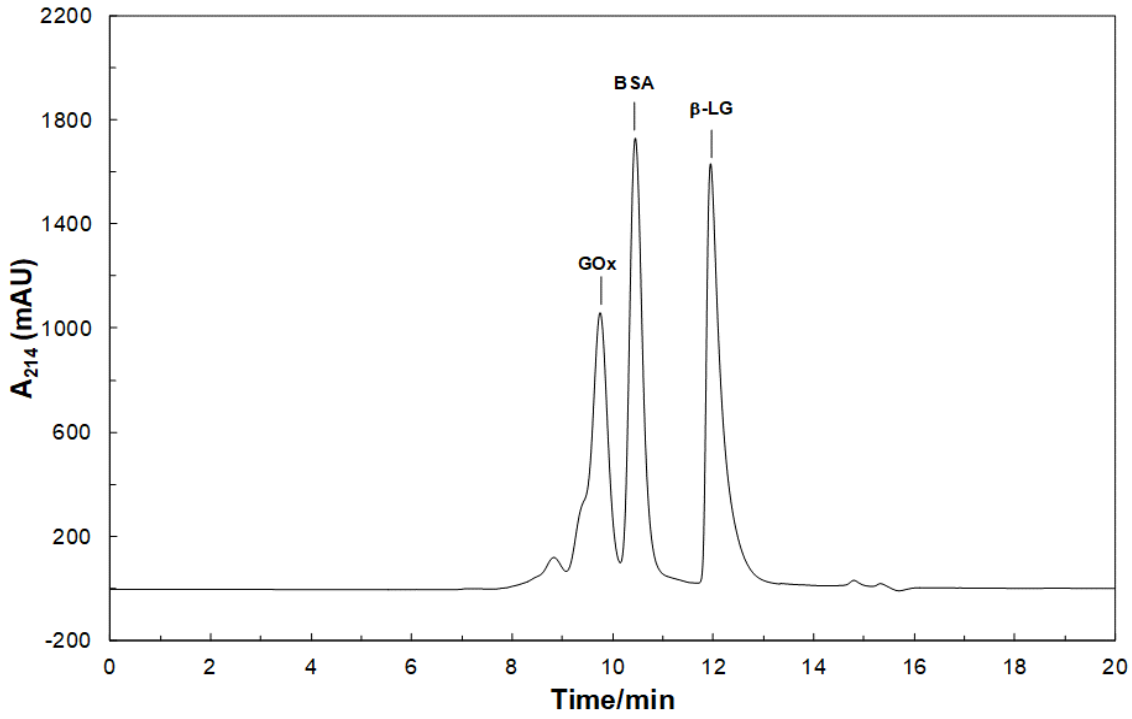


Figure 3. One-dimensional SEC separation (1D-SEC) of a mixed standard solution containing BSA, β-LG and GOx at a concentration of 500 mg L^{-1} each. Column: Yarra™ SEC-2000 300×7.8 mm + Security Guard Cartridge GFC-2000 4 x 3.0mm ID. Eluent: A) 0.1M phosphate buffer + 0.1 M NaCl at pH 6.8, isocratic mode. Flow rate: 0.7 mL min^{-1}. Injection volume: 100 µL.

The optimized 1D-SEC method was also used for performing quantitative analyses in protein extracts. The analytical performances of linearity and the chromatographic parameters are shown in Table 2.

Table 2. Performance and chromatographic parameters of proteins by 1D-SEC.

$y = a + bx$[a]					
Protein	**Retention time ± SD (min)**[b]	**a ± SD**	**b ± SD**	**R**[c]	**Linear Range (mgL^{-1})**
GOx	9.7±0.1	-5.6±2.8	1.01±0.01	0.9998	10-500
BSA	10.4±0.1	1.0±5.5	1.12±0.02	0.9994	10-500
β-LG	12.3±0.3	14.0±5.5	2.21±0.03	0.9997	10-300

a. y is the signal in mAU unit obtained from 1D-SEC separation and x is the value of concentration in mg L^{-1}
b. Mean value ± Standard Deviation for the first dimension separation evaluated by inter-day injections (n=10)
c. Correlation coefficient

A good linearity was found with a correlation coefficient higher than 0.9994. The goodness-of-fit of the data to the calibration curve was obtained in terms of response factor distribution (signal-to-concentration ratio, yi/xi) whose reference range is $(y/x)_{mean}$ ± 10%. Furthermore, the confidence interval for intercept, including the zero value (α = 95%, ν = 4), indicates the absence of systematic instrumental bias. Detection (LOD) and quantification (LOQ) limits were calculated according to the following equations: LOD= 3.3s/b and LOQ= 10s/b, where s is the standard deviation of noise levels and b is the slope of the regression line obtained from the calibration curve. From the chromatograms registered at the lowest calibration level (10 mg L^{-1}), the noise level was evaluated as Pk-to-Pk value, i.e. the difference between the maximum positive and the maximum negative amplitudes of baseline in the time window around the protein retention time. LOD and LOQ for BSA, chosen as reference protein, were 0.50 mg L^{-1} and 1.7 mg L^{-1}, respectively.

Second dimension separation based on reversed phase chromatography. For the second dimension separation, several C18 columns were used comparing the effect of different particle size, column dimension and technology. The best results were obtained by using a "*core-shell*" reversed phase column, designed for the analysis of intact proteins and polypeptides, that provided improved peak capacity, resolution, and a greater method flexibility than other fully porous columns [142].

During the optimization process of the mobile phase composition, the use of formic acid was evaluated, but better resolution and peak shape was obtained by applying water and acetonitrile binary gradients, acidified by trifluoroacetic acid. The effect of mobile phase flow-rate, column temperature, percentage of acetonitrile content and gradient ramp and shape (linear, concave and convex) were carefully considered. An increase of temperature from 30 °C to 70 °C by 10 °C step decreases the retention times for all three proteins (~ 1.5

min), BSA exhibiting a higher temperature dependence on retention. One-way ANOVA tests were performed on data (resolution, number of theoretical plates, peak area and base width) to verify the homogeneity of the mean values calculated on three replicates for each temperature value. No statistical difference was observed at 95% confidence level. Indeed, the separation between the protein couples BSA/β-LG and β-LG/GOx showed a flat trend from 30 °C to 50 °C and resolution does not significantly improve, even at 70 °C (Rs = 3.6-3.7 for β-LG/GOx and 2.4-2.6 for BSA/β-LG). Protein recovery from the stationary phase, expressed in terms of peak area, and efficiency, expressed as number of theoretical plates, displayed a similar behavior, characterized by constant values under the explored temperature range. Therefore, in order to preserve the column and extend its lifetime, a column temperature of 30 °C was selected for all the subsequent analyses. Finally, elution at 0.7 mL min^{-1}, under a composed multi-linear gradient, including a convex profile, offered distinct advantages in terms of separation time and peak resolution, as reported in Figure 4.

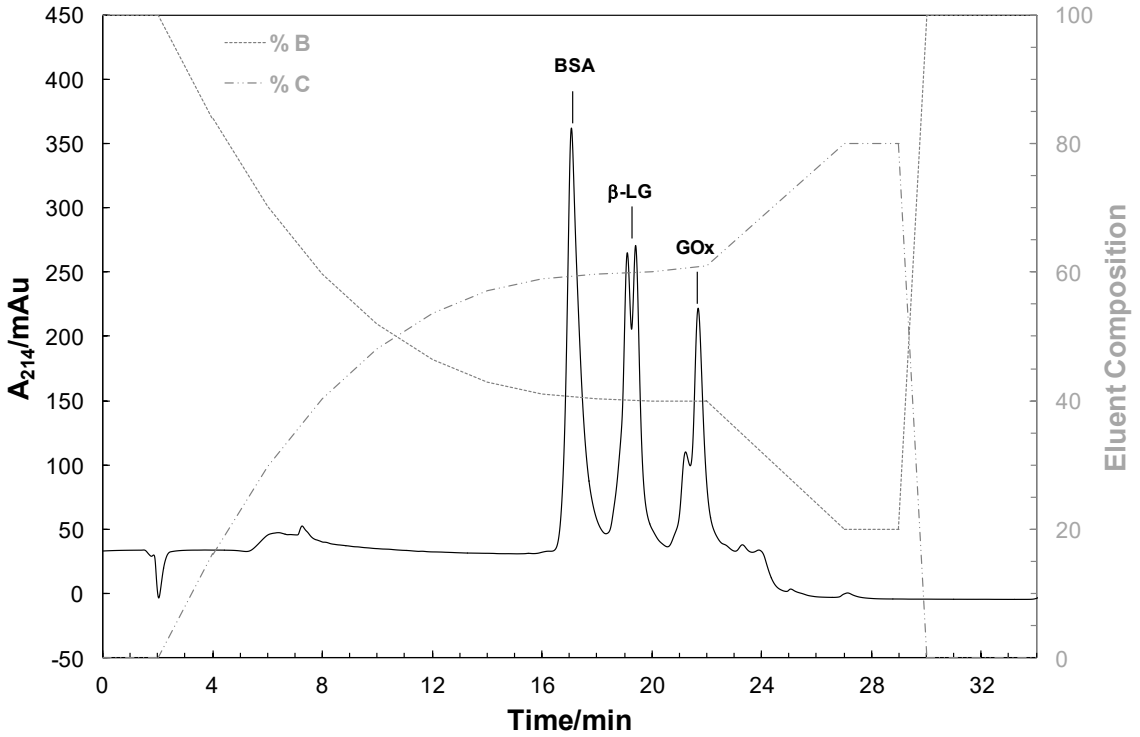


Figure 4. One-dimensional RP separation (1D-RP) of a mixed standard solution containing BSA, β-LG and GOx at a concentration of 500 mg L^{-1} each. Column: Aeris™ widepore XB-C18 150×4.6 mm, 3.6 µm + guard column ULTRA Cartridges UHPLC WIDEPORE C18 for 4.6mm. Eluent: B) 0.1% TFA in water and C) 0.1% TFA in acetonitrile, gradient mode. Flow rate: 0.7 mL min^{-1}. Injection volume: 100 µL.

A different elution order of proteins was observed, compared with SEC separation, due to the specific retention mechanisms of each column. GOx, with a hydrophobic grade of 52%, is the most retained protein and requires a higher percentage of acetonitrile for its

elution. The peak of β-LG is clearly split in two parts, corresponding to the two protein forms (the bovine genetic variants A and B), present in the commercial standard protein (coming from bovine milk), confirming the high resolving capacity and selectivity of the selected *coreshell* column for protein separation.

3.1.2. Development and optimization of a two-dimensional liquid chromatography (2D-LC) method (SEC-RP)

The development and optimization of the 2D-LC strategy was performed by testing different trapping devices placed between the two columns, since the interface between the two separation dimensions in an 2D-LC configuration is the heart of the system.

The specific valve design, the connection tubing and the mode of operation were investigated to improve the overall performance of the system in focusing the sample coming from the first dimension column onto the top of the second dimension column. The coupling of the two separation methods was carefully evaluated by protein standard analysis of BSA, β-LG and GOx, starting from the optimized conditions reached for the two 1D separation modes. A single 10-port switching valve was used as an instrumental device for coupling the two columns, alternatively connected to a single UV detection cell.

In preliminary experiments, the use of a fraction collection peek loop at different volumes, from 100 to 400 μL, was used to combine on-line the first SEC to the second RP dimension. After the appearance of the protein peak from the first column connected to the UV detector, and considering the time necessary to totally fill the collection loop (depending on the eluent flow rate and tubing size), the valve was switched and the left pump flow-rate was stopped after sample trapping, thus applying a stop-and-go approach in combination to the heart-cutting 2D-LC strategy. Then, from the collection loop, whose volume corresponded to the injection volume of the second dimension column, the protein arrived to the RP column head and gradient elution started. Unfortunately, signal spikes and baseline drifts were observed associated to re-mixing effects due to the sampling process into the second column, the large injection volume for the second dimension separation and the incompatibility of eluent composition and pH. In addition, such a stop-and-go approach caused a poor peak efficiency, as a consequence of protein diffusion processes through the stationary phase. Therefore, an innovative interfacing technology was designed consisting of a reversed phase guard column as a collection system to pack

the 1D aliquots coming from the first column, before the injection in the 2D column. In this way, the second dimension separation was decoupled from the first separation time-scale, enabling the operation of both dimensions under optimal conditions. Then, independently from column size, the trapping interface allows to limit the effects of dead volumes, peak deformation and band broadening. A second dimension column with smaller internal diameter than the first dimension column can be used and, moreover, large volumes of sample can be injected in the first column, then focused at the top of the second column.

In Figure 5, a schematic representation of instrumental set-up optimized to perform the 2D-LC separations is reported. Two different ternary pumps were connected through the 2D injector and the switching valve. Aliquots of the eluent from the first column were packed in the interface, and then injected onto the second column for further separation. A single UV cell was used in line with the first or the second column, based on the switching valve position. Finally, a collection fraction system was set at the end of the second dimension column to gather the isolated protein peaks, for their further off-line molecular characterization. When the switching valve is in position A, the sample is injected onto the SEC column; then, the eluate passes through the detector and is trapped in the collection trap until the SEC eluent flow is turned off. In valve position B, the content of the trap is injected onto the C18 column connected by the switching valve to the detector, to monitor the reversed-phase chromatographic separation. A complete experiment for the separation of three proteins in the standard mix is based on the repetition of this step sequence (switching valve position from A to B) three times, one for each protein eluted from SEC column. This experimental arrangement is an optimal combination between the two orthogonal 1D-methods and represents a good compromise between the two main 2D-LC approaches (comprehensive or heart-cutting).

The resulting SEC-RP stop-and-go method allowed to overcome the limits related to each 2D-LC mode, ensuring a high automation grade, resolving power and uncorrelated selectivity. Indeed, as a comprehensive 2D-LC, each peak is transferred to the second column and, as in multiple heart-cutting chromatography, the ideal conditions set for each 1D method can be applied. Moreover, the stop-and-go mode can be adopted, since the problems due to protein diffusion processes in the first column along its axis, when the flow rate from the left pump is stopped and late-eluting peaks are still in the SEC column, are minimized. Indeed, at a later time (when the flow-rate is again turned on) the proteins are focused and re-compacted in the trap column, before reaching the second dimension.

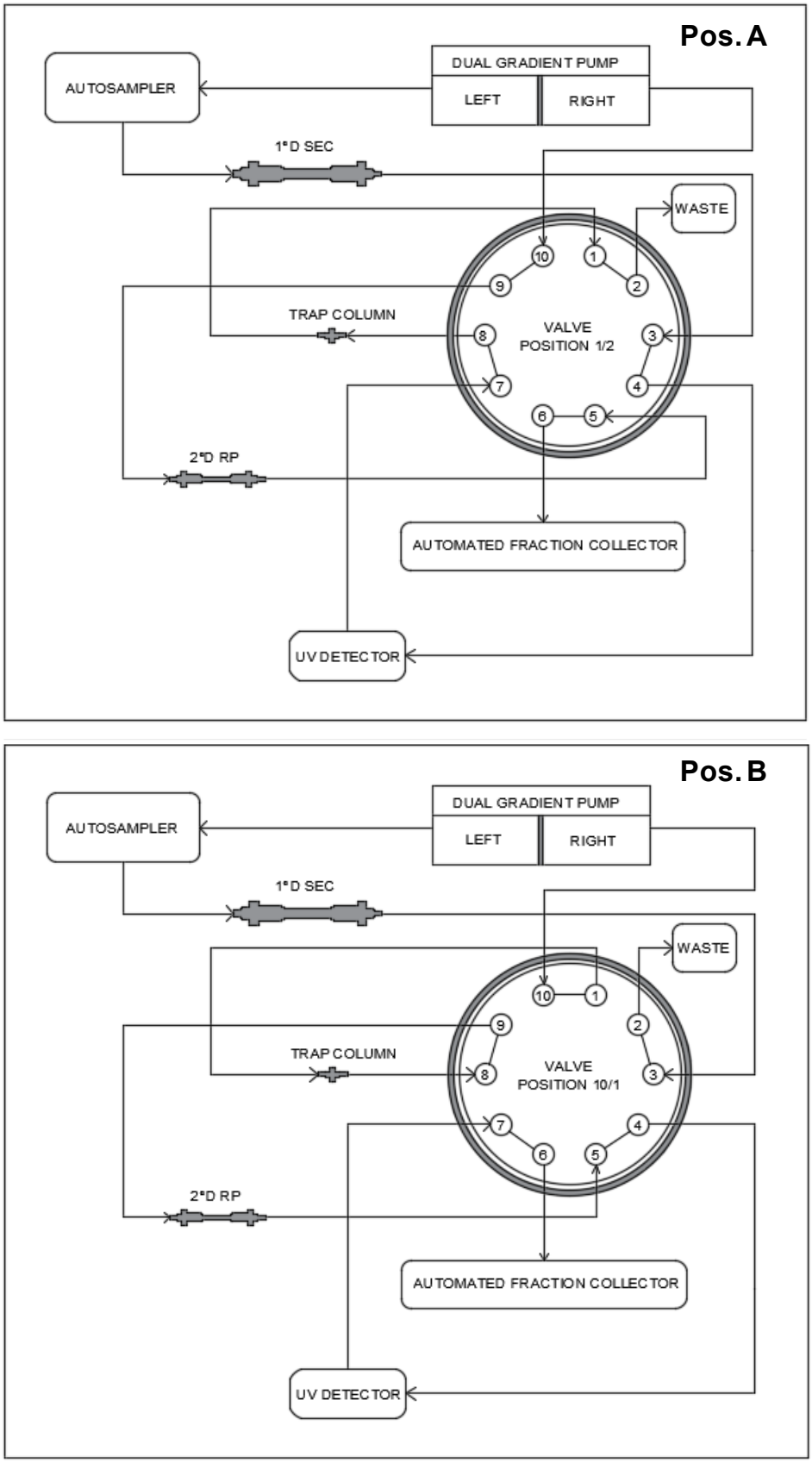


Figure 5. Schematic representation of 2D-LC experimental set-up. Left pump connected with column Yarra™ SEC-2000 300×7.8 mm + Security Guard Cartridge GFC-2000 4 x 3.0mm ID. Right pump connected with Aeris™ widepore XB-C18 150×4.6 mm, 3.6 µm. Guard column ULTRA Cartridges UHPLC WIDEPORE C18 for 4.6mm placed on the 10-port switching valve**.**

Figure 6 shows the 2D-LC separation of the mix standard of three proteins (BSA, β-LG and GOx) at a concentration of 500 mg L^{-1}.

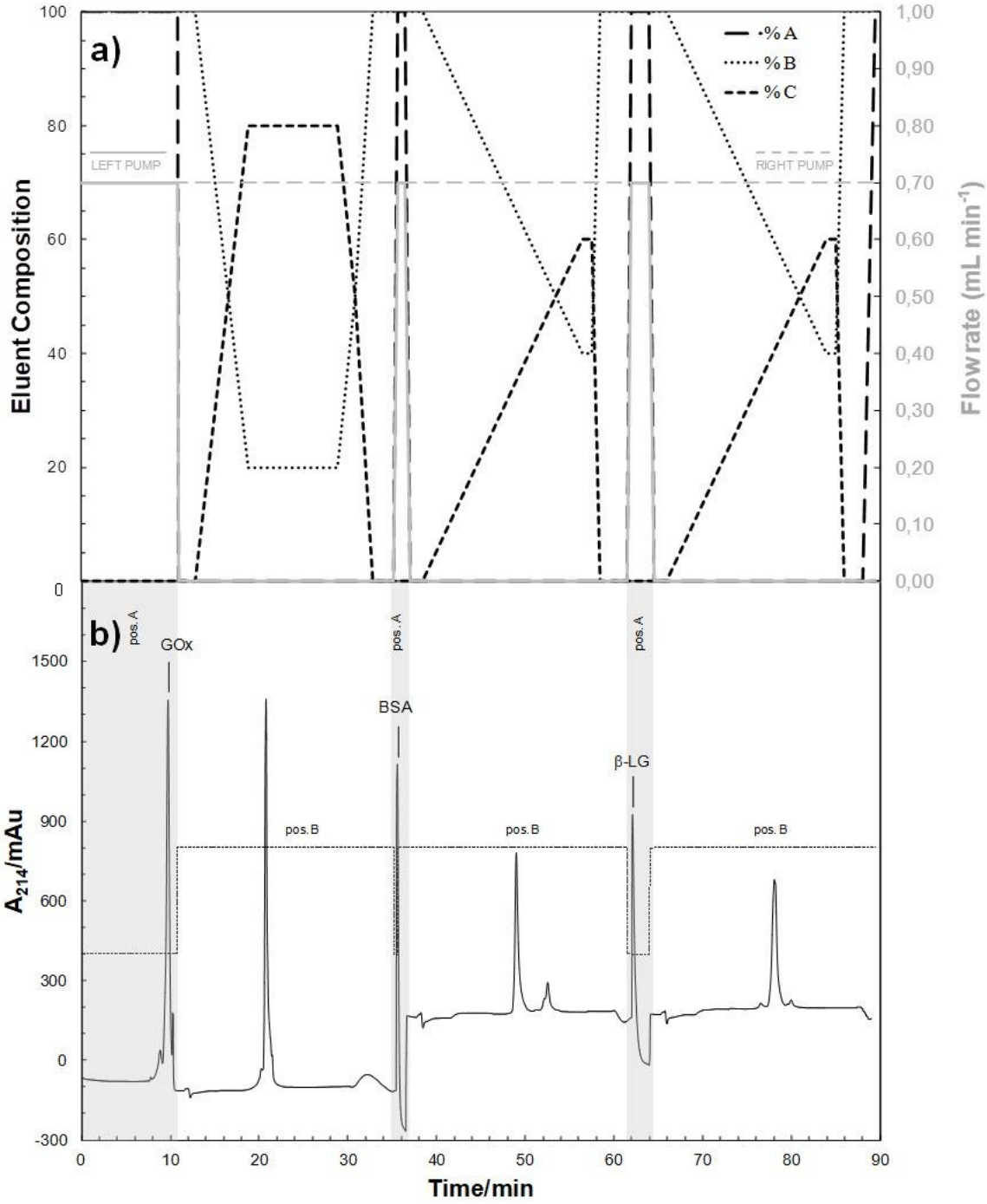


Figure 6. (a) Optimized ternary gradient elution based on: A) 0.1M phosphate buffer + 0.1 M NaCl at pH 6.8, B) 0.1% TFA in water and C) 0.1% TFA in acetonitrile for the 2D-LC SEC-RP chromatographic separation (b) of a mixed standard solution containing BSA, β-lactoglobulin β-LG and GOx at a concentration of 500 mg L^{-1} each. Injection volume: 100 µL. Left pump connected with column Yarra™ SEC-2000 300×7.8 mm + Security Guard Cartridge GFC-2000 4 x 3.0mm ID (eluent A). Right pump connected with Aeris™ widepore XB-C18 150×4.6 mm, 3.6 µm (eluent B and C). Guard column ULTRA Cartridges UHPLC WIDEPORE C18 for 4.6mm placed on the 10-port switching valve at position 1 and 8**.**

The proteins were isolated separately according to the elution order through the first column, flushed with phosphate buffer and sodium chloride; then, after valve switching, each protein was eluted in the second column by using a binary gradient based on water and acetonitrile acidified with TFA. Finally, each protein peak coming from the second dimension column (GOx, BSA and β-LG at 21.5 min, 49.2 min and 78.2 min, respectively) was collected by the fraction collector device. After solvent evaporation, the residues obtained from the three proteins were gathered, dissolved in mobile phase and re-injected into the column; the chromatographic profile, in line with the retention behavior previously observed for the protein mix, confirmed that the stop-and-go/active modulation approach can be used to on-line isolate intact proteins for their following characterization or large-scale studies and uses in biological systems.

A protein recovery percentage of 99±6% (n=6) from the guard column to the second dimension column was determined by comparing the protein amount calculated for the second dimension separation to the protein concentration obtained for the first dimension separation. The protein amount was estimated by interpolation of the BSA peak area on the calibration lines obtained by 1D-SEC and 1D-RP, as reported in Figures 7 and 8.

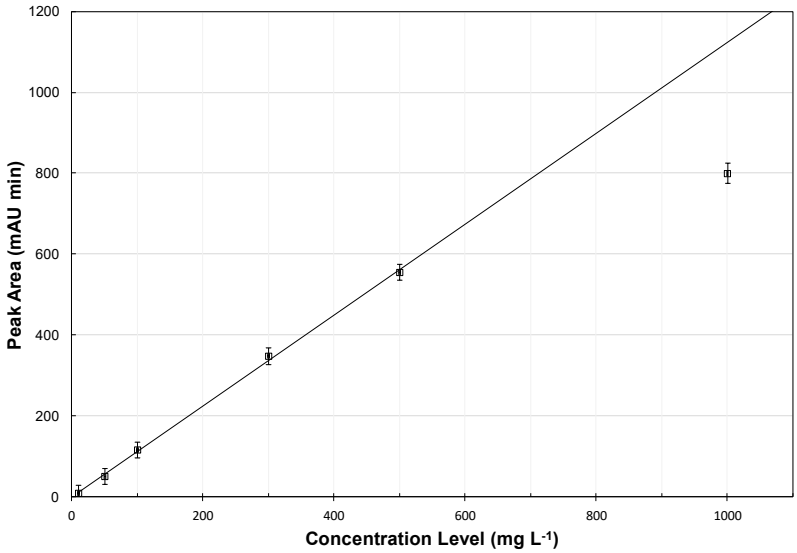


Figure 7. Calibration lines obtained for BSA by 1D-SEC. Column: Yarra™ SEC-2000, Eluent: 0.1M phosphate buffer + 0.1 M NaCl at pH 6.8. Flow rate: 0.7 mL min^{-1}. Injection volume: 100 µL.

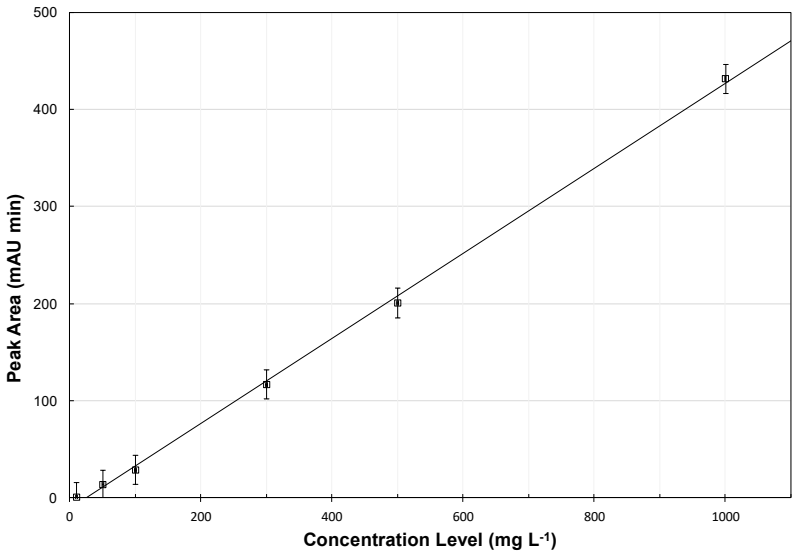


Figure 8. Calibration lines obtained for BSA by 1D-RP. Column: Aeris™ widepore XB-C18, Eluent: 0.1% TFA in water and 0.1% TFA in acetonitrile, gradient mode. Flow rate: 0.7 mL min^{-1}. Injection volume: 35 µL.

Although the fully automated workflow for protein fractionation has been described for a basis system of only three proteins, the number of cycles can be repeated several times to recovery all the protein peaks of each chromatographic run. Therefore, the proposed 2D-LC method represents a valid alternative to two-dimensional electrophoresis that is usually performed for protein isolation, allowing to overcome its intrinsic drawbacks of gel-to-gel variability, limitations for hydrophobic and alkaline proteins and problems in recovering intact proteins from the polyacrylamide gel [143].

An estimation of the protein amount that can be isolated for each peak fraction in a single chromatographic run was performed on the basis of the linearity range upper limit. For BSA and GOx an amount of 50 µg was estimated corresponding to 0.8 and 0.4 nmol, respectively, whereas 30 µg of β-LG (1.5 nmol) can be obtained from each analysis cycle, representing a sufficient amount for further MS characterization. Analogues considerations are also valid for the evaluation of the protein recovery from food samples. Although the protein identity is unknown, the amount can be approximately determined by the calibration line of BSA as a reference protein, in combination with the Log(MW) plot against the retention time (reported in Figure 2) for the estimation of the molecular weight.

3.1.3. Characterization of egg proteins

The optimized method has been used for the protein analysis in eggs, recognized as a potential allergen source. An essential requirement for achieving good performances in the analysis of real samples is the development and optimization of suitable methods for the extraction of the analytes from the matrix. For their analysis, proteins have to be disaggregated from cell or tissue extracts, denatured, reduced and properly solubilized. Generally, sample solubilization is performed in a buffer containing chaotropes (urea or thiourea), nonionic or zwitterionic detergents (Triton X-100 or CHAPS), reducing agents and, if necessary, protease inhibitors. Unfortunately, there is no universal method for sample pretreatment, and even published and standard protocols have to be adapted and further optimized for the type of sample to be analyzed.

A wide variety of extraction and fractionation tools for proteins and peptides are available based on their physicochemical and structural characteristics such as solubility, hydrophobicity, molecular weight, isoelectric point (pI). Generally, different technologies focused on cell disruption, solubilization/precipitation, and enrichment systems are needed to obtain the protein fraction of interest. Removal of interfering compounds (mainly lipids, nucleic acids, phenolic compounds, carbohydrates, proteolytic and oxidative enzymes, and pigments) is crucial. These procedures need to be optimized to minimize protein modifications and proteolysis, as well as to be compatible with subsequent analyses. Then, the efficiency of different extraction solutions and procedures based on solid-liquid and liquid-liquid extractions in combination with ultrasonication, centrifugation and filtration processes were tested in order to enhance the total protein recovery. Sample pre-treatment was optimized for the specific food sample under investigation to maximize protein extraction yield and reproducibility. An improved protein extraction method was developed based on SLE by a succession of four steps: (i) defatting by n-hexane and drying; (ii) water or saline extraction; (iii) sequential extraction using acid and alkali; (iv) isoelectric precipitation. Then, the protein extracts coming from each clean-up step were analyzed by the optimized 1D-SEC method to get an estimation of the protein content. Therefore, the sum of peak areas associated to each peak in the SEC profile was interpolated on the calibration line of BSA protein standard. A high protein amount was observed in the extract obtained after isoelectric precipitation, with a protein content mean value of 5.3 g/100 g of lyophilized egg (*intra-day* RSD = 6.5%, n = 4).

Before the 2D-LC separation, a shot-gun analysis by LC-MS/MS was performed for the identification of proteins present in the sample. Good results were obtained in terms of protein characterization, with a Mascot score ranging from 120 to 860 (well above the identity threshold) and a percentage coverage of 15-45% (Table 3), evaluated after the post-processing validation of the Peptide Spectrum Matches [135].

Five egg proteins were identified: ovalbumin, vitellogenin, ovotransferrin, apovitellenin and ovomucoid; an estimation of their molecular weights was performed on the basis of the total amino acid sequence range associated to the identified peptides by using the pI/Mw computation tool by ExPASy (Bioinformatics Resource Portal; https://web.expasy.org/compute_pi/, accessed on 31/05/2018). Reasonably, the effective analyzed proteins are larger than the estimated amino acid ranges, since missed cleavages could be observed during the enzymatic digestion in proximity of both N- and C-terminal regions. Obviously, such a situation is valid for all the egg proteins in the sample that were simultaneously submitted to the in-solution enzymatic digestion. Hence, the computed MWs (ranging from 7000 Da for apovitellenin to 200000 Da for vitellogenin) are a good approximation of the effective proteins and were used to assign the peak identity in the 1D-SEC chromatogram, by interpolation in the log plot of the molecular weights of standard proteins. For each identified peak, from the 1D-SEC profile, an estimation of the protein amount was performed by the calibration line of BSA as a reference protein. For ovalbumin, a protein amount of 93(± 8) µg was obtained from a single chromatographic run, corresponding to 2 nmol.

Table 3. List of the identified proteins in egg powder by LC-MS/MS (MASCOT search results against Swiss Prot database).

Protein	Entry name	Mascot score	No. validated Peptides	Sequence Coverage (%)	Amino Acidic range	Computed MW (kDa)
Ovalbumin	OVAL_CHICK	861	18	45	2-359	39
Vitellogenin	VIT2_CHICK	700	28	15	58-1820	207
Ovotransferrin	TRFE_CHICK	248	9	15	67-619	60
Apovitellenin	APOV1_CHICK	237	8	41	33-94	7
Ovomucoid	IOVO_CHICK	119	5	29	49-202	16

In Figure 9 the chromatographic profiles of an egg powder extract obtained after isoelectric precipitation are shown.

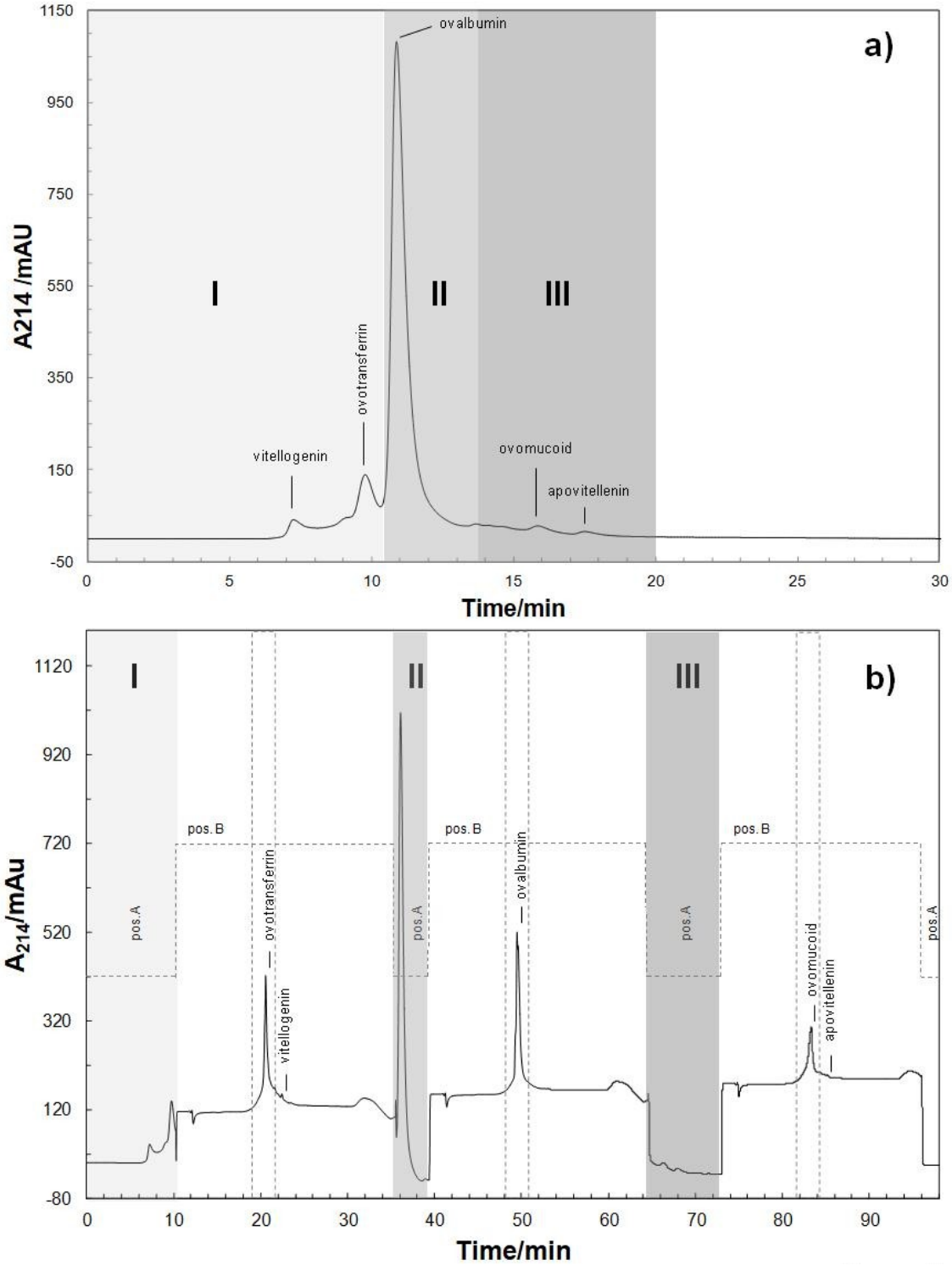


Figure 9. Chromatographic profiles of an egg powder extract obtained after isoelectric precipitation: (a) 1D-SEC separation. Column: Yarra™ SEC-2000 300×7.8 mm + Security Guard Cartridge GFC-2000 4 x 3.0 mm ID. Eluent: 0.1M phosphate buffer + 0.1 M NaCl at pH 6.8, isocratic mode. Flow rate: 0.7 mL min^{-1}. Injection volume: 100 µL; b) 2D-LC SEC-RP. Left Pump connected with column Yarra™ SEC-2000 300×7.8 mm + Security Guard Cartridge GFC-2000 4 x 3.0 mm ID. Right pump connected with Aeris™ widepore XB-C18 150×4.6 mm, 3.6 µm. Guard column ULTRA Cartridges UHPLC WIDEPORE C18 for 4.6 mm placed on the 10-port switching valve. Injection volume: 100 µL.

The experimental conditions established for the 2DLC separation of the standard protein mix were easily adapted and properly adjusted for the analysis of food extracts, based on the effective chemical characteristics of the type of sample to be analyzed. Therefore, before the 2D-LC separation, it is critical to perform in advance a 1D-SEC run to get the protein assay estimation and to note the retention times of the proteins. Then, for each food sample, the sampling time associated to the switching valve was determined on the basis of protein retention times and chromatographic behavior through the first column (Figure 9a). Compared to the optimal conditions established for the standard proteins (Figure 4), an additional method improvement was performed in terms of gradient elution through the RP column, as reported in Table 4.

Table 4. Experimental set-up for the 2D-LC separation of an egg powder extract. A) 0.1M phosphate buffer (pH 6.8) containing 0.1M sodium chloride, B) 0.1% TFA in water and C) 0.1% TFA in acetonitrile.

Time (min)	Switching valve position	LEFT PUMP		RIGHT PUMP		
		Flow-rate (mL/min)	%A	%B	%C	Flow-rate (mL/min)
0.0	A	0.7	100	100	0	0.7
10.3	B	0.7	100	100	0	0.7
10.4		0.0	100	100	0	0.7
12.3		0.0	100	100	0	0.7
18.3		0.0	100	20	80	0.7
28.3		0.0	100	20	80	0.7
30.3		0.0	100	50	50	0.7
32.3		0.0	100	100	0	0.7
35.3	A	0.0	100	100	0	0.7
35.8		0.7	100	100	0	0.7
39.4	B	0.7	100	100	0	0.7
39.5		0.0	100	100	0	0.7
41.4		0.0	100	100	0	0.7
47.4		0.0	100	20	80	0.7
57.4		0.0	100	20	80	0.7
59.4		0.0	100	50	50	0.7
61.4		0.0	100	100	0	0.7
64.4	A	0.0	100	100	0	0.7
64.9		0.7	100	100	0	0.7
73.0	B	0.7	100	100	0	0.7
75.0		0.7	100	100	0	0.7
81.0		0.7	100	20	80	0.7
91.0		0.7	100	20	80	0.7
93.0		0.7	100	50	50	0.7
95.0		0.7	100	100	0	0.7
96.0	A	0.7	100	100	0	0.7
98.0		0.7	100	100	0	0.7

Better results in terms of efficiency and protein recovery (evaluated by peak areas) were observed by increasing the elution strength with an organic percentage up to 80% for the second dimension separation. Indeed, the proposed 2D-LC approach is a prototype system to be used as a basis model for the online protein fractionation in food samples. Therefore, if necessary, the chromatographic conditions for the reversed phase second dimension separation are further optimized (e.g. by varying the organic percentage in the mobile phase) for an improved separation and recovery of the protein fractions. Nevertheless, such a method limitation also represents a key factor, confirming the method potential and its versatility. Indeed, the adjustment of the trapping time interval based on the elution times in the first dimension and the optimization of the gradient separation through the second dimension column can be quickly performed by setting the experimental conditions in a programmed spreadsheet for a general 2D-LC algorithm, as an example see Figure 10.

			LEFT PUMP		RIGHT PUMP				
Switching Valve	Time/min	%A	Flow-rate	%B	%C	Flow-rate			
pos. A	0,0	100	0,7	100	0	0,7	^{1}D+trapping	1st cycle	
pos. B	P3	100	0,7	100	0	0,7			
	C5+0,1	0	0	100	0	0,7	^{2}D		
	C5+2,0	0	0	100	0	0,7			
	C5+8,0	0	0	20	80	0,7			
	C5+18,0	0	0	20	80	0,7			
	C5+20,0	0	0	50	50	0,7			
	C5+22,0	0	0	100	0	0,7			
pos. A	C5+24,4	0	0	100	0	0,7	^{1}D+trapping	2nd Cycle	
pos. B	C5+24,8	100	0,7	100	0	0,7			
	C13+P4	100	0,7	100	0	0,7	^{2}D		
	C14+0,5	0	0	100	0	0,7			
	C14+2,0	0	0	100	0	0,7			
	C14+20,0	0	0	40	60	0,7			
	C14+21,0	0	0	40	60	0,7			
	C14+22,0	0	0	100	0	0,7			
	C14+25,0	0	0	100	0	0,7			
pos. A	C14+25,5	100	0,7	100	0	0,7	^{1}D+trapping	3rd cycle	
pos. B	C21+P5	100	0,7	100	0	0,7			
	C22+0,5	0	0	100	0	0,7	^{2}D		
	C22+2,0	0	0	100	0	0,7			
	C22+20,0	0	0	40	60	0,7			
	C22+21,0	0	0	40	60	0,7			
	C22+22,0	0	0	100	0	0,7			
	C22+24,0	0	0	100	0	0,7			
pos. A	C22+25,5	100	0,7	100	0	0,7			

INSERT peak isolation range from 1D-SEC profile (min)		Time (min)	$V_{flushed}$ (mL)	$V_{required}$ (mL)
1st retention window	xx	(R3-(0,7*0,1/2))/0,7	R3	O3*0,7
2nd retention window	xx	(R4-Q4)/0,7	0,4*0,7/2+0,5*0,7/2	O4*0,7
3rd retention window	xx	(R5-Q5)/0,7	0,7*0,5/2+0,7*0,5/2	O5*0,7

Figure 10. Instructions for the generation of a 2DLC time program at three trapping cycles.

The 2DLC cycles can be repeated several times in order to recovery all the protein peaks from each ^{1}D chromatogram. Then, by the ^{2}D separation, a sufficient amount of material associated to each protein (not only the most abundant, but also the low-intensity peaks) can be collected in few mL of organic solvent that is easily removed by evaporation. Finally each isolated protein can be reconstituted in a proper solvent or aqueous buffer for further molecular characterization or bio-analytical studies.

From a food sample at high complexity and protein content such as egg (12-17 g of protein/100 g of fresh sample), an automated recovery of intact proteins was obtained in less than 100 minutes. The peaks coming from the second dimension separation of ovotransferrin, ovalbumin and ovomucoid are clearly observed at 20.8 min, 50.0 min and 81.8 min, respectively. Small peaks for vitellogenin and apovitellenin are baseline detectable in the first and third chromatographic cycle, at retention times slightly higher than ovotranferrin and ovomucoid, respectively, due to their greater hydrophobicity grade (56% vs. 48% for vitellogenin vs. ovotranferrin and 56% vs. 46% for apovitellenin vs. ovomucoid). Finally, individual protein fractions were collected in-line, in correspondence of their retention time-window. After solvent evaporation, each protein was stored for subsequent proteome-wide analysis at the intact protein level. Hence, such an automated 2D-LC fractionation of intact proteins can be easily combined to a proteomic MS characterization or allergy tests, overcoming the problems (labor-intense steps and poor recoveries) generally observed in gel-based protein separation methods.

3.2. Development and optimization of analytical methods for protein analysis in soy flour samples

In the first part of the Ph.D. research activity, an actively modulated two-dimensional liquid chromatography method has been developed for the online protein isolation from egg samples on a preparative scale. The *stop-and-go* 2D-LC approach has been optimized by using as an interface system a reversed phase guard column placed on a single switching valve. The proposed gel-free method represented a good compromise between the *comprehensive* [144–146] and *heart-cutting* [147,148] modes in 2D-LC, allowing the isolation of intact proteins from egg samples for their MS characterization or further bio-analytical studies.

In the following paragraphs, an enhanced approach for the protein isolation from soy flour samples by two-dimensional liquid chromatography coupled with an active modulation interface, based on the use of a double switching valve system combined to a short C4 analytical trapping column, will be described. Such an interfacing technology was designed with the aim of including a trap washing step between the first and the second separation, therefore buffer residues from the eluent of the first column can be easily removed, before the second dimension separation. Hence, the proposed instrumental interface device allows reducing the remixing effects and the sample dilution process, occurring during each of the two chromatographic steps. In this way, the chromatographic problems linked to the solvent incompatibility in terms of eluent composition and pH are overcome and the analytical performances during the separation in the second column improved. Indeed, a fundamental issue in the 2D-LC experimental set-up is the interface system between the two separation dimensions, which affect the way in which fractions from the first column are collected and transferred to the second column [149]. The design of instrumental devices used as an interface between the two columns influences the volume and the solvent composition of the fractions injected in the second column [150].

3.2.1. Enhancing protein isolation by active modulation interface with double switching valve system coupled with a C4 trapping column

The experimental conditions established for the first dimension separation by SEC generally require the use of chemicals, detergents or salts that could present problems for the reversed-phase second dimension separation. Several desalting systems have been proposed, especially when employing salt based mobile phases in multiple heart-cutting two-dimensional liquid chromatography coupled with MS [151]. Then, 2D-LC platforms, based on an extensive valve/loop setup with one or two multiple-port two-position switching valves equipped with a single or more loops, have been implemented for modulation, peak parking, and trapping [152,153]. In the previous paragraphs, for the egg protein analysis, an interface system consisting in a reversed-phase guard column placed on a single switching valve was reported. Nevertheless, in order to include a trap washing step between the first and the second separation with the aim of enhancing the chromatographic performances of the RP column, the use of an innovative interface

technology, based on the presence of two 10-port switching valves and a short trapping RP column, is here described and then applied to the analysis of soy flour samples. A focusing system placed in the interface between the two dimensions is essential to retain the analytes while the rest of the effluent flows through. The selection of a proper packing material and size of the trapping column allows to create the right conditions for separating the target compounds from the other components of the buffer solutions used as an SEC eluent. This is accomplished by choosing a packed column bed that has a strong affinity for the target compounds under the elution conditions applied for the first dimension separation. On the other side, when the mobile phase composition changes and the organic percentage increases, under the gradient elution conditions set for the reversed-phase separation, the analytes elute from the trapping system and enter the second dimension column. The use of a guard cartridge packed with a bonded phase similar to the second column, as previously reported for the egg analysis, allows to focus the proteins coming from the primary column, but it is not sufficiently effective in the removal of the salts of the SEC eluent. Indeed, the guard column, as a short, disposable pre-column, allowed to remove particulates and contaminants, but due to its reduced dimension, was not able to retain proteins for the time necessary to remove salts from the SEC eluent. Indeed, during a trap washing step of just 1-2 min the target analytes moved from the guard column to the waste, whereas the second dimension column is equilibrated with the initial mobile phase. Therefore, a 5-cm reversed-phase trapping column is here proposed. Due to its chemical characteristics in the retention mechanism, a less hydrophobic C4 stationary phase was chosen in replacement of a C18 trap column, generally showing a higher protein retention that could represent a problem in an interface system between the two dimensions. Then, the double valve interface system, allows to include an effective trap washing step of 7 min at 0.7 mL min^{-1} (corresponding to 6 column volumes) after parking ^{1}D aliquots and before running each ^{2}D cycle. In Figure 11, a schematic representation of instrumental set-up optimized to perform 2D-LC separations is reported.

Two different ternary pumps were connected through the 2D injector and two switching valves. The first dimension SEC column is connected to the right valve (RV), while the second RP and the trapping column are connected to the left valve (LV). Aliquots of the eluent from the first column were parked in the C4 trap column (until the SEC eluent flow is turned off) and then injected into the second column for further separation. A restrictor coil is placed on the right switching valve to get sufficient backpressure during the trap washing step, performed by-passing the SEC column. A single UV cell was used in line

with the first or the second column, based on the switching valve positions. Finally, a collection fraction system was set at the end of the second dimension column to gather the isolated protein peaks.

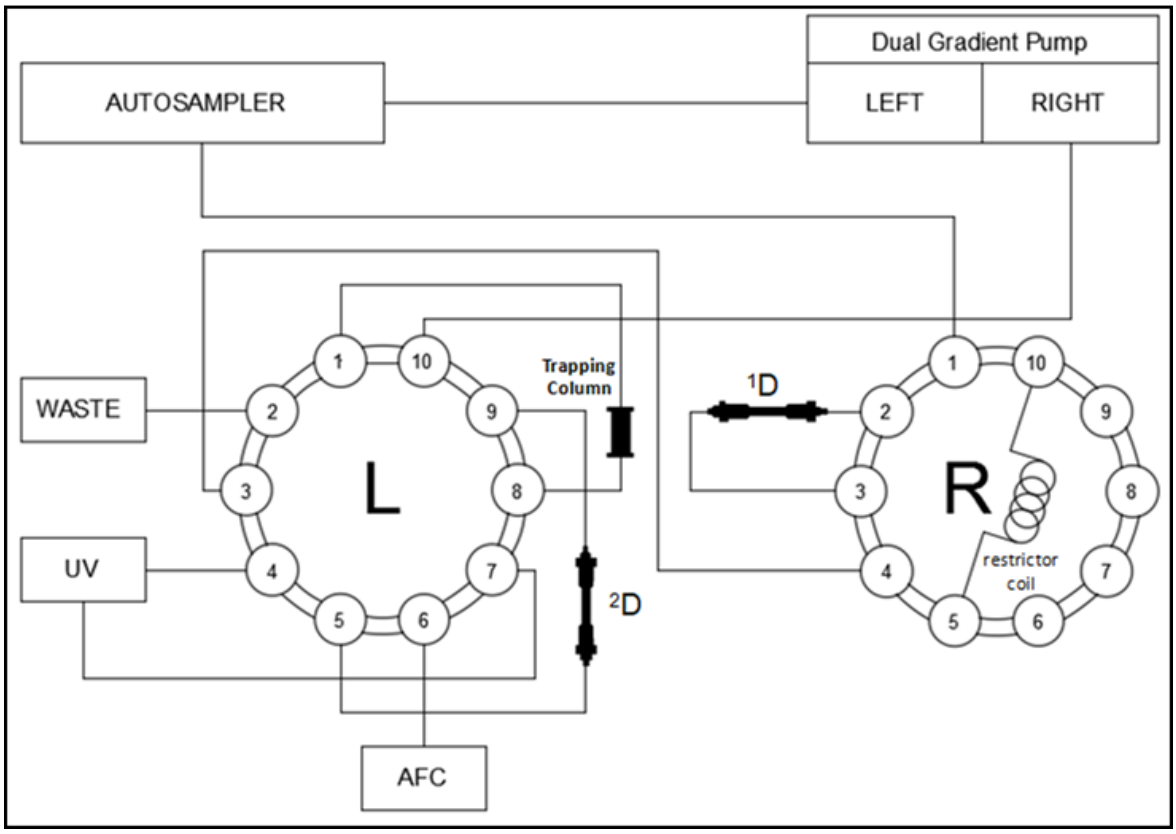


Figure 11. Schematic representation of 2D-LC experimental set-up. Left pump connected with column MabPac SEC-1 (150×4.6 mm) + Security Guard Cartridge (GFC-2000, 4 x 3.0 mm), placed on the right 10-port switching valve (RV). Right pump connected with Aeris widepore XB-C18 column (150×4.6 mm) + security guard column (ULTRA Cartridges UHPLC WIDEPORE C18), with a trapping column Eurosil C4 (50 x 4.6 mm) placed on the left 10-port switching valve (LV). An automated fraction collection device (AFC) is used to gather the protein peaks eluting from the second dimension column. First step (sample injection and 1stDimension separation): LV position 1/2; RV position 1/2. Second step (C4 trap washing): LV position 1/2; RV position 10/1. Third step (2ndDimension separation): LV position 10/1; RV position 10/1.

Each 2D-LC cycle consists of a repetition of three steps:

I. *1D separation and protein trapping*: when the left and right switching valves are both in position 1/2, the proteins are separated in the first column (eluent A), connected to the UV cell, and then focused into the C4 trap column, while the second dimension column is conditioned with the initial mobile phase (eluent B);
II. *trap washing*: after switching the right valve to position 10/1, the first dimension column is brought into the stop-flow mode, the C4 trap is washed with acidified water (eluent B) to remove salts and the second column is still flushed by acidified water;

III. *2nd D separation and gradient elution*: after switching the left valve to position 10/1, the proteins are back-flushed from the trap to the reversed phase C18 column and the second dimension separation starts in gradient mode (eluents B and C).

Obviously, the succession of these three steps has to be performed for each fraction collected from the first column. This approach minimizes band broadening, reduces the re-mixing effects, extends the mobile phase compatibility between the two columns (in terms of pH and presence of salts or additives) and improves sensitivity in the second dimension run.

3.2.2. Optimization of 2D-LC (SEC-RP) chromatographic separation

The optimization of the experimental conditions associated with the mono-dimensional SEC separation (1D-SEC) was performed by varying the mobile phase composition and flow-rates, under isocratic conditions. Although SEC-bonded stationary phases are designed to minimize interactions with biomolecules, the presence of ionic secondary interactions can dramatically impact chromatographic performances. Therefore, the use of NaCl as a buffer additive at a concentration of 0.1-0.3 M (in line with the recommended operating conditions for the MAbPAC column) was tested to evaluate the impact of secondary interactions on the retention behavior in size exclusion chromatography by gel filtration. The concentration of phosphate buffer in the mobile phase was explored in the range of 50-100 mM, at flow rates of 0.20-0.25 mL min^{-1}. The best results in terms of resolution and shape and analysis time were obtained by using phosphate buffer at a concentration of 50 mM, modified by adding 0.3 M NaCl. In Figure 12 the effect of flow rate at 0.20 and 0.25 mL min^{-1} is shown, confirming that a value of 0.25 mL min^{-1} is optimal as a balance between analysis time and peak resolution.

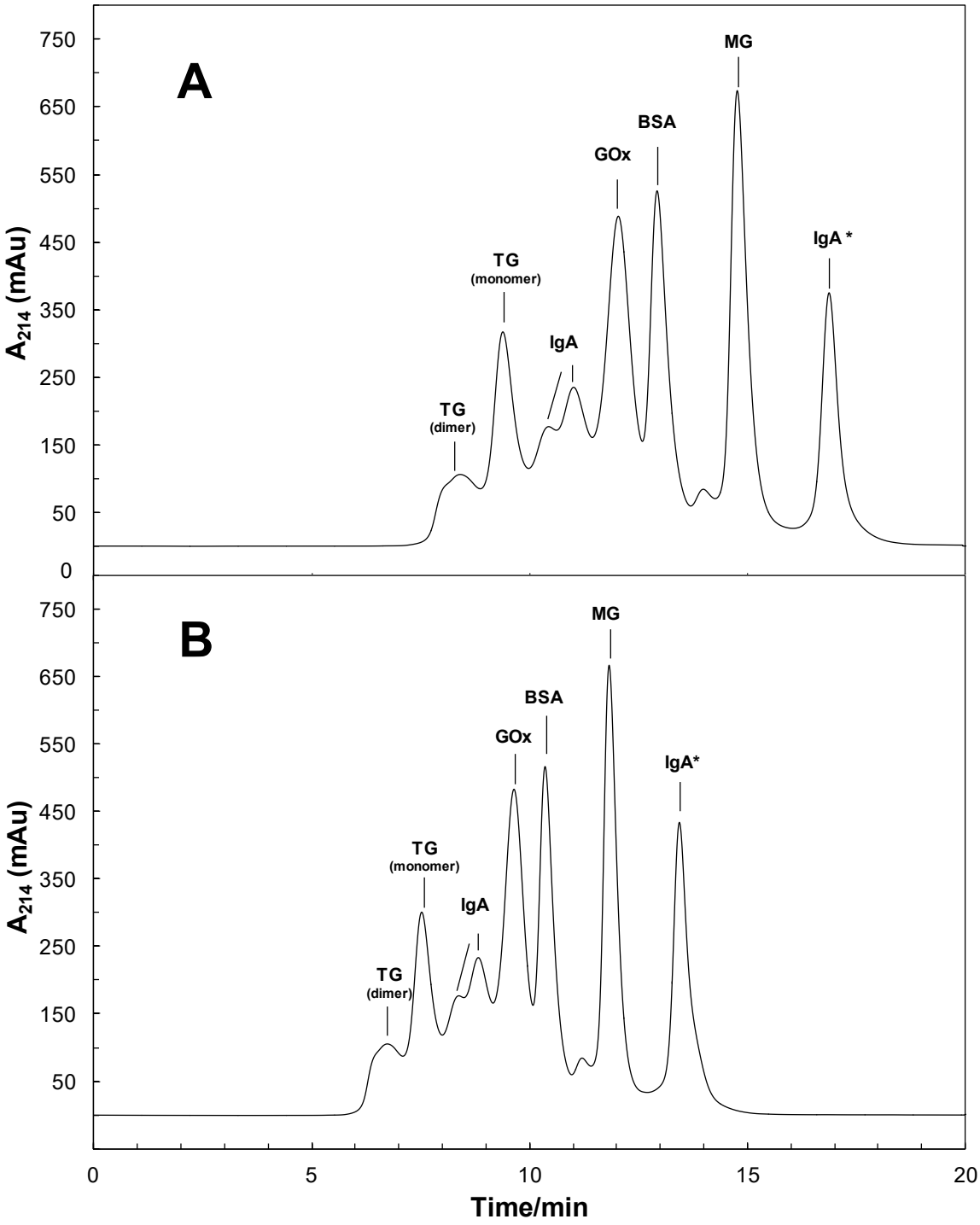


Figure 12. One-dimensional SEC separation (1D-SEC) of a mixed standard solution containing TG, GOx, BSA and MG at a concentration of 500 mg L^{-1} and IgA at a concentration of 200 mg L^{-1}. Column: MabPac SEC-1 (150×4.6 mm) + Security Guard Cartridge (GFC-2000 4 x 3.0 mm). Eluent: A) 50 mM phosphate buffer + 0.3 M NaCl at pH 6.8, isocratic mode. Injection volume: 15 µL. Flow rate: A) 0.20 mL min^{-1}. B) 0.25 mL min^{-1}.

Under the optimized conditions, the log chart of the molecular weights of standard proteins against the retention time was plotted (see Figure 13), as a useful tool for the estimation of the molecular weight of unknown proteins in food samples.

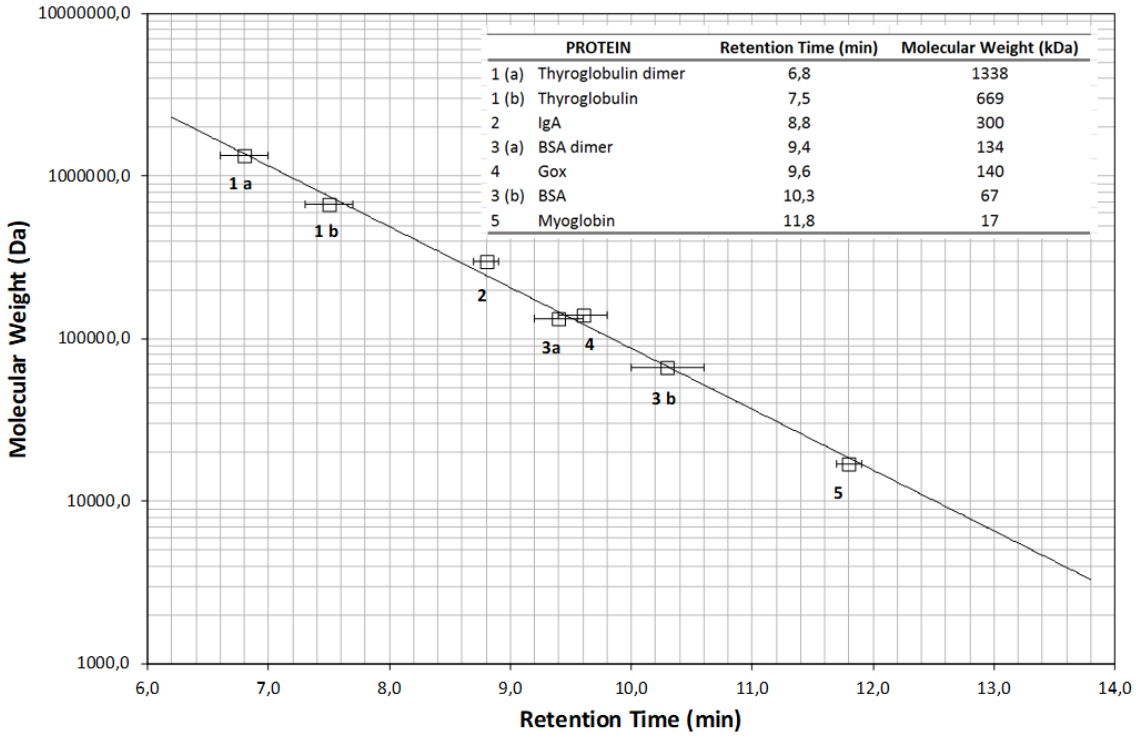


Figure 13. Protein Mixture Calibration by One-Dimensional SEC separation (1D-SEC). Column: MAbPac SEC-1™ (300×4 mm, Thermo Fisher Scientific) + Security Guard Cartridge GFC-2000 (4 x 3.0mm ID, Phenomenex). Eluent: 50 mM phosphate buffer + 0.3 M NaCl at pH 6.8, isocratic mode. Flow rate: 0.25 mL min^{-1}. Injection volume: 15 µL.

This graph was also used for the evaluation of the molecular weight of the star peak deriving from IgA standard (see Figure 12) corresponding to a protein fragment with a weight of 4300 ± 700 Da.

The optimized 1D-SEC method can be also used for performing quantitative analyses in food protein extracts. Therefore, calibration data were obtained by 1D-SEC analyses of standard solutions of BSA, chosen as a reference protein. The analytical performances of linearity are shown in Table 5.

Table 5. Calibration and chromatographic parameters by 1D-SEC.

y = a + bx[a]						
Protein	**a ± SD**	**b ± SD**	**R**[b]	**LOD**[c]	**LOQ**[c]	**Linear Range**
				(µg L^{-1})		(mg L^{-1})
BSA	-3.4±1.1	0.416±0.005	0.9997	6.2	21	0.021-500

a. y is the signal in mAU unit obtained from 1D-SEC separation and x is the value of concentration in mg L^{-1}
b. Correlation coefficient
c. Instrumental LOD and LOQ values evaluated at a signal-to-noise ratio of 3 and 10, respectively

The limit of detection (LOD) and quantification (LOQ) were determined by the chromatograms of the standard solutions of BSA at the lowest calibration level (5 mg L^{-1}). LOD and LOQ, estimated at a signal-to-noise ratio of 3 and 10, were 6.2 µg L^{-1} and 21 µg L^{-1}, respectively. Detection and quantification limits were calculated according to the following equations: LOD= 3.3s/b and LOQ= 10s/b, where s is the standard deviation of noise levels and b is the slope of the regression line obtained from the calibration curve. From the chromatograms registered at the lowest calibration level (5 mg L^{-1}), the noise level was evaluated as Pk-to-Pk value, i.e. the difference between the maximum positive and the maximum negative amplitudes of baseline in the time window around the protein retention time.

For the second dimension separations, as previously mentioned, a "*core-shell*" reversed phase column, designed for the analysis of intact proteins and polypeptides, was used by applying optimized water and acetonitrile binary gradients, acidified by trifluoroacetic acid. Increasing column temperature generally reduces protein retention and sharpens protein peak shapes, therefore the effect of the column temperature was studied in the range 30-70 °C by 10 °C steps, as shown in Figure 14. Although the Aeris C18 column is designed to tolerate even higher column temperatures (up to 90 °C), it is more advisable not operating close to the maximum temperature limit to preserve the column and extend its lifetime. An increase of temperature decreases the retention times for BSA, GOx, IgA and MG (~1.3-1.9 min from 30 ° to 70 °C), whereas no retention time dependence on temperature was observed for TG (Figure 14, panel A). In terms of peak area (Figure 14B), no statistical difference (ANOVA one-way test on three replicates for each temperature value) was observed at 95% confidence level for BSA, GOx, and MG; on the contrary, TG (in minor extent) and IgA (more evident) exhibit a significant temperature dependence. Therefore, a column temperature of 70 °C is required to facilitate the protein recovery from the stationary phase, allowing to minimize the memory effect among consecutive injections, as confirmed by the low value of residual peak area observed for TG at higher temperatures. Under the optimized 1D-RP experimental conditions, the calibration parameters obtained for the reference protein BSA are reported in Table 6.

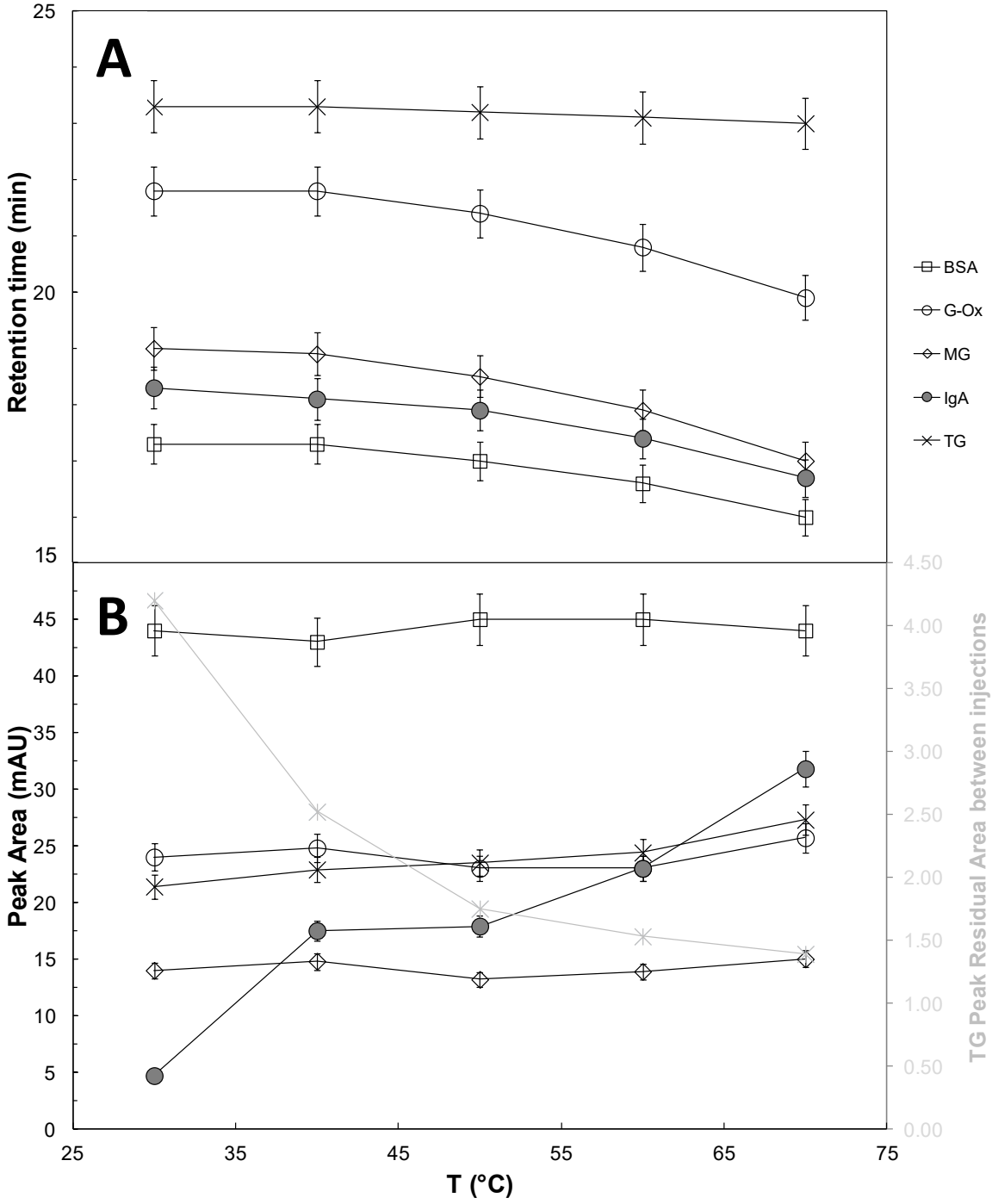


Figure 14. Effect of temperature on retention times (panel A) and peak area (panel B) of 100 mg L^{-1} mixed standard solutions. One-dimensional RP separation conditions: column Aeris™ widepore XB-C18 (150×4.6 mm) + security guard column (ULTRA Cartridges UHPLC WIDEPORE C18). Elution with 0.1% TFA in water and 0.1% TFA in acetonitrile, gradient mode. Flow rate: 0.7 mL min^{-1}. Injection volume: 100 µL.

Table 6. Calibration and chromatographic parameters by 1D-RP.

	$y = a + bx$[a]					
Protein	**a ± SD**	**b ± SD**	**R**[b]	**LOD**[c]	**LOQ**[c]	**Linear Range**
BSA	-10.7±3.6	$(437.8±7.7)10^{-3}$	0.9993	($\mu g\ L^{-1}$)		($mg\ L^{-1}$)
				7.5	23	0.023-1000

a. y is the signal in mAU unit obtained from 1D-SEC separation and x is the value of concentration in mg L^{-1}

b. Correlation coefficient

c. Instrumental LOD and LOQ values evaluated at a signal-to-noise ratio of 3 and 10, respectively

As a result of the optimization process of each one dimensional LC method, a good coverage in the two-dimensional separation space was obtained (Figure 15), therefore the 1D-SEC and 1D-RP methods were combined in a fully automated workflow for protein fractionation coupled with two switching valves.

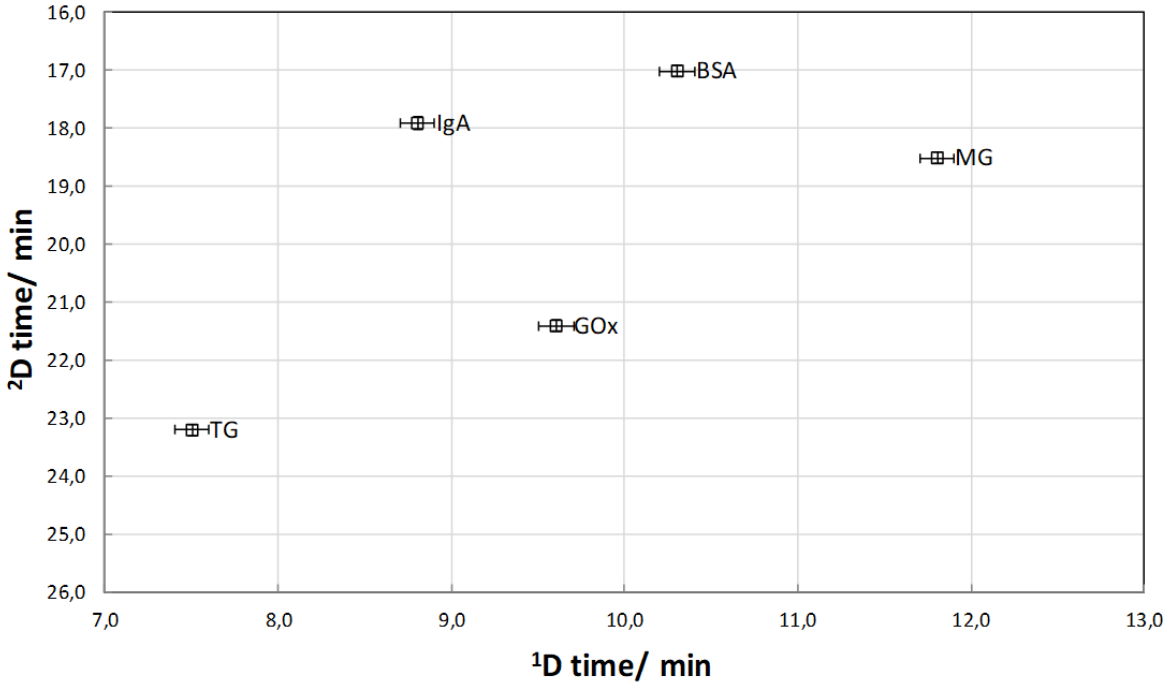


Figure 15. Retention time coordinates in a two-dimensional separation space obtained from a SEC-RP separation of a mixed protein standard solution. ^{1}D: SEC separation by column MabPac SEC-1 (150×4.6 mm) + security guard cartridge GFC-2000 (4 x 3.0 mm ID, Phenomenex), flushed with 50 mM phosphate buffer + 0.3 M NaCl at pH 6.8. Flow rate: 0.25 mL min^{-1}. ^{2}D: RP separation by column Aeris™ widepore XB-C18 (150×4.6 mm) + security guard column (ULTRA Cartridges UHPLC WIDEPORE C18). Elution with 0.1% TFA in water and 0.1% TFA in acetonitrile, gradient mode. Flow rate: 0.7 mL min^{-1}.

In Figure 16, the 2D-LC profiles obtained for the protein standard mix by applying three trapping cycles under the optimized conditions reached for the two 1D separation modes are reported. In each cycle, the proteins were isolated separately according to the elution order through the first column (^{1}D) as blocks of one or more proteins, then each protein

block was eluted in the second column (^{2}D) and collected by the fraction collector device for their following molecular characterization or large-scale biological investigations. For the protein couple BSA/MG (second cycle) it seems that a resolution worsening is observed in the second dimension. Anyway, this is the price to be paid for the improvement of the separation between GOx and BSA that in the first dimension show partially overlapping peaks. As it is evident from the profile of the third cycle, a poor recovery was observed for IgA* from the C18 second dimension column in the second dimension separation, that is a well-known issue for monoclonal antibodies in reversed-phase liquid chromatography [154], although the reasons behind this behavior are not yet fully understood.

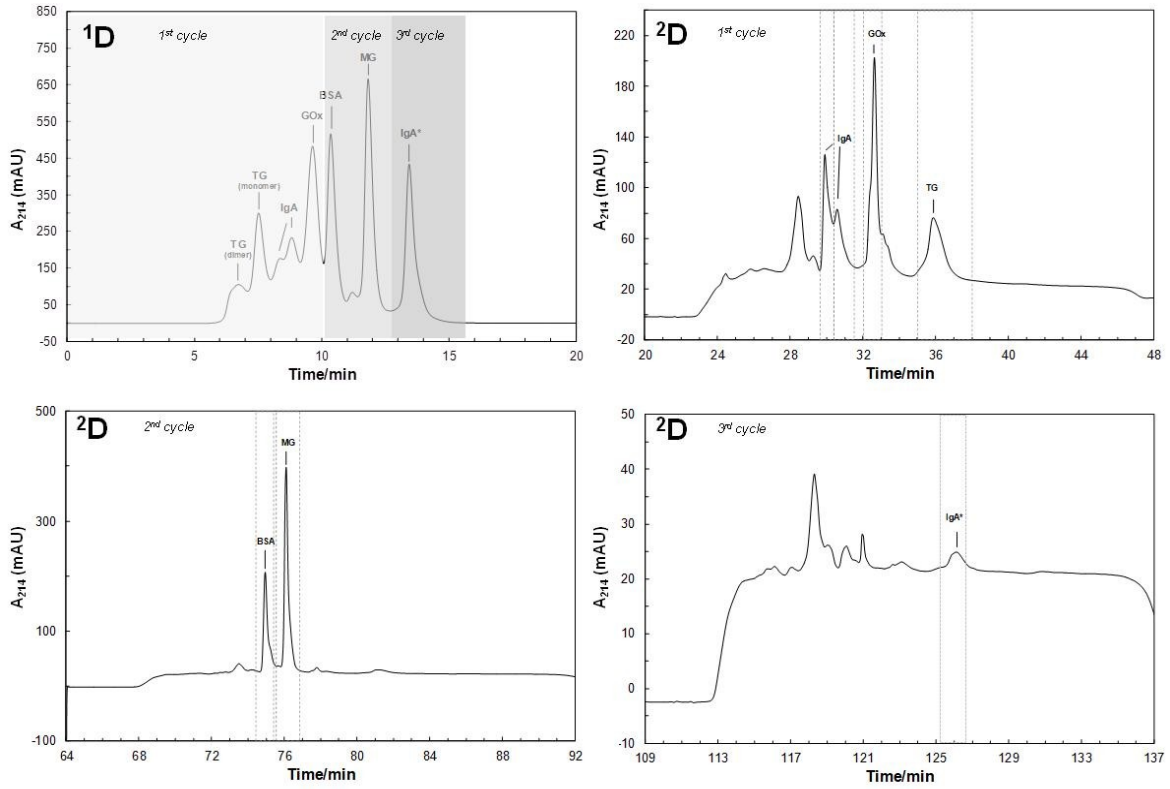


Figure 16. 2D chromatographic separation SEC-RP of a mixed standard solution containing 500 mg L^{-1} TG, 200 mg L^{-1} IgA, 500 mg L^{-1} GO_X, 500 mg L^{-1} BSA, and 500 mg L^{-1} MG. Injection volume: 15 µL. Left pump connected with column MabPac SEC-1 (150×4.6 mm) + security guard cartridge, flushed with eluent A. Right pump connected with Aeris™ widepore XB-C18 (150×4.6 mm) + security guard cartridge, flushed at 70 °C with eluents B and C. Trap column Eurosil C4 (50 x 4,6 mm) placed on the left 10-port switching valve at port connections 1 and 8.

A total analysis time of 15 min for the first separation and additional 30 min of the second dimension separation for each trapped protein block is obtained. Although the method, set for standard protein fractionating, has been described for only three different fractions in the first dimension (that represents a good compromise between peak resolution and total

analysis time), the number of cycles can be repeated more times, if necessary, when analyzing complex matrices. Indeed, the method, set for the standard proteins, is easily adaptable to the analysis of food extracts. On the basis of the effective chemical characteristics of the type of sample to be analyzed, the sampling times associated to the switching valves are properly adjusted and the number of cycles can be repeated several times to recovery all the protein peaks of each chromatographic run. Therefore, before the 2D-LC separation, it is critical to perform in advance a 1D-SEC run to get the protein assay estimation and the retention times of the proteins. The general 2D-LC algorithm, on the basis of the chromatographic behaviors of the protein peaks present in the 1D-SEC profile, is easily programmed by setting the switching times and gradient conditions in a spreadsheet. Then, the set-up of the valve switching is automatically checked and managed by software and a custom-made excel algorithm. As an example, the excel workflow set for the 2D-LC separation of a standard mix of bovine serum albumin, β-lactoglobulin and glucose oxidase by a single trapping cycle is reported in Figure 17.

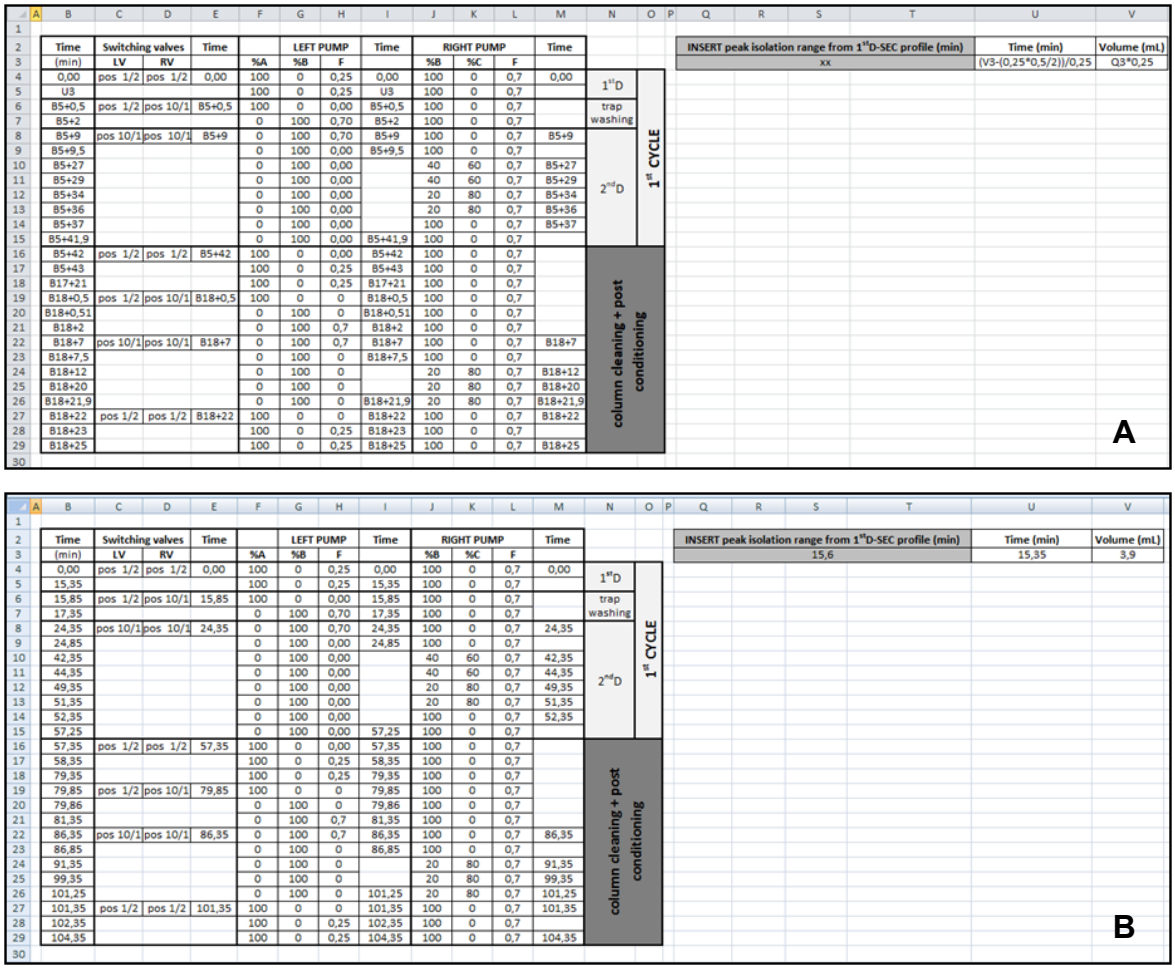


Figure 17. Instructions (A) for the generation of (B) a 2D-LC time program for the 2D-LC separation of a standard mix of bovine serum albumin (BSA), β-lactoglobulin (β-LG) and glucose oxidase (GOx) by a single trapping cycle.

In Figure 18, the corresponding 2D chromatographic profiles are also shown.

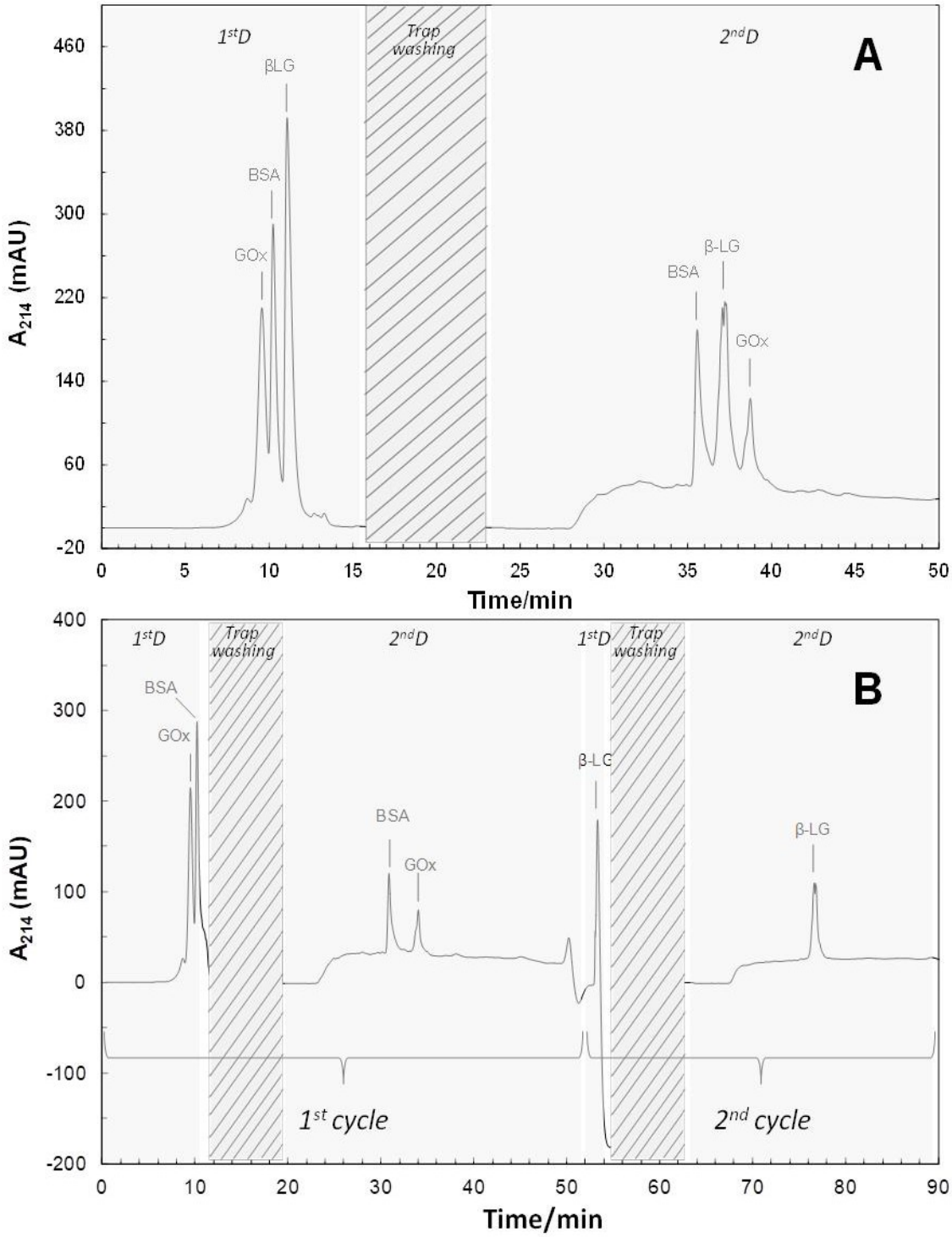


Figure 18. 2D chromatographic separations SEC-RP of a mixed standard solution containing BSA, β-LG and GOx at a concentration of 500 mg L^{-1}, each. Injection volume: 15 µL. Left pump connected with column MabPac SEC-1 (150×4.6 mm) + security guard cartridge GFC-2000 (4 x 3.0mm ID, Phenomenex), flushed with 50 mM phosphate buffer + 0.3 M NaCl at pH 6.8. Right pump connected with Aeris™ widepore XB-C18 (150×4.6 mm) + security guard cartridge, flushed with acidified water and acetonitrile with trifluoroacetic acid. Trapping column Eurosil C4 (50 x 4,6 mm, Knauer) placed on the left 10-port switching valve at port connections 1 and 8. A) Single trapping cycle; B) two-step trapping cycles.

3.2.3. Isolation and characterization of soy flour proteins

The optimized 2D-LC method has been used for the protein analysis in soy flour, known as a potential allergen source. Many methods have been used to extract plant proteins, such as traditional alkali extraction, salt extraction [155], reverse micelle extraction [156], organic solvent extraction [157], and enzyme-assisted extraction [158]. However, these methods have several latent disadvantages, such as a low extraction rate, long extraction time, complex extraction process, high extraction cost, and environmental pollution.

In the present research activity, the protein extraction procedure from soy flour samples was at first performed by adopting the innovative Microextraction by Packed Sorbent (MEPS), that is a relatively new technique for sample preparation, being a miniaturization of the conventional solid phase extraction (SPE) technique. In MEPS the sorbent material, about 4 mg of silica based particles, is introduced into the barrel of a syringe as a plug with polyethylene filters on both sides. This method is very easy to use, fully automatable, inexpensive, solvent-saver and fast, when compared to other conventional extraction methods. Many factors, such as sorbent type, volumes and composition of washing and elution solutions can affect the performance of MEPS. Moreover, additional steps for post-cleaning and re-conditioning have to be included to enable multiple uses of MEPS sorbent. Before analyzing soy flour samples, preliminary experiments by MEPS were performed with standard solutions of BSA in the "*draw-eject*" mode, using methanol and water for the conditioning step and 0.1% TFA in ACN/H_20 (80:20) for the protein elution. Different sorbents (C8, C18) were investigated; MEPS-C18 provided a higher extraction recovery than C8 sorbent (24% against 10%), nevertheless, these values show that the proteins are scarcely adsorbed on the solid phase. Also increasing the number of aspirate-dispense cycles (from 20 to 60) or loading less amount of standard solution, the adsorption did not improve, demonstrating that the MEPS technique was not useful for the extraction of intact proteins. Therefore alternative ways for the protein extraction were tested.

A rapid extraction method based on the use of an eco-friendly and biodegradable solvent for the recovery of proteins from soy flour samples was investigated. Several procedures have been described in the literature for the protein extraction from soy samples [159–163], nevertheless, the use of aqueous polyethylene glycol (PEG) as a green and biodegradable reaction medium was chosen for the recovery of the hydrophilic protein fraction, as already reported for samples at high protein content such as almonds [133]. Recently, PEG has drawn upon increasing attention as a green and environmentally

friendly solvent for its biodegradability, low flammability, stability, good miscibility with water or organic solvents, and solubility for various organic compounds, especially for PEG with a low degree of polymerization [164]. Aqueous solutions of PEG-200 (i.e. with an average molecular weight of 200 Da) assured lower viscosities, which are beneficial for mass transfer during the extraction process. Furthermore, the concentration of PEG 200 in water was found to have an impact on the protein recovery from food samples. The optimization of the extraction procedure was carried out on aliquots of the same soy flour sample commercially available in a local market. Spiked samples with BSA were prepared by adding proper amounts of the reference protein to the soy flour before the extraction process. Then, the recovery percentages were evaluated by comparing the concentration of spiked samples, determined by the external calibration regression line, with the nominal fortification level. The use of PEG-200 at different percentages by weight (from 20% to 40%) was investigated; it was noted that an increase of the extraction temperature was necessary in order to reduce viscosity when higher PEG levels than 30% were used. The optimal extraction conditions were obtained by using PEG at a concentration of 25% in an ultrasonic bath for 60 minutes at a temperature of 45 °C.

A sample clean-up by SPE with different sorbents (strong cation exchange, SCX, and reversed-phase, C18) was evaluated, following a typical extraction protocol as shown in Figure 19. A poor adsorption of proteins during the loading step was observed using SCX (sulfonic acid functionalized hydrophilic modified styrene polymer, 30 mg/1mL, Supelco) cartridges, also changing the activation conditions and the sample solvent composition (pure water or acidified by TFA or HCl). Indeed, recovery percentages, calculated both on spiked samples and by comparing the peak area of BSA standard solutions eluted from the SCX cartridge to that obtained from an equivalent protein solution not passed through the cartridge, was not more than 10%. Therefore, a solid phase treatment based on C18 cartridges (end-capped with a trimethylsilyl group in order to reduce the secondary interactions) was adopted. An elution step with aqueous methanol assured a recovery percentage of 97 ± 10% (n=6), demonstrating the efficiency of the proposed extraction procedure that results easy to use, fast, quite inexpensive and solvent-saver.

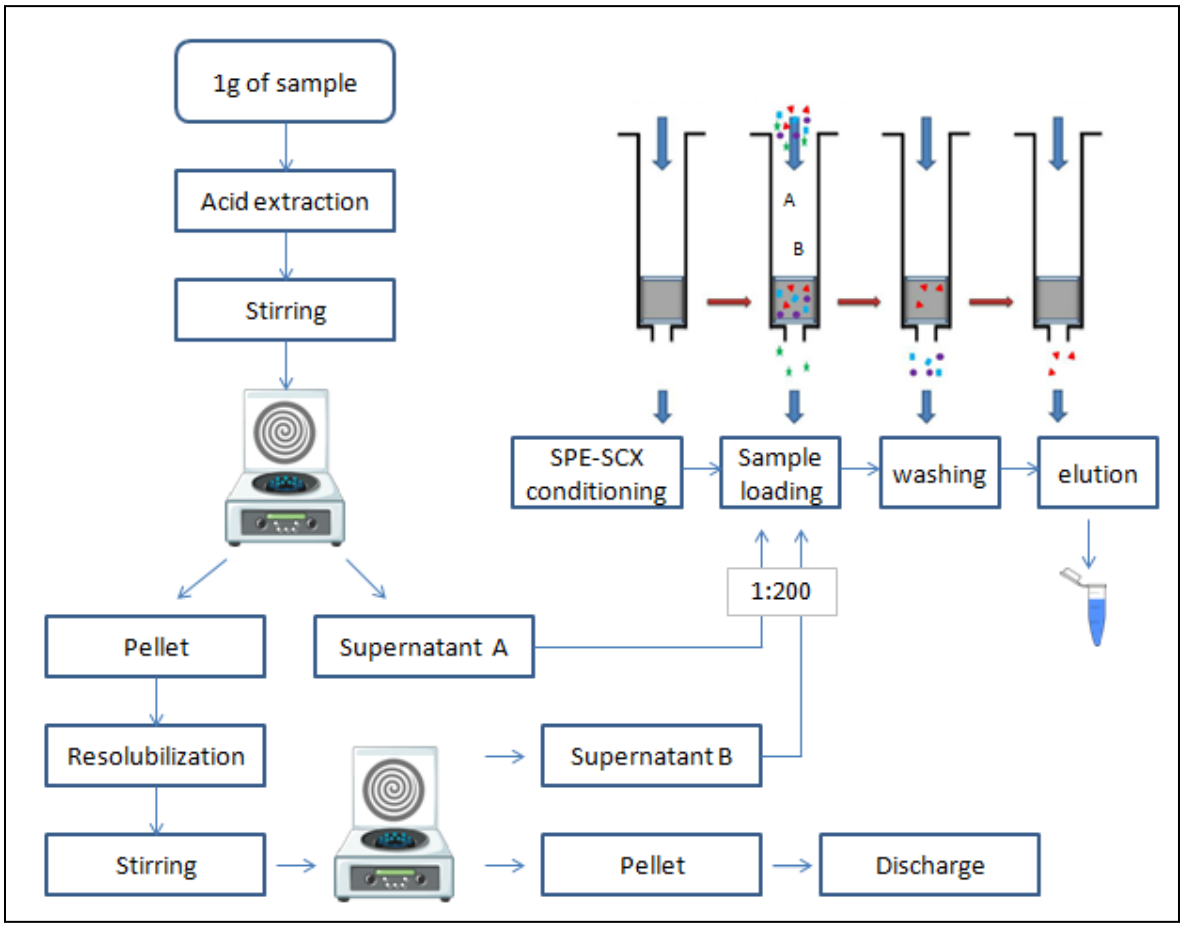


Figure 19. Schematic representation of the protein extraction procedure by SCX.

In Figure 20, the chromatographic profiles by 2D-LC of a hydrophilic protein extract obtained from soy flour are shown. Before the 2D-LC separation, the optimized 1D-SEC method was used not only to select the retention windows, but also to get an estimation of the protein identity and content. Indeed, the molecular weight of each protein peak was estimated by interpolation on the protein calibration mixture (Figure 13) and a comparison with literature data was also performed. It has been reported that soybean seeds contain a number of allergic proteins, such as lipoxidase, glycinin, and β-conglycinin, formed mainly by globulins [159,165,166]. Glycinin, also named 11S globulin, is a hexameric high molecular weight storage protein (~ 360 kDa) and the molecular weights of each of its six subunits are approximately 60 kDa; β-conglycinin, also known as 7S globulin [167], is a soybean storage protein composed of three trimers, namely α, α′ and β subunits, having a total molecular weight of approximately 220 kDa. Other allergic proteins of soy are 2S albumins, consisting of low molecular weight polypeptides characterized by a cysteine skeleton held together by four disulfide bonds. Many 2S albumins from different species have been identified as allergens in sesame seeds, walnuts, rapeseeds, cashew nuts, peanuts

and soybean [168–170]. Another small proteins, the so-called Kunitz soybean trypsin inhibitor (STI, with a molecular weight of 24 kDa), and Bowman-Birk inhibitor (BBI, with a molecular weight of 7–8 kDa), have been identified as a potent allergen capable of inducing food anaphylaxis [171,172]. These considerations, in addition to the fact that, in the present work, we have isolated the hydrophilic fraction from soy, leads us to suppose that, presumably, the peak at 11.8 min in the 1D-SEC profile (Figure 20, panel ^{1}D) corresponds to STI (~25 kDa), whereas the peak at the retention time of 12.5 min are suspected to be 2S albumin, with a molecular weight of 18 kDa. The formation of abundant low-MW peaks (< 1 kDa) at retention times higher than 15 min, after the dashed line, could be a result of protein hydrolysis.

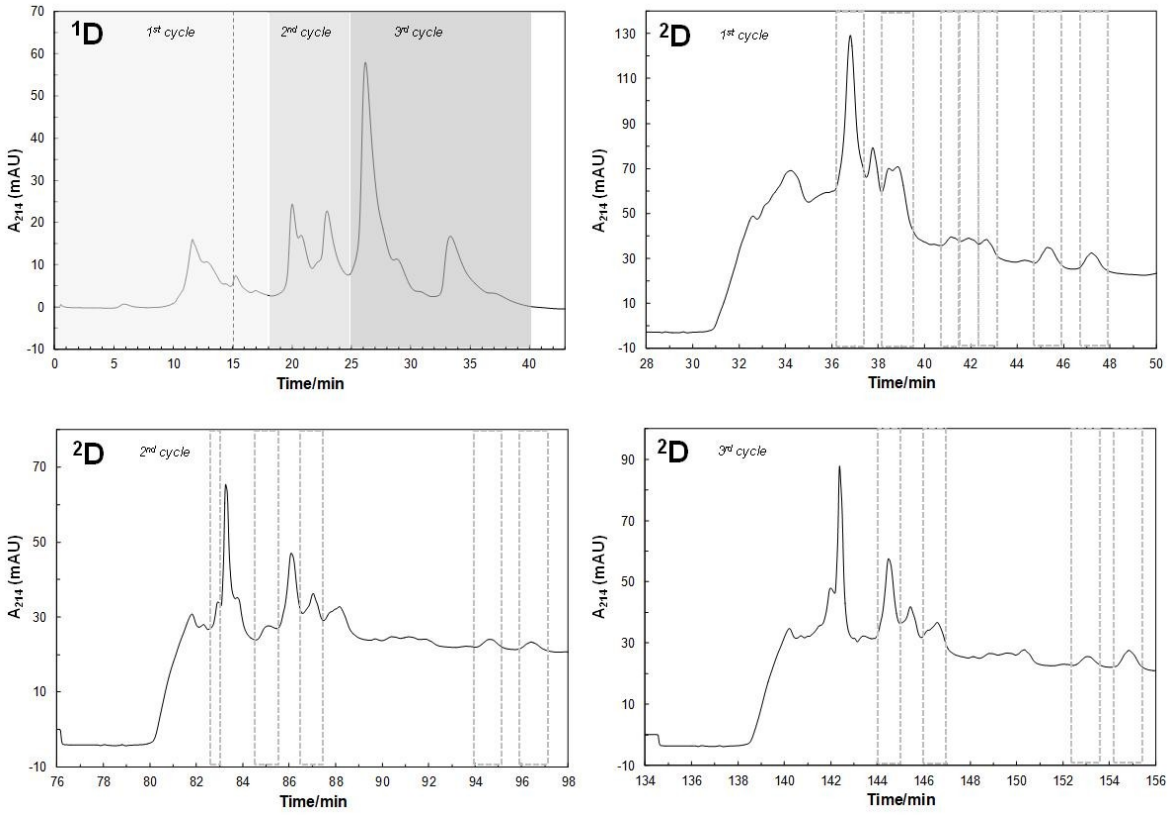


Figure 20. Chromatographic profiles of a hydrophilic protein extract obtained from soy flour sample: ^{1}D-SEC separation. Column: MabPac (150×4.6 mm) + Security Guard Cartridge (GFC-2000 4 x 3.0 mm). Eluent: 50 mM phosphate buffer + 0.3 M NaCl at pH 6.8, isocratic mode. Flow rate: 0.25 mL min^{-1}. Injection volume: 15 µL; ^{2}D chromatographic separations (first, second and third cycle): Left pump connected with column MabPac (150×4.6 mm) + Security Guard Cartridge (GFC-2000 4 x 3.0 mm). Right pump connected with Aeris™ widepore XB-C18 (150×4.6 mm) + guard column (ULTRA Cartridges UHPLC WIDEPORE C18) at 70 °C. Eurosil C4 trap column (50 x 4.6 mm) placed on the left 10-port switching valve at port connections 1 and 8. Injection volume: 15 µL.

An evaluation of the protein content, isolated and collected from soy flour, was performed by interpolation on the regression line of BSA by 1D-RP. A protein amount of ~ 4 μg was estimated, confirming that the proposed 2D-LC method allows to collect protein material as intact species, enough for further investigations.
The 2D-LC cycles can be repeated several times with the aim of recovering all the protein peaks from the first dimension chromatogram. Then, by the second dimension separation, a sufficient amount of protein (not only the most abundant but also the low-intensity peaks) can be collected in a few mL of organic solvent that can be easily removed by evaporation. Finally, each isolated protein can be reconstituted in a proper solvent or aqueous buffer for its molecular characterization by mass spectrometry or allergy testing.

3.3. Ultra high resolution analysis of fennel proteins

The last part of the research activity of the Ph.D. project was focused on the fennel protein analysis by ultra-high-resolution FT-ICR mass spectrometry. As already mentioned in the previous paragraphs, mass spectrometry (MS) based methods have been suggested as confirmatory tools for an accurate protein identification in the field of food quality and safety, over recent years [173]. Emphasis is placed on food processing, in the determination of possible contaminants like bacteria and fungi, and in allergen detection [174]. Among the different proteomic strategies, bottom-up analysis remains the workhorse for protein characterization; nevertheless, it results in a greatly increased complexity of the generated peptide mixture, requiring highly sensitive and efficient methods which 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 [175]. Moreover, the large dynamic range and unmatched sensitivity of FTICR-MS currently provides the highest quality data for molecular identifications [176]. Nevertheless, to date, the advantages of FT-ICR/MS techniques have not been fully exploited for proteome investigations and among the exponentially increasing number of proteomics literature works, only a limited number of applications have been reported, by direct infusion analysis [177] or coupled with chromatographic separations [178,179].
The first way of using mass spectrometry data for proteomic determinations is known as peptide mass fingerprint (PMF), based on the mass measurement of enzymatically digested

proteins and the comparison of such data with theoretical fingerprints from protein databases [180,181]. During the last decades, PMF database searching has become the preferred method of choice for high throughput protein identifications. Nevertheless, the risk of false positives is greater with PMF than other mass spectrometry analyses such as MS/MS ion search [182]. As for any analytical process, uncertainty in a measurement translates into uncertainty in the derived results from that measurement. When using PMF, false discovery is closely related to mass measurement error; inaccuracies in the measurement of peptide m/z ratios have the unavoidable consequence of leading to errors in sequence associations at both the peptide and protein levels. An effective way to minimize the possibility of incorrect sequence assignments is the use of accurate mass measurement (i.e., with mass error < 10 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 FT-ICR mass analyzer is used, allowing to assign detected m/z-values to their unambiguous peptide sequences and decreasing the score and peptide sequence coverage of the highest ranked random protein match from the database [183].

3.3.1. Liquid Chromatography analyses and MS characterization by FT-ICR

Starting from the 2DLC method, properly developed, optimized and applied to food products known as potential allergen sources (such as egg and soy flour), the isolation and characterization of fennel proteins were also performed. Indeed, although fennel (*Foeniculum vulgare Mill.*) has attracted attention as a medicinal plant with an enormous amount of health benefits [184], it is recently recognized as an allergenic source, especially in the Mediterranean area [185].

Liquid Chromatography analyses. Before MS characterization, chromatographic analyses were performed on fennel extracts by using size-exclusion MAbPac column under isocratic conditions with 50 mM phosphate buffer + 0.3 M NaCl at pH 6.8 at a flow rate of 0.25 mL min^{-1}. For the quantitative analyses, calibration data were obtained by three series of 1D-SEC analyses on three different days, by injecting seven working standard solutions of BSA (injection volume 15 µL) at concentrations of 5, 10, 50, 100, 300, 500 and 1000

mg L^{-1} (see Table 5). The protein assay estimation in the fennel extract was then performed by interpolation on the BSA calibration line, using the sum of peak areas associated with each peak in the SEC fennel profile. A protein amount of 6.3±0.5 g/L was obtained. Moreover, for the estimation of the molecular weight (MW) of unknown fennel proteins the log plot of the molecular weights of standard proteins against the retention time (Figure 13) was used.

In Figure 21 the SEC separation of fennel extract is shown. Individual protein fractions were collected in-line, in correspondence of their retention time-windows, for the subsequent proteome-wide analysis. Hence, after solvent evaporation, the most abundant protein fractions at low molecular weight were isolated, tryptically digested and submitted to proteomic FT-ICR MS characterization.

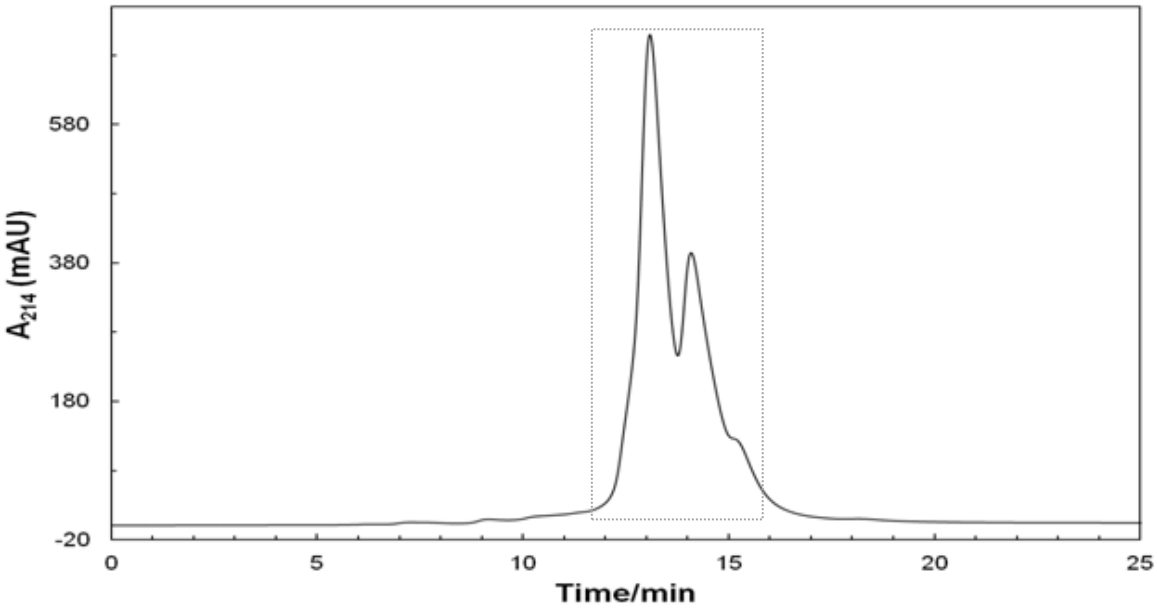


Figure 21. 1D-SEC separation of fennel extract fraction at low MWs. Column: MabPac SEC-1 (150×4.6 mm) + Security Guard Cartridge (GFC-2000 4 x 3.0 mm). Eluent: A) 50 mM phosphate buffer + 0.3 M NaCl at pH 6.8, isocratic mode. Injection volume: 15 µL. Flow rate: 0.25 mL min^{-1}.

The MS spectrum of the tryptic peptide mixture is displayed in Figure 22. The FT-ICR analysis of the low-MW fractions has led to the identification of three fennel proteins with molecular weights ranging from 4 to 8 kDa, belonging to the ribosomal protein class (Table 7).

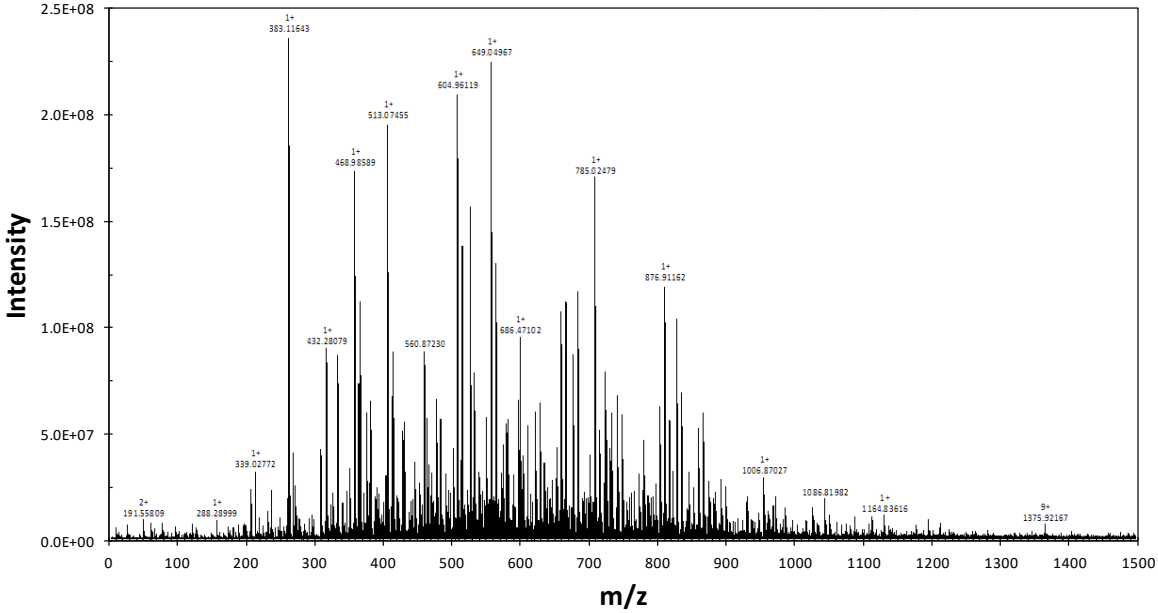


Figure 22. ESI(+)-FTICR mass spectrum of a tryptic mixture (acetonitrile/water (70:30, v/v) + 0.2% FA) of a fennel extract fraction at low MWs.

Table 7. List of the low-MW fennel proteins identified by FT-ICR.

Entry Name	Protein Code	Sequence Coverage %	MASCOT Score	No. Matched Peptides	MW (Da)	Length (No. residues)
Ribosomal protein L32 (chloroplast)	YP_009235928.1	66	107	6	5997	53
Ribosomal protein L33 (chloroplast)	YP_009235900.1	33	116	12	7687	66
Ribosomal protein L36 (chloroplast)	YP_009235913.1	70	157	9	4460	37

Afterward, a shot-gun analysis was performed for the identification of all the proteins present in the whole raw fennel extract (without prior fractionation by 2DLC), taking advantage of the elevated sensitivity and mass accuracy of FTICR-MS.

3.3.2. Shotgun analysis of fennel extract by direct infusion FT-ICR

Before analyzing fennel extracts, a UBIQ standard solution was tested to set all the experimental parameters for the direct infusion FT-ICR analysis as well as for the database searching. To this goal, several experiments were performed for the optimization of the

MS/MS fragmentation parameters such as accumulation time in the linear hexapole ion trap and collisional energy. The best results were obtained by setting a value of 50 scans, an ion accumulation time of 1.8 s and collision energy of 10 eV. The quality of UBIQ mass spectrum (Figure 23), dominated by a few high-intensity principal signals, was confirmed by the excellent results obtained by Peptide Mass Fingerprint Mascot database searching. As shown in Figure 24, a protein score of 152 was obtained associated to a percentage coverage of 80%, with the identification of 7 peptides uniformly distributed along the protein sequence.

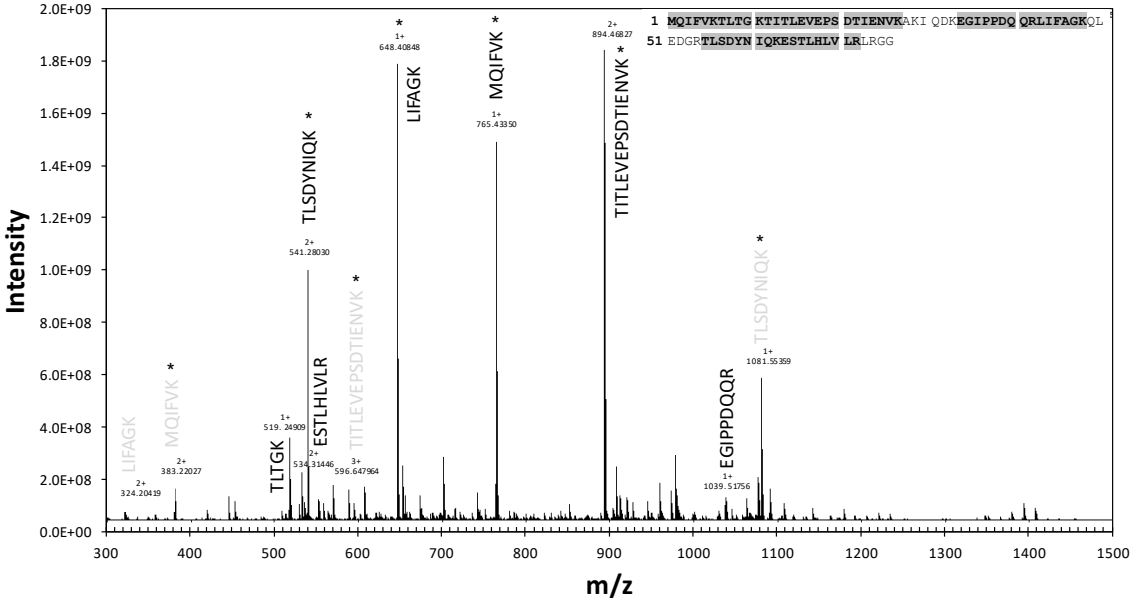


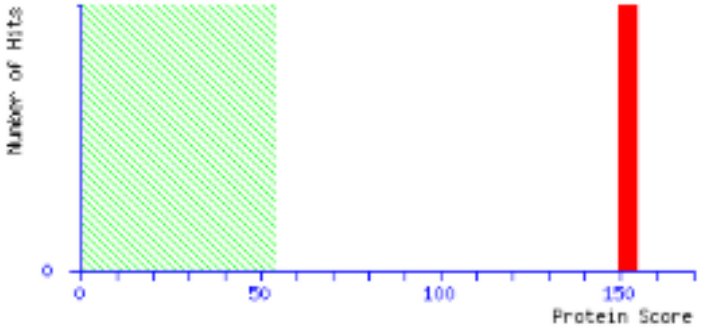
Figure 23. ESI(+)-FTICR mass spectrum of a tryptic mixture (acetonitrile/water (70:30, v/v) + 0.2% FA) of UBIQ standard. Stars indicate the signals of the peptides later subjected to MS/MS analysis.

MATRIX SCIENCE Mascot Search Results

```
User            : MARIA TERESA MELFI
Email           : mariateresa.melfi@unifg.it
Search title    :
Database        : SwissProt 2019_08 (560823 sequences; 201585439 residues)
Taxonomy        : Other mammalia (13203 sequences)
Timestamp       : 3 Oct 2019 at 13:55:35 GMT
Top Score       : 152 for UBIQ_CAMDR, Ubiquitin OS=Camelus dromedarius OX=9838 PE=3 SV=2
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 54 are significant (p<0.05).



Protein sequence coverage: 80%

Matched peptides shown in ***bold red***.

```
  1 MQIFVKTLTG KTITLEVEPS DTIENVKAKI QDKEGIPPDQ QRLIFAGKQL
 51 EDGRTLSDYN IQKESTLHLV LRLRGG
```

Unformatted sequence string: 76 residues (for pasting into other applications).

Sort by ◉ residue number ○ increasing mass ○ decreasing mass
Show ◉ matched peptides only ○ predicted peptides also

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	M	Peptide
1 - 6	765.4326	764.4254	764.4255	-0.0001	0	-.MQIFVK.T
1 - 6	765.4335	764.4262	764.4255	0.0007	0	-.MQIFVK.T
1 - 6	765.4335	764.4262	764.4255	0.0007	0	-.MQIFVK.T
12 - 27	1787.9278	1786.9205	1786.9200	0.0005	0	K.TITLEVEPSDTIENVK.A
34 - 42	1039.5168	1038.5095	1038.5094	0.0001	0	K.EGIPPDQQR.L
43 - 48	648.4077	647.4004	647.4006	-0.0002	0	R.LIFAGK.Q
43 - 48	648.4085	647.4012	647.4006	0.0006	0	R.LIFAGK.Q
43 - 48	648.4085	647.4012	647.4006	0.0006	0	R.LIFAGK.Q
49 - 54	717.3533	716.3460	716.3453	0.0007	0	K.QLEDGR.T
49 - 54	717.3533	716.3460	716.3453	0.0007	0	K.QLEDGR.T
55 - 63	1081.5527	1080.5454	1080.5451	0.0003	0	R.TLSDYNIQK.E
64 - 72	1067.6209	1066.6136	1066.6135	0.0001	0	K.ESTLHLVLR.L

Figure 24. Mascot search results view of UBIQ Protein Identification by Peptide Mass Fingerprint search mode. NCBI database; Other mammalia category; No variable and fixed modifications; enzyme trypsin; 2 allowed missed cleavages; peptide mass tolerance of 5 ppm.

For selected precursor ions, collision-induced dissociation (CID) analyses were also performed, therefore the most intense ions were isolated for fragmentation. The collisional energy was optimized for individual molecular species to achieve maximum S/N ratio. Systematic experiments MS/MS were performed by changing the collision energy (eV) and monitoring the intensities of the fragment ions. As an example, Figure 25 shows the MS/MS spectrum of the triply charged precursor ion at m/z 597.30 corresponding to the 16 residue long peptide TITLEVEPSDTIENVK. CID fragmentation yielded 10 inter-residue cleavages and 13 ion fragments of b/y-type with the identification of 4 complementary couples. Therefore, a fragmentation efficiency of 43% was obtained, defined 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.

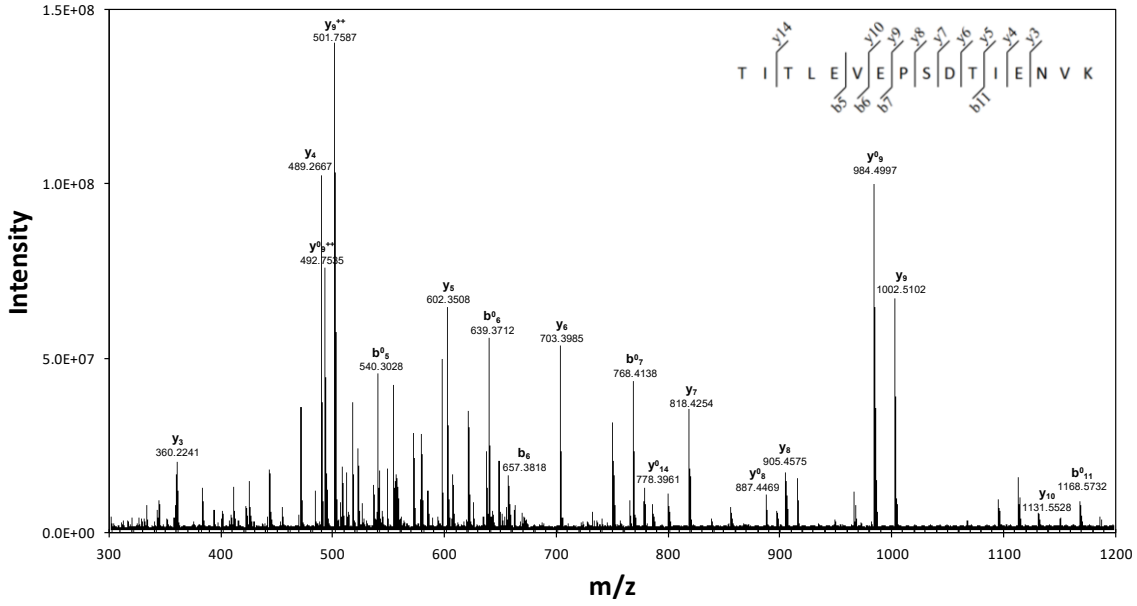
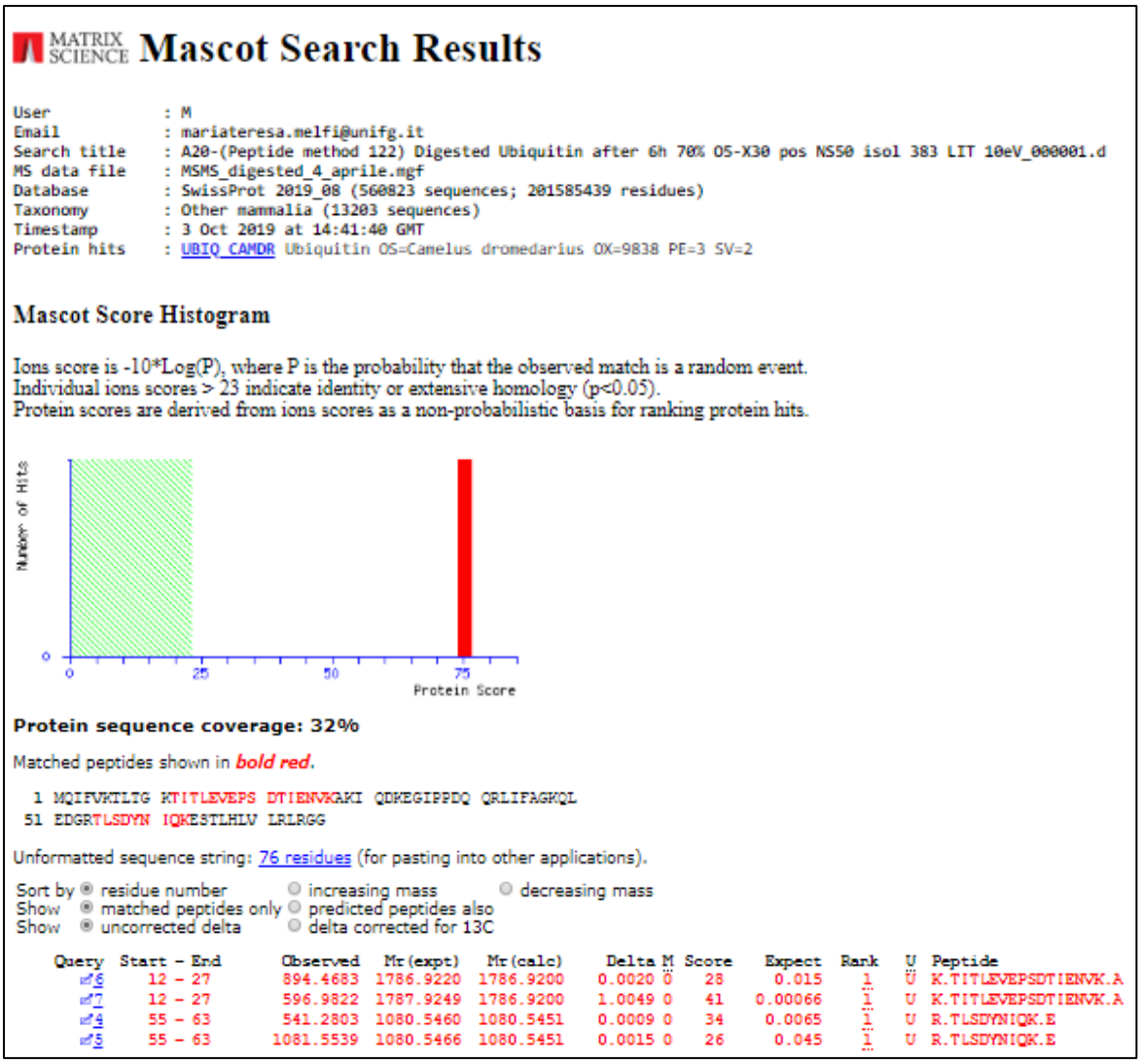


Figure 25. Collision induced dissociation (CID) mass spectrum of the precursor peptide at m/z 597.2982 from UBIQ standard . Collisional energy = 10 eV.

The acquired MS/MS datasets were used in the database searching for protein identification but a poor peptide match was obtained: only two peptides were successfully sequenced by MS/MS ion search against the seven peptides identified by PMF as shown in Figure 26.



MATRIX SCIENCE **Mascot Search Results**

User : M
Email : mariateresa.melfi@unifg.it
Search title : A20-(Peptide method 122) Digested Ubiquitin after 6h 70% 05-X30 pos NS50 isol 383 LIT 10eV_000001.d
MS data file : MSMS_digested_4_aprile.mgf
Database : SwissProt 2019_08 (560823 sequences; 201585439 residues)
Taxonomy : Other mammalia (13203 sequences)
Timestamp : 3 Oct 2019 at 14:41:40 GMT
Protein hits : UBIQ_CAMDR Ubiquitin OS=Camelus dromedarius OX=9838 PE=3 SV=2

Mascot Score Histogram

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 23 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.

Protein sequence coverage: 32%

Matched peptides shown in ***bold red***.

1 MQIFVKTLTG KTITLEVEPS DTIENVKAKI QDKEGIPPDQ QRLIFAGKQL
51 EDGRTLSDYN IQKESTLHLV LRLRGG

Unformatted sequence string: 76 residues (for pasting into other applications).

Sort by residue number increasing mass decreasing mass
Show matched peptides only predicted peptides also
Show uncorrected delta delta corrected for 13C

Query	Start – End	Observed	Mr(expt)	Mr(calc)	Delta	M	Score	Expect	Rank	U	Peptide
6	12 – 27	894.4683	1786.9220	1786.9200	0.0020	0	28	0.015	1	U	K.TITLEVEPSDTIENVK.A
7	12 – 27	596.9822	1787.9249	1786.9200	1.0049	0	41	0.00066	1	U	K.TITLEVEPSDTIENVK.A
4	55 – 63	541.2803	1080.5460	1080.5451	0.0009	0	34	0.0065	1	U	R.TLSDYNIQK.E
5	55 – 63	1081.5539	1080.5466	1080.5451	0.0015	0	26	0.045	1	U	R.TLSDYNIQK.E

Figure 26. Mascot search results view of UBIQ Protein Identification by MS/MS Ion search mode. NCBI database; Other mammalia category; No variable and fixed modifications; enzyme trypsin; 2 allowed missed cleavages; peptide mass tolerance of 5 ppm.

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 peptide mass fingerprint (PMF) approach using data from MS spectra of tryptic protein digests, taking full advantage of the high resolving power and accuracy of the FT-ICR mass analyzer.

For the identification of all the proteins present in the fennel sample (without prior fractionation by liquid chromatography), a shot-gun analysis was performed on the peptide mixture obtained by the enzymatic digestion of the whole raw extract. Considering the complexity of the fennel extract whose MS spectrum, shown in Figure 27, is characterized by an elevated number of ion signals exceeding the maximum limit of 1200 set for the free MASCOT database searching, the MS spectrum was reduced to a peak list, including only

mass data associated with the most intense peptide ions (intensity threshold of 10^6 a.u. and signal-to-noise ratio of 4).

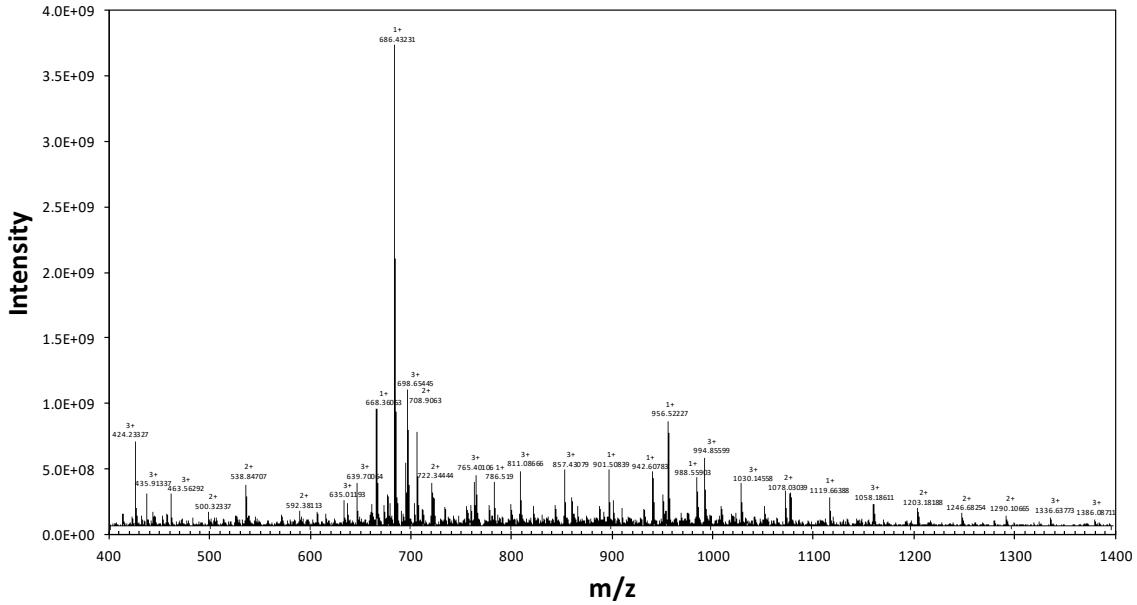


Figure 27. ESI(+)-FTICR mass spectrum of a tryptic mixture (acetonitrile/water (70:30, v/v) + 0.2% FA) of a fennel extract.

This mass list was submitted to PMF searching, but no statistically significant hits were observed, probably as a consequence 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 [183,186]. Indeed, in large data sets, there are likely to be several shared mass values, that match to more than one of the proteins in the mixture. In addition, although details of the Mascot search engine and 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 6686534 sequences - October 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 then adopted, as reported in the following paragraph.

3.3.3. Characterization of fennel proteins

Starting from the NCBI protein database for fennel (*Foeniculum vulgare*) containing 231 proteins (updated to 3 July 2019), also including redundant/partial/isoform sequences, a custom-made proteome database of 92 fennel proteins was generated, as reported in Table 8, where the proteins are listed in increasing order of MWs.

Table 8. List of 92 proteins of fennel (*Foeniculum vulgare*) from the not-redundant custom-made proteome database.

PROTEIN name	NCBI code	MW (Da)
RecName: Full=Non-specific lipid-transfer protein; Short=LTP; AltName: Allergen=Foe v 3	B3EWP9.1	1186
RecName: Full=Pathogenesis-related protein; Short=PRP; AltName: Allergen=Foe v 1	C0HJB6.1	1217
cytochrome b6/f complex subunit VIII (chloroplast)	YP_009235872.1	3170
cytochrome b6/f complex subunit VI (chloroplast)	YP_009235897.1	3359
PsbA, partial	AAT06832.1	3749
photosystem II protein T (chloroplast)	YP_009235906.1	3818
photosystem I subunit VIII (chloroplast)	YP_009235889.1	3951
cytochrome b6/f complex subunit V (chloroplast)	YP_009235898.1	4130
photosystem II protein J (chloroplast)	YP_009235893.1	4131
photosystem II protein I (chloroplast)	YP_009235863.1	4168
photosystem II protein M (chloroplast)	YP_009235873.1	4311
photosystem II cytochrome b559 beta subunit (chloroplast)	YP_009235895.1	4424
ribosomal protein L36 (chloroplast)	YP_009235913.1	4460
photosystem II protein L (chloroplast)	YP_009235894.1	4497
photosystem II protein N (chloroplast)	YP_009235907.1	4722
photosystem I subunit IX (chloroplast)	YP_009235899.1	4809
ribosomal protein L32 (chloroplast)	YP_009235928.1	5997
t-anol/isoeugenol synthase, partial	QDG10108.1	6426
photosystem II protein Z (chloroplast)	YP_009235876.1	6511
photosystem II protein K (chloroplast)	YP_009235862.1	6972
ribosomal protein L33 (chloroplast)	YP_009235900.1	7692
photosystem II phosphoprotein (chloroplast)	YP_009235908.1	7754
ATP synthase CF0 subunit III (chloroplast)	YP_009235866.1	7990
photosystem I subunit VII (chloroplast)	YP_009235931.1	8939
ribosomal protein S16 (chloroplast)	YP_009235861.1	9124
translational initiation factor 1 (chloroplast)	YP_009235914.1	9181
photosystem II cytochrome b559 alpha subunit (chloroplast)	YP_009235896.1	9424
ribosomal protein S19 (chloroplast)	YP_009235920.1	10545

ribosomal protein L23 (chloroplast)	YP_009235922.1	10724
ribosomal protein S15 (chloroplast)	YP_009235937.1;	10796
NADH-plastoquinone oxidoreductase subunit 4L (chloroplast)	YP_009235932.1	11217
ribosomal protein S14 (chloroplast)	YP_009235877.1	11750
ribosomal protein S18 (chloroplast)	YP_009235901.1	11986
RNA polymerase beta subunit, partial (chloroplast)	AFK09970.1	12859
ribosomal protein L14 (chloroplast)	YP_009235916.1	13511
ribosomal protein S12 (chloroplast)	YP_009235858.1	13738
RNA polymerase beta' subunit, partial (chloroplast)	AFK10055.1	13824
NADH-plastoquinone oxidoreductase subunit 3 (chloroplast)	YP_009235884.1	13841
ribosomal protein S11 (chloroplast)	YP_009235912.1	14972
ribosomal protein L16 (chloroplast)	YP_009235917	15289
ATP synthase CF1 epsilon subunit (chloroplast)	YP_009235885.1	15292
ribosomal protein L20 (chloroplast)	YP_009235902.1	15295
ribosomal protein S8 (chloroplast)	YP_009235915.1	15829
ribosomal protein S7 (chloroplast)	YP_009235925.1	17357
cytochrome b6/f complex subunit IV (chloroplast)	YP_009235910.1	17459
NADH-plastoquinone oxidoreductase subunit J (chloroplast)	YP_009235882.1	18605
ribosomal protein L22 (chloroplast)	YP_009235919.1	18796
NADH-plastoquinone oxidoreductase subunit 6 (chloroplast)	YP_009235933.1	19157
ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit, partial (plastid)	APC26521.1	19291
NADH-plastoquinone oxidoreductase subunit I (chloroplast)	YP_009235934.1	19567
photosystem I assembly protein Ycf3 (chloroplast)	YP_009235880.1	19600
RNA polymerase C, partial (chloroplast)	ACB88304.1	19693
ATP synthase CF0 subunit I (chloroplast)	YP_009235865.1	20527
homogentisate geranylgeranyl transferase, partial	ADG26669.1	20569
photosystem I assembly protein ycf4 (chloroplast)	YP_009235890.1	21194
clp protease proteolytic subunit (chloroplast)	YP_009235904.1	22188
ribosomal protein S4 (chloroplast)	YP_009235881.1	23341
ribosomal protein S3 (chloroplast)	YP_009235918.1	25095
NADH-plastoquinone oxidoreductase subunit K (chloroplast)	YP_009235883.1	25448
phenylalanine ammonia lyase, partial	QDG10107.1	25740
ribosomal protein S2 (chloroplast)	YP_009235868.1	26874
chloroplast envelope membrane protein (chloroplast)	YP_009235891.1	26989
ATP synthase CF0 subunit IV (chloroplast)	YP_009235867.1	27023
maturase K, partial (chloroplast)	ACB88392.1	27760
ribosomal protein L2 (chloroplast)	YP_009235921.1	30100
cytochrome f (chloroplast)	YP_009235892.1	35170
cytochrome c heme attachment protein (chloroplast)	YP_009235929.1	36869

delta12-fatty acid acetylenase, partial	AAO38034.1	37861
RNA polymerase alpha subunit (chloroplast)	YP_009235911.1	38479
photosystem II protein D1 (chloroplast)	YP_009235859.1	38937
photosystem II protein D2 (chloroplast)	YP_009235874.1	39550
NADH-plastoquinone oxidoreductase subunit 1 (chloroplast)	YP_009235935.1	40317
NADH-plastoquinone oxidoreductase subunit 4 (chloroplast)	YP_009235930.1	42191
NADH-plastoquinone oxidoreductase subunit 7 (chloroplast)	YP_009235936.1	45734
photosystem II CP43 chlorophyll apoprotein (chloroplast)	YP_009235875.1	51879
ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit (chloroplast)	YP_009235887.1	52596
ATP synthase CF1 beta subunit (chloroplast)	YP_009235886.1	53589
Acetyl-CoA carboxylase carboxyltransferase beta subunit (chloroplast)	YP_009235888.1	55333
ATP synthase CF1 alpha subunit (chloroplast)	YP_009235864.1	55426
photosystem II CP47 chlorophyll apoprotein (chloroplast)	YP_009235905.1	56004
NADH-plastoquinone oxidoreductase subunit 2 (chloroplast)	YP_009235924.1	56654
maturase K (chloroplast)	YP_009235860.1	61157
hypothetical chloroplast RF1 (chloroplast)	YP_009235926.1	74606
RNA polymerase beta (chloroplast)	YP_009235870.1	78281
photosystem I P700 apoprotein A2 (chloroplast)	YP_009235878.1	82345
photosystem I P700 apoprotein A1 (chloroplast)	YP_009235879.1	83158
NADH-plastoquinone oxidoreductase subunit 5 (chloroplast)	YP_009235927.1	85465
RNA polymerase beta subunit (chloroplast)	YP_009235871.1	120866
nuclear matrix constituent protein 1-like	BAF64423.1	128731
RNA polymerase beta' subunit (chloroplast)	YP_009235869.1	157541
hypothetical chloroplast RF1 (chloroplast)	YP_009235938.1	217228
hypothetical chloroplast RF2 (chloroplast)	YP_009235923.1	247783

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 spices (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, it is reported in the literature that fennel allergic patients could also suffer from allergy to *Apiaceae* spices and birch and/or mugwort pollinosis, due to the presence of analogues proteins [66,68,70,187]. The list of the common allergens coming from

Apiaceae spices or associated with mugwort-birch-pollen-allergy-syndrome is reported in Table 9.

Table 9. Common allergens in *Apiaceae* spices and proteins associated to mugwort-birch-pollen-allergy-syndrome.

Organism	Protein name	Swiss-Prot code	MW (Da)
Apium graveolens (Celery)	Allergen Api g 5	P81943	9441
Apium graveolens (Celery)	Profilin	Q9XF37	14276
Apium graveolens (Celery)	Major allergen Api g 1, isoallergen 1	P49372	16321
Apium graveolens (Celery)	Major allergen Api g 1, isoallergen 2	P92918	17090
Apium graveolens (Celery)	Cofactor-independent phosphoglyceromutase	Q9SDL3	61125
Petroselinum crispum (Parsley)**	Pathogenesis-related protein 1	Q40795	16356
Daucus carota (Wild carrot) ***	Non-specific lipid-transfer protein	P27631	12504
Betula pendula (European white birch)	Major pollen allergen Bet v 1-A	P15494	17571
Betula pendula (European white birch)	Major pollen allergen Bet v 1-B	P45431	17537
Artemisia vulgaris (Mugwort)	Major pollen allergen Art v 1	Q84ZX5	13404

** A more specific fennel but partial PRP appears in the UniProt Knowledgebase under accession number C0HJB6
*** A more specific fennel but partial LTP appears in the UniProt Knowledgebase under accession number B3EWP9

For each protein of the custom-made database, a complete list of theoretical peptide sequences coming 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. As an example, the in-silico enzymatic digestion performed on the photosystem II protein Z (chloroplast) (NCBI code YP_009235876.1) is reported in Figure 28.

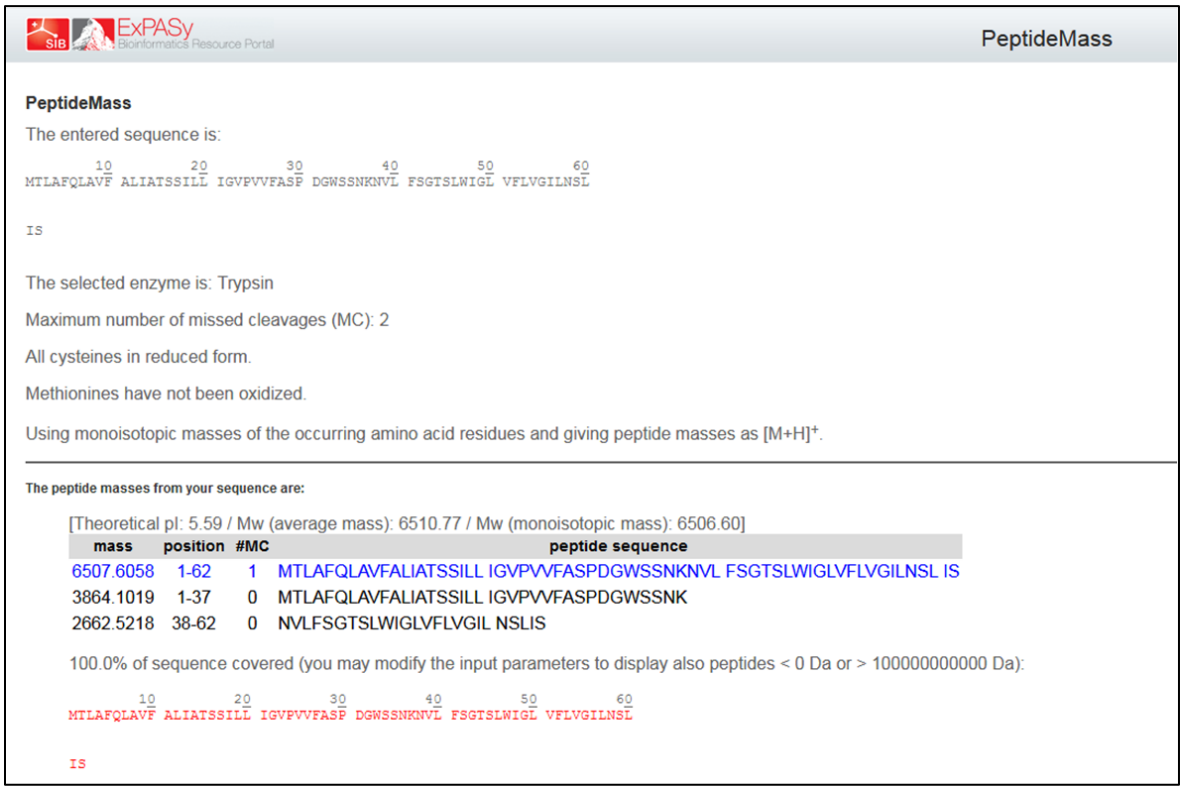


Figure 28. In-silico tryptic digestion view of photosystem II protein Z (chloroplast) from NCBI database by PeptideMass on-line tool from UniProtKB.

A filtering criterion, based on a custom-designed MATLAB algorithm, was applied to match the experimental m/z ratios with the in-silico enzymatic digestion data (in the form of $[M+H]^+$, $[M+2H]^{2+}$ and $[M+3H]^{3+}$ ions). 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 ratios (R= 400,000; 250,000; 190,000; 150,000; 130,000; 100,000 at 400; 600; 800, 1000; 1200; 1400 m/z, respectively). 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 ratios. For example, an error of 5 ppm at m/z 1000 is equivalent to 5 mDa error, which is reduced to 2.5 mDa at m/z 500 and is further reduced to 1.25 mDa at m/z 250. Thus, the true mDa error depends on the examined m/z ion to be searched in the database. The reduction of considered mDa search tolerance mass error is important, especially for low m/z ratios < 800, since the mass resolving power of FT-ICR technique increases in a mathematical power function when moving to lower m/z ratios, 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 to strike a balance between greater information capture and reduced number of incorrect sequence assignments, minimizing the risk of false-positive results.

In Figure 29 the MS spectrum of the fennel tryptic peptide mixture is reported; the amino acidic sequences matching the in silico peptide digests were also displayed.

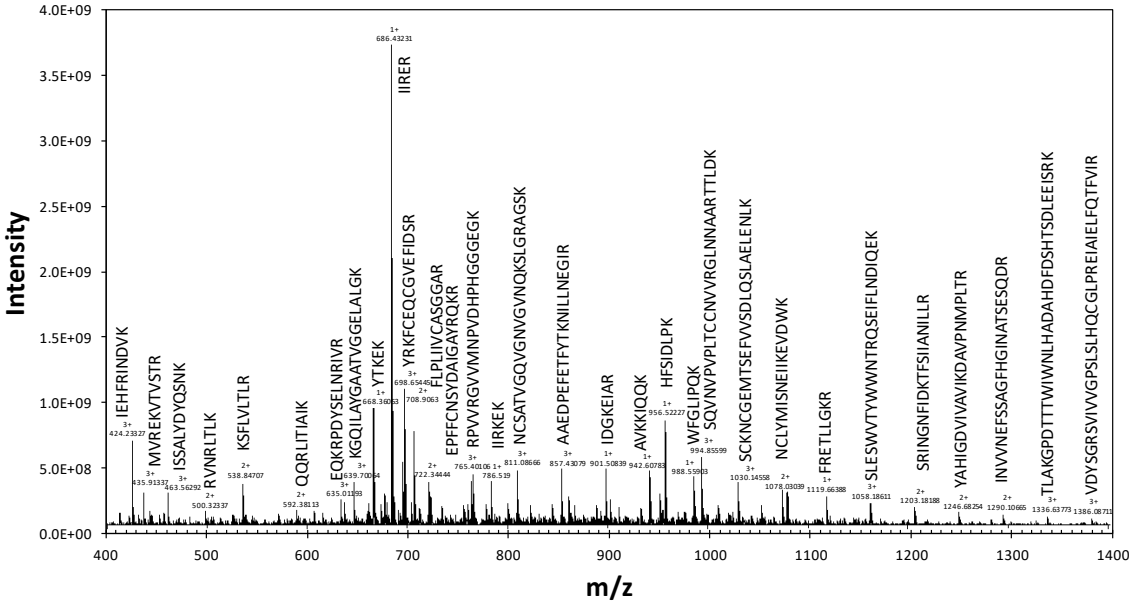
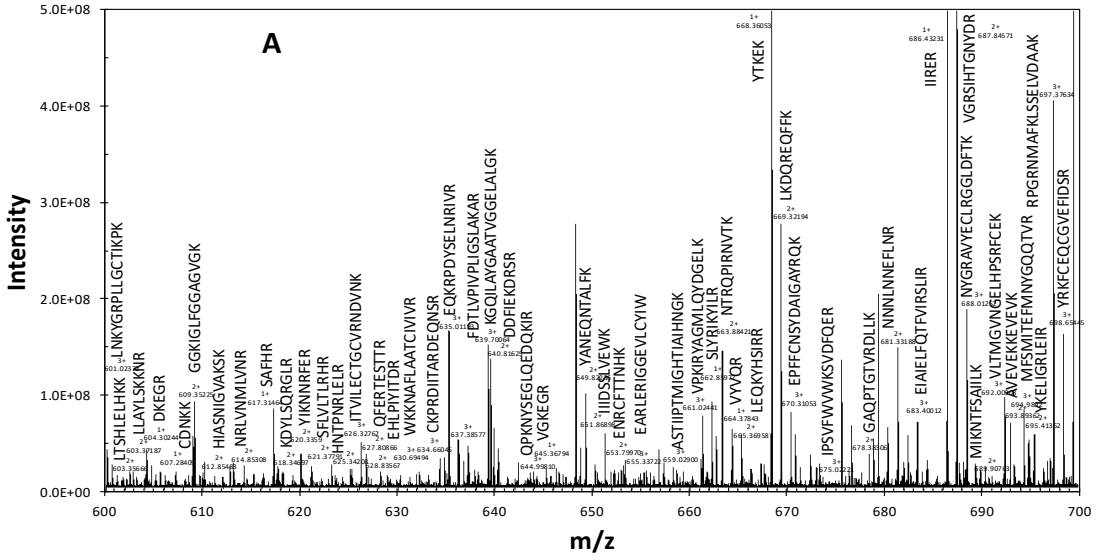


Figure 29. ESI(+)-FTICR mass spectrum of a tryptic mixture (acetonitrile/water (70:30, v/v) + 0.2% FA) of a fennel extract. Sequences matching the in silico peptide digests are shown.

Figure 30 shows the expanded views of the fennel MS spectrum in the mass segments from 600-700 (panel A) to 1000-1200 amu (panel D).



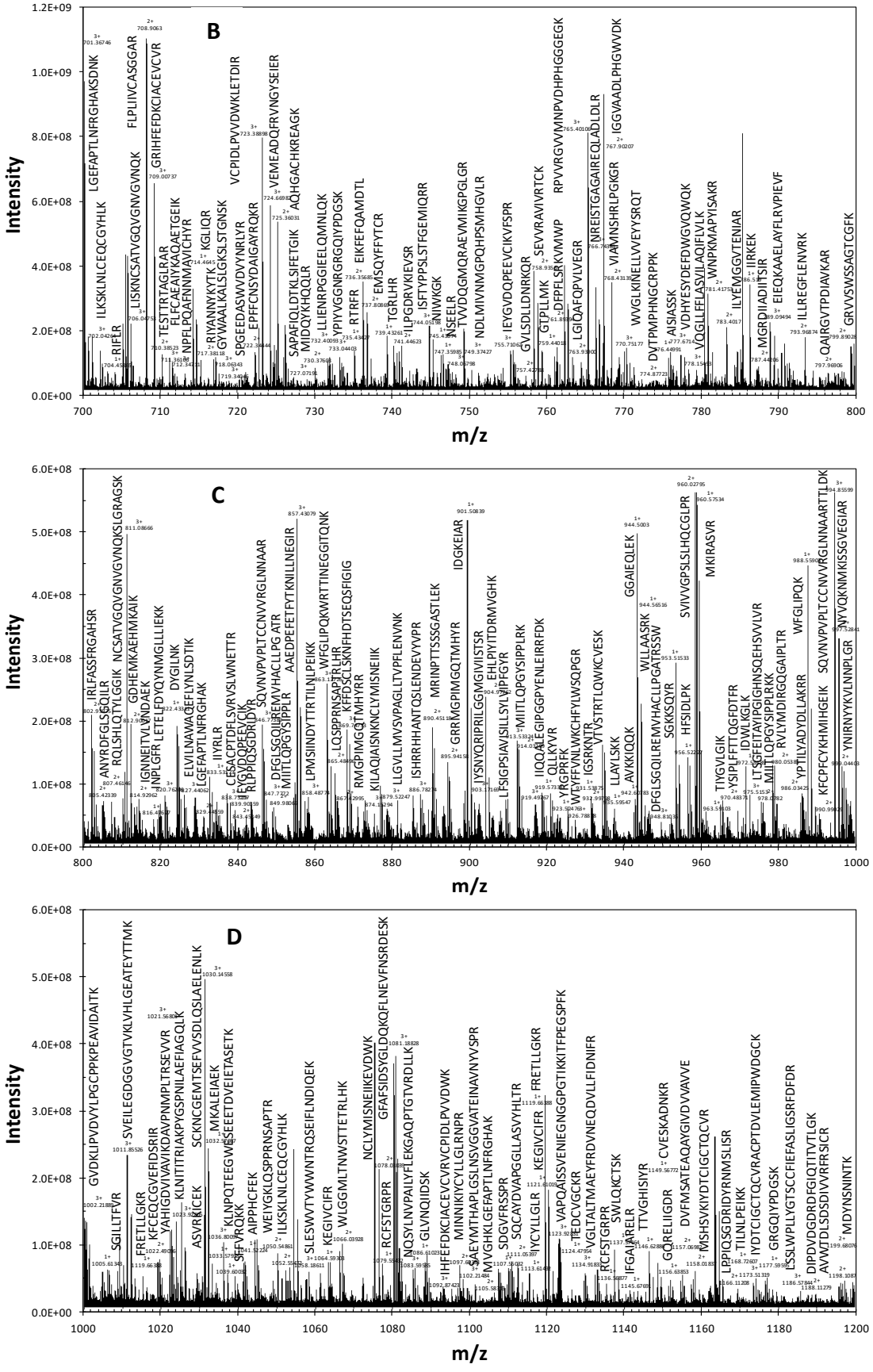


Figure 30. Expanded views of the mass segments 600-700 (panel A), 700-800 (panel B), 800-1000 (panel C) and 1000-1200 m/z (panel D) of the ESI(+)-FTICR mass spectrum of a tryptic mixture (acetonitrile/water (70:30, v/v) + 0.2% FA) of a fennel extract. Sequences matching the in silico peptide digests are shown.

Finally, in order to confirm the protein identity and corroborate the method reliability, the matched experimental mass datasets with the theoretically digested NCBI fennel proteins, obtained by the Matlab algorithm, were submitted to PMF against the full NCBI database. Before database searching, the matched multi-charged m/z values deriving from the Matlab processing have been converted in MH^+ ions (considering the charge state of the corresponding in-silico peptides from Expasy Peptide Mass tool, according to the following equation: MH^+=c*m/z +(1-c)proton mass, where c is the charge state), since this is the query format required for PMF. The submission of the mass data to MASCOT search engine returned fennel proteins as top scoring hits; other potential matching proteins were totally absent in most of the cases; only in 8 protein view results on a total of 61 successful protein identifications, incorrectly assigned proteins were also observed but they are 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[182]. The higher is Δscore, the more accurate will be the identification associated with the first protein hit compared to the second-highest score protein. 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 (that the observed match is a random event) 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. The complete list of all the identified proteins from multiple analyses of the same fennel sample (according to the digestion protocol A and B) is displayed in Table 10. 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 score (84-847), with a sequence coverage higher than 14% and a number of identified peptides ranging from 3 to 109 for individual proteins.

Table 10. List of the identified proteins from fennel (*Foeniculum vulgare*) extract by direct injection ESI qQ-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	10	5997	53
Ribosomal protein L33 (chloroplast)	YP_009235900.1	87	146	7	7687	66
Photosystem I subunit VII (chloroplast)	YP_009235931.1	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	9	9181	77
Ribosomal protein S19 (chloroplast)	YP_009235920.1	70	172	8	10545	92
Ribosomal protein L23 (chloroplast)	YP_009235922.1	82	195	9	10724	93
Ribosomal protein S15 (chloroplast)	YP_009235937.1	72	181	10	10796	90
Ribosomal protein S14 (chloroplast)	YP_009235877.1	80	215	11	11750	100
Ribosomal protein S18 (chloroplast)	YP_009235901.1	82	218	13	11986	101
RNA polymerase beta subunit, partial (chloroplast)	AFK09970.1	31	134	5	12859	117
Ribosomal protein L14 (chloroplast)	YP_009235916.1	85	203	9	13511	122
Ribosomal protein S12 (chloroplast)	YP_009235858.1	55	139	6	13738	123
RNA polymerase beta' subunit, partial (chloroplast)	AFK10055.1	75	200	10	13824	123
Ribosomal protein S11 (chloroplast)	YP_009235912.1	55	147	9	14972	138
Ribosomal protein L16 (chloroplast)	YP_009235917.1	63	194	11	15289	135
ATP synthase CF1 epsilon subunit (chloroplast)	YP_009235885.1	21	86	4	15292	140
Ribosomal protein L20 (chloroplast)	YP_009235902.1	70	200	12	15295	128
Ribosomal protein S8 (chloroplast)	YP_009235915.1	66	172	10	15829	134
Ribosomal protein S7 (chloroplast)	YP_009235925.1	59	239	14	17357	155
Cytochrome b6/f complex subunit IV (chloroplast)	YP_009235910.1	84	90	3	17459	160

NADH-plastoquinone oxidoreductase subunit J (chloroplast)	YP_009235882.1	41	165	8	18605	158	
NADH-plastoquinone oxidoreductase subunit I (chloroplast)	YP_009235934.1	40	113	6	19567	167	
Photosystem I assembly protein Ycf3 (chloroplast)	YP_009235880.1	42	102	5	19600	168	
RNA polymerase C, partial (chloroplast)	ACB88304.1	54	214	12	19693	176	
ATP synthase CF0 subunit I (chloroplast)	YP_009235865.1	59	212	12	20527	181	
Homogentisate geranylgeranyl transferase, partial	ADG26669.1	50	106	5	20569	186	
Photosystem I assembly protein ycf4 (chloroplast)	YP_009235890.1	54	137	7	21194	184	
CLP protease proteolytic subunit (chloroplast)	YP_009235904.1	29	84	4	22188	197	
Ribosomal protein S4 (chloroplast)	YP_009235881.1	62	290	16	23341	201	
Ribosomal protein S3 (chloroplast)	YP_009235918.1	70	269	16	25095	218	
NADH-plastoquinone oxidoreductase subunit K (chloroplast)	YP_009235883.1	53	227	13	25448	225	
Phenylalanine ammonia-lyase, partial	QDG10107.1	43	270	15	25740	232	
Ribosomal protein S2 (chloroplast)	YP_009235868.1	38	149	10	26874	236	
Chloroplast envelope membrane protein (chloroplast)	YP_009235891.1	37	124	6	26989	229	
Maturase K, partial (chloroplast)	ACB88392.1	63	180	11	27760	238	
Ribosomal protein L2 (chloroplast)	YP_009235921.1	47	157	11	30100	274	
Cytochrome f (chloroplast)	YP_009235892.1	48	264	14	35170	320	
Delta12-fatty acid acetylenase, partial	AAO38034.1	20	94	6	37861	324	
RNA polymerase alpha subunit (chloroplast)	YP_009235911.1	50	236	15	38479	336	
Photosystem II protein D2 (chloroplast)	YP_009235874.1	36	140	7	39550	353	
NADH-plastoquinone oxidoreductase subunit 1 (chloroplast)	YP_009235935.1	29	150	8	40317	363	
NADH-plastoquinone oxidoreductase subunit 7 (chloroplast)	YP_009235936.1	73	362	20	45734	395	
Photosystem II CP43 chlorophyll apoprotein (chloroplast)	YP_009235875.1	26	158	9	51879	473	
Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit, partial (chloroplast)	YP_009235887.1	40	224	17	52596	475	
ATP synthase CF1 beta subunit (chloroplast)	YP_009235886.1	65	333	20	53589	498	

Acetyl-CoA carboxylase carboxyltransferase beta subunit (chloroplast)	YP_009235888.1	43	163	11	55333	489
ATP synthase CF1 alpha subunit (chloroplast)	YP_009235864.1	49	269	15	55426	507
Photosystem II CP47 chlorophyll apoprotein (chloroplast)	YP_009235905.1	32	228	14	56004	508
Maturase K (chloroplast)	YP_009235860.1	52	328	26	61157	518
Hypothetical chloroplast RF1 (chloroplast)	YP_009235926.1	49	384	27	74606	635
RNA polymerase beta (chloroplast)	YP_009235870.1	64	666	51	78281	678
Photosystem I P700 apoprotein A2 (chloroplast)	YP_009235878.1	24	171	11	82345	734
Photosystem I P700 apoprotein A1 (chloroplast)	YP_009235879.1	44	270	17	83158	750
NADH-plastoquinone oxidoreductase subunit 5 (chloroplast)	YP_009235927.1	14	107	9	85465	750
RNA polymerase beta subunit (chloroplast)	YP_009235871.1	53	516	47	120866	1070
Nuclear matrix constituent protein 1-like	BAF64423.1	56	737	85	128731	1119
RNA polymerase beta' subunit (chloroplast)	YP_009235869.1	53	693	72	157541	1386
Hypothetical chloroplast RF1 (chloroplast)	YP_009235938.1	56	847	109	217228	1817
Hypothetical chloroplast RF2 (chloroplast)	YP_009235923.1	48	717	98	247783	2119

As an example, the identification of the Ribosomal protein L33 (chloroplast) is reported in Figure 31. Amino acidic sequences corresponded to the identified peptides are highlighted in bold red.

MATRIX SCIENCE **MASCOT Search Results**

Protein View: YP_009235900.1

ribosomal protein L33 (chloroplast) [Foeniculum vulgare]

Database: NCBIprot
Score: 146
Expect: 2.4e-08
Monoisotopic mass (M_r): 7687
Calculated pI: 9.74
Taxonomy: Foeniculum vulgare

This protein sequence matches the following other entries:

- YP_009235985.1 from Anethum graveolens
- ABU85204.1 from Anethum graveolens
- AMD83936.1 from Foeniculum vulgare
- AMD84021.1 from Anethum graveolens

Sequence similarity is available as an NCBI BLAST search of YP_009235900.1 against nr.

Search parameters

Enzyme: Trypsin: cuts C-term side of KR unless next residue is P.
Mass values searched: 8
Mass values matched: 8

Protein sequence coverage: 87%

Matched peptides shown in ***bold red***.

```
  1 MAKGKDVRIT VILECTGCVR NDVNKVSTGI SRYITEKNRH NTPNRLELRK
 51 FCPFCYKHMI HGEIKK
```

Unformatted sequence string: 66 residues (for pasting into other applications).

Sort by (•) residue number ( ) increasing mass ( ) decreasing mass
Show (•) matched peptides only ( ) predicted peptides also

Start – End	Observed	Mr(expt)	Mr(calc)	ppm	M	Peptide
1 – 5	534.3083	533.3010	533.2995	2.78	1	-.MAKGK.D
4 – 8	574.3285	573.3213	573.3235	-3.84	1	K.GKDVR.I
9 – 25	1876.9671	1875.9598	1875.9547	2.74	1	R.ITVILECTGCVRNDVNK.V
33 – 39	923.4956	922.4883	922.4872	1.18	1	R.YITEKNR.H
40 – 45	738.3613	737.3540	737.3569	-3.92	0	R.HNTPNR.L
40 – 49	1249.6732	1248.6659	1248.6687	-2.22	1	R.HNTPNRLELR.K
40 – 49	1249.6763	1248.6690	1248.6687	0.24	1	R.HNTPNRLELR.K
51 – 66	1980.9727	1979.9654	1979.9572	4.14	2	K.FCPFCYKHMIHGEIKK.-

Figure 31. Mascot search results view of Ribosomal protein L33 (chloroplast) identification from matched experimental mass dataset of fennel by Peptide Mass Fingerprint searching. NCBI database; Viridiplantae category; No variable and fixed modifications; enzyme trypsin; 2 allowed missed cleavages; peptide mass tolerance of 5 ppm.

As far as the allergenic profile is concerned, nine homologues proteins (see Table 11) among the total ten investigated allergens (Table 9) were well identified in the fennel extract, thus 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 similarities to the recognized allergenic proteins from different sources.

Table 11. List of allergenic no-fennel specific proteins identified in fennel extract by direct injection ESI qQ-FT-ICR-MS analysis and PMF Mascot database searching.

Protein Name	Organism	Protein Code	Sequence Coverage %	MASCOT Score	No. Matched Peptides	MW (Da)
Allergen Api g 5	*Apium graveolans*	P81943	51	88	4	9441
Major allergen Api g 1, isoallergen 1	*Apium graveolans*	P49372	39	136	6	16321
Major allergen Api g 1, isoallergen 2	*Apium graveolans*	P92918	49	133	6	17090
Cofactor-independent phosphoglyceromutase	*Apium graveolans*	Q9SDL3	43	291	19	61125
Pathogenesis-related protein 1	*Petroselinum crispum*	Q40795	35	103	5	16356
Non specific lipid transfer protein	*Dacus carota*	P27631	49	84	4	12504
Major pollen allergen Bet v 1-A	*Betula pendula*	P15494	69	158	7	17571
Major pollen allergen Bet v 1-B	*Betula pendula*	P45431	62	163	8	17537
Major pollen allergen Art v 1	*Artemisia vulgaris*	Q84ZX5	41	126	7	13404

In conclusion, 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, calculation of tryptic digestions and matching the calculated post-digestion peptides with the experimental FT-ICR mass spectrum of tryptic fennel protein digestion mixture as well as the subsequent Mascot database searching in peptide mass fingerprint PMF mode represents the most informative approach for a rapid and accurate protein characterization in fennel protein extracts.

3.3.4. Immunoblotting analysis on 40 sera of fennel allergy patients

The *Foeniculum vulgare* extract was also used in immunoblotting analysis to detect immunoreactive putative bands, maybe responsible for fennel allergy. Therefore, the extract was at first analyzed by SDS-PAGE, then by antibody-based assays for Immunoglobulin E detection in order to clarify the effective allergenic molecules. This experimental activity was performed by the research team of Professor Macchia at the Policlinic of Bari (Italy) and is included in the Ph.D. thesis of dr. Mariangela Di Giacomo (Thesis title: Fennel (*Foeniculum vulgare*): a novel food allergen of the Mediterranean Diet; Ph.D. course in Qualità degli Alimenti e Nutrizione Umana, XXVIII cycle).

Among a population of 189 adult patients with food allergy diagnosis, *Foeniculum vulgare* allergy was diagnosed in 57 patients (30%), who reported symptoms clearly associated with fennel consumption and exhibited positive Skin Prick Test with fennel extracts. Immunoblotting analysis of the putative fennel allergens was performed with sera from 40 fennel allergy subjects. Only patients with IgE values > 0.1 kU/l were considered. To detect immune-reactive bands, serum obtained from each patient (the primary antibody) and an anti-human IgE peroxidase-conjugated polyclonal antibody (as the secondary antibody) were used.

The IgE-binding proteins were detected by enhanced chemiluminescence (ECL) and, upon recording on RX films, the molecular weight of the immunoreactive bands was determined by comparison with markers at known molecular weight, loaded on the same gel. The immunoblotting analysis revealed quite different allergenic protein profiles, depending on the patient's serum used as the primary antibody. Putative allergenic proteins involved in *F. vulgare* allergy were detected, as reported in Figure 32: a band with an apparent molecular mass of ~ 45-50 kDa was detected by sera from 10 patients and another immunoreactive band with an apparent molecular weight of ~ 27-34 kDa was detected by sera from 8 fennel allergy patients.

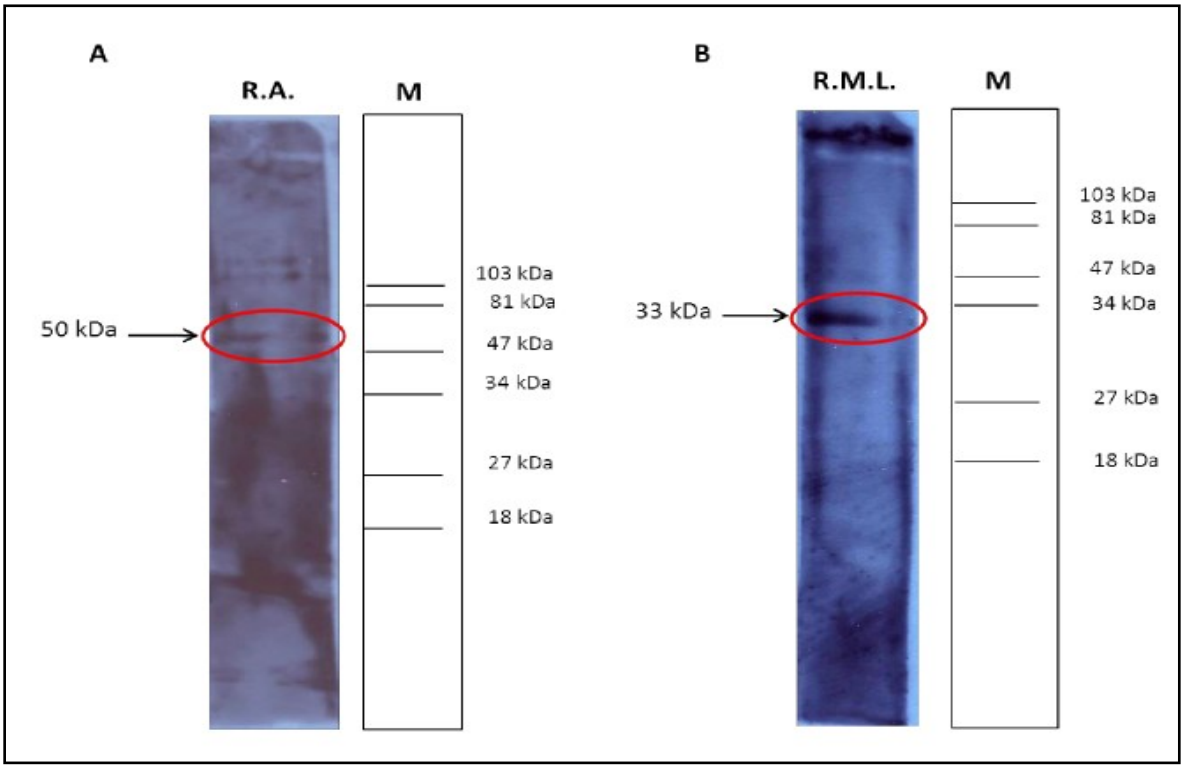


Figure 32. A) The 50 kDa immunoreactive band detected by serum from the patient R.A.; B) the band with the apparent molecular weight of 33 kDa observed in the patient R.M.L. In both cases, an anti-human IgE peroxidase-conjugated polyclonal antibody (Sigma, Milan, Italy) was used as the secondary antibody. These bands were visualized by ECL (exposure time 10' and 15', respectively).

These results (considering the gel-method variability) are in accordance with literature data regarding the existence of a 60 kDa fennel allergen [185] corresponding to the protein cofactor-independent phosphoglyceromutase, whose presence in our fennel extract was confirmed by FT-ICR analysis (see Table 11). As far as the smaller allergenic proteins in the mugwort-fennel-spice-allergy-syndrome [66,68,70,187] are concerned, no concluding remarks can be deduced by antibody-based assays (Figure 32), since their molecular weights are below the lower limit of 18 kDa in the gel. On the other side, eight homologues allergenic proteins with molecular weight below 18 kDa were found in the raw fennel extract by MS analysis. Then, 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 searching. As can be noted from Figure

33, 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*), supporting the hypothesis of a common molecular basis at the origin of the cross-sensitization of patients with spice-pollen allergy.

A)

1	XXXXXXXXXX	XXXXXXXX**LF**	**KAFILDGDNL**	**FPKVAPQAIS**	**SVENIEGNGG**
51	**PGTIK**X**ISFP**	**EGFPFKYVK**X	XXXXXXXXXX	X**YNYSVIEGG**	**PIGDTLEKIS**
101	**NEIKIVATPD**	**GGSILK**XXXX	**YHTKGDHEVK**	**AEQVKASKEM**	**GETLLR**XXXX
151	XXXXXXXXXX				

B)

1	MGVFNYETET	TSVIPAAR**LF**	**KAFILDGDNL**	**FPKVAPQAIS**	**SVENIEGNGG**
51	**PGTIK**K**ISFP**	**EGFPFKYVK**D	RVDEVDHTNF	K**YNYSVIEGG**	**PIGDTLEKIS**
101	**NEIKIVATPD**	**GGSILK**ISNK	**YHTKGDHEVK**	**AEQVKASKEM**	**GETLLR**AVES
151	YLLAHSDAYN				

Figure 33. Primary structure of Major pollen allergen Bet v 1-A of: A) *Foeniculum vulgare* and B) *Betula pendula*.

This further investigation could significantly improve the scientific impact of the present project, providing additional information (at the moment not clarified in the literature) about the presence of common allergens among different organisms and the determination of their homology grade with the analogues proteins in the other plants.

4. CONCLUSION

The control activities in food production rely on the availability of analytical methods capable of detecting traces of allergenic proteins. Most routine food allergen analysis is undertaken by Polymerase Chain Reaction (PCR) and enzyme-linked immunosorbent assay (ELISA) enabling detection and (semi-) quantification. However, cross-reactivity problems could lead to false-positive results due to a possible interaction of the antibodies with the food matrix. Although many efforts have been made during the last few years leading to the identification of new allergenic sources and individual molecules, this knowledge is still incomplete and does not allow the definition of a comprehensive pattern. The fragmentary information on allergenic molecules contained in allergenic sources also determine socioeconomic effects in the clinical field, affecting costs of healthcare assistance and public advertisements, as well as in private sector, with consequences in food manufacturing, processing and marketing, cost management by pharmaceuticals companies, waste reduction in purchasing of food products and drugs by consumers.

The hallmark characteristics of the analytical methods for allergen detection are the complexity of food matrices and the unknown identity of the key analytes. Efforts to reduce the intrinsic limits and drawbacks of proteomics are challenging because the protein separation from complex media is technically demanding and time-consuming. Indeed, the dynamic range of proteins in a biological system extends from one copy to more than a million copies per cell; furthermore, the heterogeneity of the proteins in terms of molecular range and isoelectric point resulting in different physical and chemical behaviors, requests an additional effort during sample handling and protein determination. It is not by chance that at the moment the food allergen lists include the most abundant proteins in the natural sources or showing high structural stability to proteolysis and high temperatures.

In recent years, mass spectrometry-based methods have been considered a promising analytical strategy for food allergens monitoring thanks to their high specificity, sensitivity and accuracy. Nevertheless, the characterization of proteins from food samples always needs a concerted application of several technologies, where MS measurements represent only the final step. Accurate analytical methods for protein isolation are essential in the field of food safety for the recovery of allergens from food samples, before protein identification by MS.

The present Ph.D. project aimed at the development of analytical methods for the characterization of proteins as putative allergens in animal and vegetable food samples by Two-Dimensional Liquid Chromatography (2D-LC) and Mass Spectrometry.

During the first year of research activity, an automated food protein isolation approach by Two-Dimensional Liquid Chromatography with an active modulation interface was developed for the protein isolation on a preparative scale. An innovative trapping interface made by a reversed phase guard column installed in a switching valve, placed between two separation columns, was proposed. Therefore, the proteins coming from the first column were trapped and focused before injection in the second column, reducing the effects of dead volumes and band broadening. Then, a stop-and-go/active modulation SEC-RP LC approach was used for an on-line isolation of intact proteins from egg samples. A size-exclusion column was used for the first dimension separation and combined to a reversed phase *widepore* column for the second dimension separation. Protein peaks coming from the second separation column were collected through a fraction collector connected to the UV cell. The optimized separation workflow, as a good compromise between the comprehensive and multiple heart-cutting modes, takes on the challenging task of analyzing complex food matrices rich in proteins greatly different in concentration, mass and polarity, ensuring high resolving power and automation grade.

During the second Ph.D. year, the use of two multi-port switching valves was proposed as an enhanced approach for protein isolation by actively modulated two-dimensional SEC-RP liquid chromatography. A C4 trapping column installed in the active modulation interface allows reducing the re-mixing effects and the problems related to the incompatibility of eluent composition and pH between the two separation columns. An automated isolation of individual proteins as intact species from soy flour samples was performed by the optimized gel-free method in order to collect a sufficient protein amount for further molecular characterization by MS or allergy testing. Therefore, the innovative method, proposed for the protein separation, allows performing an automated isolation of individual proteins for their subsequent MS characterization (by both top-down or bottom-up proteomics) or for further bio-analytical investigations. These procedures represent the starting point for allergen characterization in food products, allowing the upgrade of the pattern of allergenic molecules/sources and improving official controls.

Finally, the research activity of the last Ph.D. year was focused on the protein characterization in extracts of fennel, a novel food allergen belonging to the Mediterranean

diet. An untargeted shotgun approach by FT-ICR-MS was developed in order to characterize the whole fennel proteome. Few microliters of extract have been digested with trypsin and analyzed by MS, without recurring to any fractionation or purification process. A direct electrospray FT-ICR-MS analysis of the fennel extract sample allowed to obtain unambiguous peptide sequence assignments in a few seconds. 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. A data pre-processing was proposed to compare the experimental mass peak list with theoretical data deriving from the silico enzymatic digestion of all the known fennel proteins. Then, the protein identity was confirmed by database searching in PMF mode of the matched experimental mass peak lists, taking advantage of the very high mass accuracy provided by FT-ICR-MS technique. The molecular characterization by MS combined to the immunoblotting IgE assays allowed to confirm the presence of nine allergenic proteins in the fennel sample.

On the basis of these considerations, the research activity of the present project, through the development of proper analytical tools for protein separation and characterization to be used for the identification of new allergens, should actively contribute to: i) elucidate comprehensive allergen profiles and upgrade the pattern of allergenic molecules/sources; ii) evaluate the presence at trace levels in food matrices, iii) check/confirm the declaration on the commercial label and iv) study allergen modifications occurring during and after food processing. Therefore, the expected results of the present project are supposed to provide improvements in the field of public health, with a special impact on the quality of life of sensitized/allergic individuals.

5. REFERENCES

[1] P.S.C. Leung, S.-A. Shu, C. Chang, The Changing Geoepidemiology of Food Allergies, Clin. Rev. Allergy Immunol. 46 (2014) 169–179. https://doi.org/10.1007/s12016-014-8411-5.

[2] D.D. Metcalfe, H.A. Sampson, R.A. Simon, G. Lack, eds., Food Allergy: Adverse Reactions to Foods and Food Additives, John Wiley & Sons Ltd, Chichester, UK, 2013. http://doi.wiley.com/10.1002/9781118744185.

[3] F. Ferreira, T. Hawranek, P. Gruber, N. Wopfner, A. Mari, Allergic cross-reactivity: from gene to the clinic, Allergy. 59 (2004) 243–267. https://doi.org/10.1046/j.1398-9995.2003.00407.x.

[4] A.B. Fishbein, R.L. Fuleihan, The hygiene hypothesis revisited: does exposure to infectious agents protect us from allergy?, Curr. Opin. Pediatr. 24 (2012) 98–102. https://doi.org/10.1097/MOP.0b013e32834ee57c.

[5] A. Kemp, B. Björkstén, Immune deviation and the hygiene hypothesis: A review of the epidemiological evidence, Pediatr. Allergy Immunol. 14 (2003) 74–80. https://doi.org/10.1034/j.1399-3038.2003.00017.x.

[6] M. Rottem, Y. Shoenfeld, Asthma as a Paradigm for Autoimmune Disease, Int. Arch. Allergy Immunol. 132 (2003) 210–214. https://doi.org/10.1159/000074301.

[7] M. Rottem, M.E. Gershwin, Y. Shoenfeld, Allergic Disease and Autoimmune Effectors Pathways, Dev. Immunol. 9 (2002) 161–167. https://doi.org/10.1080/1044667031000137638.

[8] L. Zuidmeer-Jongejan, M. Fernandez-Rivas, L.K. Poulsen, A. Neubauer, J. Asturias, L. Blom, J. Boye, C. Bindslev-Jensen, M. Clausen, R. Ferrara, P. Garosi, H. Huber, B.M. Jensen, S. Koppelman, M.L. Kowalski, A. Lewandowska-Polak, B. Linhart, B. Maillere, A. Mari, A. Martinez, C.E. Mills, C. Nicoletti, D.-J. Opstelten, N.G. Papadopoulos, A. Portoles, N. Rigby, E. Scala, H.J. Schnoor, S.T. Sigurdardottir, G. Stavroulakis, F. Stolz, I. Swoboda, R. Valenta, R. van den Hout, S.A. Versteeg, M. Witten, R. van Ree, FAST: towards safe and effective subcutaneous immunotherapy of persistent life-threatening food allergies, Clin. Transl. Allergy. 2 (2012). https://doi.org/10.1186/2045-7022-2-5.

[9] K.J. Allen, B.C. Remington, J.L. Baumert, R.W.R. Crevel, G.F. Houben, S. Brooke-Taylor, A.G. Kruizinga, S.L. Taylor, Allergen reference doses for precautionary labeling (VITAL 2.0): Clinical implications, J. Allergy Clin. Immunol. 133 (2014) 156–164. https://doi.org/10.1016/j.jaci.2013.06.042.

[10] K.J. Allen, P.J. Turner, R. Pawankar, S. Taylor, S. Sicherer, G. Lack, N. Rosario, M. Ebisawa, G. Wong, E.N.C. Mills, K. Beyer, A. Fiocchi, H.A. Sampson, Precautionary labelling of foods for allergen content: are we ready for a global framework?, World Allergy Organ. J. 7 (2014) 10. https://doi.org/10.1186/1939-4551-7-10.

[11] M. Pele, M. Brohée, E. Anklam, A.J.V. Hengel, Peanut and hazelnut traces in cookies and chocolates: Relationship between analytical results and declaration of food allergens on product labels, Food Addit. Contam. 24 (2007) 1334–1344. https://doi.org/10.1080/02652030701458113.

[12] S.L. Taylor, S.L. Hefle, K. Farnum, S.W. Rizk, J. Yeung, M.E. Barnett, F. Busta, F.R. Shank, R. Newsome, S. Davis, C.M. Bryant, Analysis and Evaluation of Food Manufacturing Practices Used to Address Allergen Concerns, Compr. Rev. Food Sci. Food Saf. 5 (2006) 138–157. https://doi.org/10.1111/j.1541-4337.2006.00012.x.

[13] J. Lidholm, B.K. Ballmer-Weber, A. Mari, S. Vieths, Component-resolved diagnostics in food allergy:, Curr. Opin. Allergy Clin. Immunol. 6 (2006) 234–240. https://doi.org/10.1097/01.all.0000225166.90768.d6.
[14] A. Mari, C. Rasi, P. Palazzo, E. Scala, Allergen databases: Current status and perspectives, Curr. Allergy Asthma Rep. 9 (2009) 376–383. https://doi.org/10.1007/s11882-009-0055-9.
[15] K. Török, V. Horváth, Á. Horváth, L. Hajas, Z. Bugyi, S. Tömösközi, Investigation of incurred single- and multi-component model food matrices for determination of food proteins triggering allergy and coeliac disease, Eur. Food Res. Technol. 239 (2014) 923–932. https://doi.org/10.1007/s00217-014-2289-6.
[16] C. Radauer, M. Bublin, S. Wagner, A. Mari, H. Breiteneder, Allergens are distributed into few protein families and possess a restricted number of biochemical functions, J. Allergy Clin. Immunol. 121 (2008) 847–852.e7. https://doi.org/10.1016/j.jaci.2008.01.025.
[17] K. Hoffmann-Sommergruber, E.N.C. Mills, Food allergen protein families and their structural characteristics and application in component-resolved diagnosis: New data from the EuroPrevall project, Anal. Bioanal. Chem. 395 (2009) 25–35. https://doi.org/10.1007/s00216-009-2953-z.
[18] G. Martos, R. López-Fandiño, E. Molina, Immunoreactivity of hen egg allergens: Influence on in vitro gastrointestinal digestion of the presence of other egg white proteins and of egg yolk, Food Chem. 136 (2013) 775–781. https://doi.org/10.1016/j.foodchem.2012.07.106.
[19] Y. Mine, P. Rupa, Immunological and biochemical properties of egg allergens, Worlds Poult. Sci. J. 60 (2004) 321–330. https://doi.org/10.1079/WPS200420.
[20] B. Jacobsen, K. Hoffmann-Sommergruber, T.T. Have, N. Foss, P. Briza, C. Oberhuber, C. Radauer, S. Alessandri, A.C. Knulst, M. Fernandez-Rivas, V. Barkholt, The panel of egg allergens, Gal d 1-Gal d 5: Their improved purification and characterization, Mol. Nutr. Food Res. 52 (2008) S176–S185. https://doi.org/10.1002/mnfr.200700414.
[21] A. Maeno, H. Matsuo, K. Akasaka, Tyrosine/tyrosinate fluorescence at 700 MPa: A pressure unfolding study of chicken ovomucoid at pH 12, Biophys. Chem. 183 (2013) 57–63. https://doi.org/10.1016/j.bpc.2013.07.008.
[22] S. Quirce, F. Marañón, A. Umpiérrez, M. De Las Heras, E. Fernández-Caldas, J. Sastre, Chicken serum albumin (Gal d 5*) is a partially heat-labile inhalant and food allergen implicated in the bird-egg syndrome, Allergy Eur. J. Allergy Clin. Immunol. 56 (2001) 754–762. https://doi.org/10.1034/j.1398-9995.2001.056008754.x.
[23] A. Amo, R. Rodríguez-Pérez, J. Blanco, J. Villota, S. Juste, I. Moneo, M.L. Caballero, Gal d 6 is the second allergen characterized from egg yolk, J. Agric. Food Chem. 58 (2010) 7453–7457. https://doi.org/10.1021/jf101403h.
[24] L.M. Bartnikas, W. Phipatanakul, Turning up the heat on skin testing for baked egg allergy, Clin. Exp. Allergy. 43 (2013) 1095–1096. https://doi.org/10.1111/cea.12183.
[25] A. Urisu, H. Ando, Y. Morita, E. Wada, T. Yasaki, K. Yamada, K. Komada, S. Torii, M. Goto, T. Wakamatsu, Allergenic activity of heated and ovomucoid-depleted egg white, J. Allergy Clin. Immunol. 100 (1997) 171–176. https://doi.org/10.1016/S0091-6749(97)70220-3.
[26] C. Escudero, S. Sánchez-García, P. Rodríguez del Río, C. Pastor-Vargas, C. García-Fernández, I. Pérez-Rangel, A. Ramírez-Jiménez, M.D. Ibáñez, Dehydrated egg white: An allergen source for improving efficacy and safety in the diagnosis and

treatment for egg allergy, Pediatr. Allergy Immunol. 24 (2013) 263–269. https://doi.org/10.1111/pai.12052.
[27] J. Golias, M. Schwarzer, M. Wallner, M. Kverka, H. Kozakova, D. Srutkova, K. Klimesova, P. Sotkovsky, L. Palova-Jelinkova, F. Ferreira, L. Tuckova, Heat-induced structural changes affect OVA-antigen processing and reduce allergic response in mouse model of food allergy, PLoS ONE. 7 (2012). https://doi.org/10.1371/journal.pone.0037156.
[28] G. Martos, I. Lopez-Exposito, R. Bencharitiwong, M.C. Berin, A. Nowak-Węgrzyn, Mechanisms underlying differential food allergy response to heated egg, J. Allergy Clin. Immunol. 127 (2011) 990–997.e2. https://doi.org/10.1016/j.jaci.2011.01.057.
[29] R. Jiménez-Saiz, S. Benedé, E. Molina, I. López-Expósito, Effect of Processing Technologies on the Allergenicity of Food Products, Crit. Rev. Food Sci. Nutr. 55 (2015) 1902–1917. https://doi.org/10.1080/10408398.2012.736435.
[30] X.J. Ma, H.B. Chen, J.Y. Gao, C.Q. Hu, X. Li, Conformation affects the potential allergenicity of ovalbumin after heating and glycation, Food Addit. Contam. - Part Chem. Anal. Control Expo. Risk Assess. 30 (2013) 1684–1692. https://doi.org/10.1080/19440049.2013.822105.
[31] M. Shin, Y. Han, K. Ahn, The influence of the time and temperature of heat treatment on the allergenicity of egg white proteins, Allergy Asthma Immunol. Res. 5 (2013) 96–101. https://doi.org/10.4168/aair.2013.5.2.96.
[32] Y. Mine, J.W. Zhang, Comparative studies on antigenicity and allergenicity of native and denatured egg white proteins, J. Agric. Food Chem. 50 (2002) 2679–2683. https://doi.org/10.1021/jf0112264.
[33] S. Hildebrandt, H.D. Kratzin, R. Schaller, R. Fritsché, H. Steinhart, A. Paschke, In vitro determination of the allergenic potential of technologically altered hen's egg, J. Agric. Food Chem. 56 (2008) 1727–1733. https://doi.org/10.1021/jf0725981.
[34] L. Manzocco, M.C. Nicoli, Macromolecular crowding affects protein photosensitivity: The case of egg white immunoreactivity, Food Chem. 132 (2012) 982–988. https://doi.org/10.1016/j.foodchem.2011.11.083.
[35] L. Manzocco, A. Panozzo, M.C. Nicoli, Effect of ultraviolet processing on selected properties of egg white, Food Chem. 135 (2012) 522–527. https://doi.org/10.1016/j.foodchem.2012.05.028.
[36] J.-W. Lee, J.-H. Seo, J.-H. Kim, S.-Y. Lee, M.-W. Byun, Comparison of the changes of the antigenicities of a hen's egg albumin by a gamma and an electron beam irradiation, Radiat. Phys. Chem. 76 (2007) 879–885. https://doi.org/10.1016/j.radphyschem.2006.06.010.
[37] T. Holzhauser, O. Wackermann, B.K. Ballmer-Weber, C. Bindslev-Jensen, J. Scibilia, L. Perono-Garoffo, S. Utsumi, L.K. Poulsen, S. Vieths, Soybean (Glycine max) allergy in Europe: Gly m 5 (β-conglycinin) and Gly m 6 (glycinin) are potential diagnostic markers for severe allergic reactions to soy, J. Allergy Clin. Immunol. 123 (2009) 452–458.e4. https://doi.org/10.1016/j.jaci.2008.09.034.
[38] M. Ebisawa, P. Brostedt, S. Sjölander, S. Sato, M.P. Borres, K. Ito, Gly m 2S albumin is a major allergen with a high diagnostic value in soybean-allergic children, J. Allergy Clin. Immunol. 132 (2013) 976–978.e5. https://doi.org/10.1016/j.jaci.2013.04.028.
[39] J. Kleine-Tebbe, A. Wangorsch, L. Vogel, D.N. Crowell, U.-F. Haustein, S. Vieths, Severe oral allergy syndrome and anaphylactic reactions caused by a Bet v 1-related PR-10 protein in soybean, SAM22, J. Allergy Clin. Immunol. 110 (2002) 797–804. https://doi.org/10.1067/mai.2002.128946.

[40] R. González, L. Zapatero, F. Caravaca, J. Carreira, Identification of soybean proteins responsible for respiratory allergies, Int. Arch. Allergy Immunol. 95 (1991) 53–57. https://doi.org/10.1159/000235454.
[41] G.S. Ladics, G.J. Budziszewski, R.A. Herman, C. Herouet-Guicheney, S. Joshi, E.A. Lipscomb, S. McClain, J.M. Ward, Measurement of endogenous allergens in genetically modified soybeans - Short communication, Regul. Toxicol. Pharmacol. 70 (2014) 75–79. https://doi.org/10.1016/j.yrtph.2014.06.006.
[42] T. Ogawa, M. Samoto, K. Takahashi, Soybean allergens and hypoallergenic soybean products, J. Nutr. Sci. Vitaminol. (Tokyo). 46 (2000) 271–279. https://doi.org/10.3177/jnsv.46.271.
[43] A. Paschke, K. Zunker, M. Wigotzki, H. Steinhart, Determination of the IgE-binding activity of soy lecithin and refined and non-refined soybean oils, J. Chromatogr. B. Biomed. Sci. App. 756 (2001) 249–254. https://doi.org/10.1016/S0378-4347(01)00085-8.
[44] R.K. Bush, S.L. Taylor, J.A. Nordlee, W.W. Busse, Soybean oil is not allergenic to soybean-sensitive individuals, J. Allergy Clin. Immunol. 76 (1985) 242–245. https://doi.org/10.1016/0091-6749(85)90709-2.
[45] Y.S. Song, C. Martinez-Villaluenga, E.G. De Mejia, Quantification of human IgE immunoreactive soybean proteins in commercial soy ingredients and products, J. Food Sci. 73 (2008) T90–T99. https://doi.org/10.1111/j.1750-3841.2008.00848.x.
[46] A.W. Burks, L.W. Williams, R.M. Helm, W. Thresher, J.R. Brooks, H.A. Sampson, Identification of soy protein allergens in patients with atopic dermatitis and positive soy challenges; Determination of change in allergenicity after heating or enzyme digestion, Adv. Exp. Med. Biol. 289 (1991) 295–307.
[47] M. Amigo-Benavent, J.M. Silván, F.J. Moreno, M. Villamiel, M.D. Del Castillo, Protein quality, antigenicity, and antioxidant activity of soy-based foodstuffs, J. Agric. Food Chem. 56 (2008) 6498–6505. https://doi.org/10.1021/jf800697n.
[48] A. Gomaa, J.I. Boye, Impact of thermal processing time and cookie size on the detection of casein, egg, gluten and soy allergens in food, Food Res. Int. 52 (2013) 483–489. https://doi.org/10.1016/j.foodres.2013.01.019.
[49] A. Ohishi, K. Watanabe, M. Urushibata, K. Utsuno, K. Ikuta, K. Sugimoto, H. Harada, Detection of soybean antigenicity and reduction by twin-screw extrusion, J. Am. Oil Chem. Soc. 71 (1994) 1391–1396. https://doi.org/10.1007/BF02541361.
[50] R. Yamanishi, H. Tsuji, N. Bando, Y. Yamada, Y. Nadaoka, T. Huang, K. Nishikawa, S. Emoto, T. Ogawa, Reduction of the allergenicity of soybean by treatment with proteases, J. Nutr. Sci. Vitaminol. (Tokyo). 42 (1996) 581–587. https://doi.org/10.3177/jnsv.42.581.
[51] K. Tsumura, W. Kugimiya, N. Bando, M. Hiemori, T. Ogawa, Preparation of Hypoallergenic Soybean Protein with Processing Functionality by Selective Enzymatic Hydrolysis, Food Sci. Technol. Res. 5 (1999) 171–175. https://doi.org/10.3136/fstr.5.171.
[52] H.-W. Lee, E.-H. Keum, S.-J. Lee, D.-E. Sung, D.-H. Chung, S.-I. Lee, S. Oh, Allergenicity of proteolytic hydrolysates of the soybean 11S globulin, J. Food Sci. 72 (2007) C168–C172. https://doi.org/10.1111/j.1750-3841.2007.00307.x.
[53] K. Tsumura, Improvement of the physicochemical properties of soybean proteins by enzymatic hydrolysis, Food Sci. Technol. Res. 15 (2009) 381–388. https://doi.org/10.3136/fstr.15.381.
[54] Y. Zhao, G.X. Qin, Z.W. Sun, B. Zhang, T. Wang, Stability and immunoreactivity of glycinin and β-conglycinin to hydrolysis in vitro, Food Agric. Immunol. 21 (2010) 253–263. https://doi.org/10.1080/09540101003758954.

[55] E.L. Van Boxtel, L.A.M. Van Den Broek, S.J. Koppelman, H. Gruppen, Legumin allergens from peanuts and soybeans: Effects of denaturation and aggregation on allergenicity, Mol. Nutr. Food Res. 52 (2008) 674–682. https://doi.org/10.1002/mnfr.200700299.
[56] J. Frias, Y.S. Song, C. Martínez-Villaluenga, E. Gonzalez De Mejia, C. Vidal-Valverde, Fermented soyabean products as hypoallergenic food, Proc. Nutr. Soc. 67 (2008). https://doi.org/10.1017/S0029665108006484.
[57] K. Arita, E.E. Babiker, H. Azakami, A. Kato, Effect of chemical and genetic attachment of polysaccharides to proteins on the production of IgG and IgE, J. Agric. Food Chem. 49 (2001) 2030–2036. https://doi.org/10.1021/jf001120t.
[58] E.F.E. Babiker, A. Hiroyuki, N. Matsudomi, H. Iwata, T. Ogawa, N. Bando, A. Kato, Effect of Polysaccharide Conjugation or Transglutaminase Treatment on the Allergenicity and Functional Properties of Soy Protein, J. Agric. Food Chem. 46 (1998) 866–871. https://doi.org/10.1021/jf9705072.
[59] L. L'Hocine, J.I. Boye, S. Jouve, Ionic strength and pH-induced changes in the immunoreactivity of purified soybean glycinin and its relation to protein molecular structure, J. Agric. Food Chem. 55 (2007) 5819–5826. https://doi.org/10.1021/jf070281v.
[60] H. Li, K. Zhu, H. Zhou, W. Peng, Effects of high hydrostatic pressure treatment on allergenicity and structural properties of soybean protein isolate for infant formula, Food Chem. 132 (2012) 808–814. https://doi.org/10.1016/j.foodchem.2011.11.040.
[61] C. Cuadrado, B. Cabanillas, M.M. Pedrosa, M. Muzquiz, J. Haddad, K. Allaf, J. Rodriguez, J.F. Crespo, C. Burbano, Effect of instant controlled pressure drop on ige antibody reactivity to peanut, lentil, chickpea and soybean proteins, Int. Arch. Allergy Immunol. 156 (2011) 397–404. https://doi.org/10.1159/000324443.
[62] K. Takács, E. Guillamon, M.M. Pedrosa, C. Cuadrado, C. Burbano, M. Muzquiz, J. Haddad, K. Allaf, A. Maczó, M. Polgár, É. Gelencsér, Study of the effect of instant controlled pressure drop (DIC) treatment on IgE-reactive legume-protein patterns by electrophoresis and immunoblot, Food Agric. Immunol. 25 (2014) 173–185. https://doi.org/10.1080/09540105.2012.759539.
[63] S.B. Badgujar, V.V. Patel, A.H. Bandivdekar, Foeniculum vulgare Mill: A review of its botany, phytochemistry, pharmacology, contemporary application, and toxicology, BioMed Res. Int. 2014 (2014). https://doi.org/10.1155/2014/842674.
[64] R. Asero, Relevance of pollen-specific IgE levels to the development of Apiaceae hypersensitivity in patients with birch pollen allergy, Allergy Eur. J. Allergy Clin. Immunol. 52 (1997) 560–564. https://doi.org/10.1111/j.1398-9995.1997.tb02600.x.
[65] R. Asero, Fennel, cucumber, and melon allergy successfully treated with pollen-specific injection immunotherapy, Ann. Allergy. Asthma. Immunol. 84 (2000) 460–462. https://doi.org/10.1016/S1081-1206(10)62281-8.
[66] E. Jensen-Jarolim, A. Leitner, R. Hirschwehr, D. Kraft, B. Wüthrich, O. Scheiner, J. Graf, C. Ebner, Characterization of allergens in Apiaceae spices: Anise, fennel, coriander and cumin, Clin. Exp. Allergy. 27 (1997) 1299–1306. https://doi.org/10.1111/j.1365-2222.1997.tb01175.x.
[67] K. Hoffmann-Sommergruber, G. Riordain, H. Ahorn, C. Ebner, M. Laimer Da Camara Machado, H. Pühringer, O. Scheiner, H. Breiteneder, 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, Clin. Exp. Allergy. 29 (1999) 840–847. https://doi.org/10.1046/j.1365-2222.1999.00529.x.
[68] E.A. Pastorello, L. Farioli, C. Stafylaraki, J. Scibilia, M.G. Giuffrida, A. Mascheri, M. Piantanida, C. Baro, L. Primavesi, M. Nichelatti, J.W. Schroeder, V. Pravettoni,

Fennel allergy is a lipid-transfer protein (LTP)-related food hypersensitivity associated with peach allergy, J. Agric. Food Chem. 61 (2013) 740–746. https://doi.org/10.1021/jf303291k.
[69] E.A. Pastorello, A.M. Robino, Clinical role of lipid transfer proteins in food allergy, Mol. Nutr. Food Res. 48 (2004) 356–362. https://doi.org/10.1002/mnfr.200400047.
[70] F. Borghesan, G. Mistrello, S. Amato, M.G. Giuffrida, D. Villalta, R. Asero, Mugwort-fennel-allergy-syndrome associated with sensitization to an allergen homologous to Api g 5, Eur. Ann. Allergy Clin. Immunol. 45 (2013) 130–137.
[71] S. Dölle, K. Lehmann, D. Schwarz, W. Weckwert, C. Scheler, E. George, P. Franken, M. Worm, Allergenic activity of different tomato cultivars in tomato allergic subjects, Clin. Exp. Allergy. 41 (2011) 1643–1652. https://doi.org/10.1111/j.1365-2222.2011.03841.x.
[72] M. Maeda-Yamamoto, K. Ema, M. Monobe, I. Shibuichi, Y. Shinoda, T. Yamamoto, T. Fujisawa, The Efficacy of Early Treatment of Seasonal Allergic Rhinitis with Benifuuki Green Tea Containing O-methylated Catechin before Pollen Exposure: An Open Randomized Study, Allergol. Int. 58 (2009) 437–444. https://doi.org/10.2332/allergolint.08-OA-0066.
[73] M.L. Sanz, M.D. De Las Marinas, J. Fernández, P.M. Gamboa, Lupin allergy: a hidden killer in the home: Lupin allergy, Clin. Exp. Allergy. 40 (2010) 1461–1466. https://doi.org/10.1111/j.1365-2222.2010.03590.x.
[74] J. Trcka, S.G. Schäd, S. Scheurer, A. Conti, S. Vieths, G. Gross, A. Trautmann, Rice-Induced Anaphylaxis: IgE-Mediated Allergy against a 56-kDa Glycoprotein, Int. Arch. Allergy Immunol. 158 (2012) 9–17. https://doi.org/10.1159/000330641.
[75] M. Curin, R. Reininger, I. Swoboda, M. Focke, R. Valenta, S. Spitzauer, Skin Prick Test Extracts for Dog Allergy Diagnosis Show Considerable Variations Regarding the Content of Major and Minor Dog Allergens, Int. Arch. Allergy Immunol. 154 (2011) 258–263. https://doi.org/10.1159/000321113.
[76] M. Focke, K. Marth, R. Valenta, Molecular composition and biological activity of commercial birch pollen allergen extracts, Eur. J. Clin. Invest. 39 (2009) 429–436. https://doi.org/10.1111/j.1365-2362.2009.02109.x.
[77] S. Hildebrandt, H.D. Kratzin, R. Schaller, R. Fritsché, H. Steinhart, A. Paschke, In Vitro Determination of the Allergenic Potential of Technologically Altered Hen's Egg, J. Agric. Food Chem. 56 (2008) 1727–1733. https://doi.org/10.1021/jf0725981.
[78] K.R. Kottapalli, P. Payton, R. Rakwal, G.K. Agrawal, J. Shibato, M. Burow, N. Puppala, Proteomics analysis of mature seed of four peanut cultivars using two-dimensional gel electrophoresis reveals distinct differential expression of storage, anti-nutritional, and allergenic proteins, Plant Sci. 175 (2008) 321–329. https://doi.org/10.1016/j.plantsci.2008.05.005.
[79] J.N. Larsen, S. Dreborg, Standardization of Allergen Extracts, in: M.G. Jones, P. Lympany (Eds.), Allergy Methods Protoc., Humana Press, Totowa, NJ, 2008: pp. 133–145. http://link.springer.com/10.1007/978-1-59745-366-0_12 (accessed January 13, 2017).
[80] A. Matthes, M. Schmitz-Eiberger, Apple (Malus domestica L. Borkh.) Allergen Mal d 1: Effect of Cultivar, Cultivation System, and Storage Conditions, J. Agric. Food Chem. 57 (2009) 10548–10553. https://doi.org/10.1021/jf901938q.
[81] V. van Kampen, R. Merget, S. Rabstein, I. Sander, T. Bruening, H.C. Broding, C. Keller, H. Muesken, A. Overlack, G. Schultze-Werninghaus, J. Walusiak, M. Raulf-Heimsoth, Comparison of wheat and rye flour solutions for skin prick testing: a multi-centre study (Stad 1), Clin. Exp. Allergy. 39 (2009) 1896–1902. https://doi.org/10.1111/j.1365-2222.2009.03342.x.

[82] L. Monaci, A. Visconti, Immunochemical and DNA-based methods in food allergen analysis and quality assurance perspectives, Trends Food Sci. Technol. 21 (2010) 272–283. https://doi.org/10.1016/j.tifs.2010.02.003.
[83] K. Török, L. Hajas, V. Horváth, E. Schall, Z. Bugyi, S. Kemény, S. Tömösközi, Identification of the factors affecting the analytical results of food allergen ELISA methods, Eur. Food Res. Technol. 241 (2015) 127–136. https://doi.org/10.1007/s00217-015-2441-y.
[84] H. Van, Food allergen detection methods and the challenge to protect food-allergic consumers, Anal. Bioanal. Chem. 389 (2007) 111–118. https://doi.org/10.1007/s00216-007-1353-5.
[85] T.-J. Fu, N. Maks, Impact of thermal processing on elisa detection of peanut allergens, J. Agric. Food Chem. 61 (2013) 5649–5658. https://doi.org/10.1021/jf304920h.
[86] L. Monaci, M. Brohée, V. Tregoat, H. Van, Influence of baking time and matrix effects on the detection of milk allergens in cookie model food system by ELISA, Food Chem. 127 (2011) 669–675. https://doi.org/10.1016/j.foodchem.2010.12.113.
[87] M. Koeberl, D. Clarke, A.L. Lopata, Next Generation of Food Allergen Quantification Using Mass Spectrometric Systems, J. Proteome Res. 13 (2014) 3499–3509. https://doi.org/10.1021/pr500247r.
[88] G. Mamone, G. Picariello, F. Addeo, P. Ferranti, Proteomic analysis in allergy and intolerance to wheat products, Expert Rev. Proteomics. 8 (2011) 95–115. https://doi.org/10.1586/epr.10.98.
[89] T. Rabilloud, Two-dimensional gel electrophoresis in proteomics: Old, old fashioned, but it still climbs up the mountains, Proteomics. 2 (2002) 3–10. https://doi.org/10.1002/1615-9861(200201)2:1<3::AID-PROT3>3.0.CO;2-R.
[90] J.L. López, Two-dimensional electrophoresis in proteome expression analysis, J. Chromatogr. B. 849 (2007) 190–202. https://doi.org/10.1016/j.jchromb.2006.11.049.
[91] E.P. Romijn, J. Krijgsveld, A.J.R. Heck, Recent liquid chromatographic-(tandem) mass spectrometric applications in proteomics, J. Chromatogr. A. 1000 (2003) 589–608. https://doi.org/10.1016/S0021-9673(03)00178-X.
[92] F. Xie, R.D. Smith, Y. Shen, Advanced proteomic liquid chromatography, J. Chromatogr. A. 1261 (2012) 78–90. https://doi.org/10.1016/j.chroma.2012.06.098.
[93] I. Neverova, E. Van, Role of chromatographic techniques in proteomic analysis, J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 815 (2005) 51–63. https://doi.org/10.1016/j.jchromb.2004.11.009.
[94] O. Carrette, P.R. Burkhard, J.-C. Sanchez, D.F. Hochstrasser, State-of-the-art two-dimensional gel electrophoresis: A key tool of proteomics research, Nat. Protoc. 1 (2006) 812–823. https://doi.org/10.1038/nprot.2006.104.
[95] G. Picariello, G. Mamone, A. Cutignano, A. Fontana, L. Zurlo, F. Addeo, P. Ferranti, Proteomics, Peptidomics, and Immunogenic Potential of Wheat Beer (Weissbier), J. Agric. Food Chem. 63 (2015) 3579–3586. https://doi.org/10.1021/acs.jafc.5b00631.
[96] D. Apostolovic, T.A.T. Tran, C. Hamsten, M. Starkhammar, T. Cirkovic Velickovic, M. van Hage, Immunoproteomics of processed beef proteins reveal novel galactose-α-1,3-galactose-containing allergens, Allergy. 69 (2014) 1308–1315. https://doi.org/10.1111/all.12462.
[97] K.M. Odedra, Milk allergy in adults and children, Nurs. Stand. R. Coll. Nurs. G. B. 1987. 29 (2015) 43–48. https://doi.org/10.7748/ns.29.44.43.e9729.
[98] K.A. Hettinga, F.M. Reina, S. Boeren, L. Zhang, G.H. Koppelman, D.S. Postma, J.J.M. Vervoort, A.H. Wijga, Difference in the Breast Milk Proteome between

Allergic and Non-Allergic Mothers, PLOS ONE. 10 (2015) e0122234. https://doi.org/10.1371/journal.pone.0122234.
[99] J. Goliáš, Z. Humlová, P. Halada, V. Hábová, I. Janatková, L. Tučková, Identification of Rice Proteins Recognized by the IgE Antibodies of Patients with Food Allergies, J. Agric. Food Chem. 61 (2013) 8851–8860. https://doi.org/10.1021/jf402759f.
[100] J.M. Tomm, D. van, C. Jende, J.C. Simon, R. Treudler, B. von, M. Averbeck, Identification of new potential allergens from nile perch (Lates niloticus) and Cod (Gadus morhua), J. Investig. Allergol. Clin. Immunol. 23 (2013) 159–167.
[101] I. François, K. Sandra, P. Sandra, Comprehensive liquid chromatography: Fundamental aspects and practical considerations-A review, Anal. Chim. Acta. 641 (2009) 14–31. https://doi.org/10.1016/j.aca.2009.03.041.
[102] J.C. Giddings, Two-dimensional separations:concept and promise, Anal. Chem. 56 (1984) 1258A–1268A+1270A.
[103] K. Horváth, J.N. Fairchild, G. Guiochon, Generation and limitations of peak capacity in online two-dimensional liquid chromatography, Anal. Chem. 81 (2009) 3879–3888. https://doi.org/10.1021/ac802694c.
[104] M. Sarrut, A. D'Attoma, S. Heinisch, Optimization of conditions in on-line comprehensive two-dimensional reversed phase liquid chromatography. Experimental comparison with one-dimensional reversed phase liquid chromatography for the separation of peptides, J. Chromatogr. A. 1421 (2015) 48–59. https://doi.org/10.1016/j.chroma.2015.08.052.
[105] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography, Anal. Chem. 77 (2005) 6426–6434. https://doi.org/10.1021/ac050923i.
[106] A.F.G. Gargano, M. Duffin, P. Navarro, P.J. Schoenmakers, Reducing Dilution and Analysis Time in Online Comprehensive Two-Dimensional Liquid Chromatography by Active Modulation, Anal. Chem. 88 (2016) 1785–1793. https://doi.org/10.1021/acs.analchem.5b04051.
[107] R. Simon, S. Passeron, J. Lemoine, A. Salvador, Hydrophilic interaction liquid chromatography as second dimension in multidimensional chromatography with an anionic trapping strategy: Application to prostate-specific antigen quantification, J. Chromatogr. A. 1354 (2014) 75–84. https://doi.org/10.1016/j.chroma.2014.05.063.
[108] S.R. Wilson, M. Jankowski, M. Pepaj, A. Mihailova, F. Boix, G. Vivo Truyols, E. Lundanes, T. Greibrokk, 2D LC Separation and Determination of Bradykinin in Rat Muscle Tissue Dialysate with On-Line SPE-HILIC-SPE-RP-MS, Chromatographia. 66 (2007) 469–474. https://doi.org/10.1365/s10337-007-0341-4.
[109] A. Holm, E. Storbråten, A. Mihailova, B. Karaszewski, E. Lundanes, T. Greibrokk, Combined solid-phase extraction and 2D LC-MS for characterization of the neuropeptides in rat-brain tissue, Anal. Bioanal. Chem. 382 (2005) 751–759. https://doi.org/10.1007/s00216-005-3146-z.
[110] Z. Li, K. Chen, M. Guo, D. Tang, Two-dimensional liquid chromatography and its application in traditional Chinese medicine analysis and metabonomic investigation: Liquid Chromatography, J. Sep. Sci. 39 (2016) 21–37. https://doi.org/10.1002/jssc.201500634.
[111] M. Verstraeten, M. Pursch, P. Eckerle, J. Luong, G. Desmet, Thermal Modulation for Multidimensional Liquid Chromatography Separations Using Low-Thermal-Mass Liquid Chromatography (LC), Anal. Chem. 83 (2011) 7053–7060. https://doi.org/10.1021/ac201207t.

[112] Y. Wang, R. Lehmann, X. Lu, X. Zhao, G. Xu, Novel, fully automatic hydrophilic interaction/reversed-phase column-switching high-performance liquid chromatographic system for the complementary analysis of polar and apolar compounds in complex samples, J. Chromatogr. A. 1204 (2008) 28–34. https://doi.org/10.1016/j.chroma.2008.07.010.
[113] S.S. Jakobsen, J.H. Christensen, S. Verdier, C.R. Mallet, N.J. Nielsen, Increasing Flexibility in Two-Dimensional Liquid Chromatography by Pulsed Elution of the First Dimension: A Proof of Concept, Anal. Chem. 89 (2017) 8723–8730. https://doi.org/10.1021/acs.analchem.7b00758.
[114] D. Nardiello, C. Palermo, A. Natale, M. Quinto, D. Centonze, Strategies in protein sequencing and characterization: Multi-enzyme digestion coupled with alternate CID/ETD tandem mass spectrometry, Anal. Chim. Acta. 854 (2015) 106–117. https://doi.org/10.1016/j.aca.2014.10.053.
[115] X. Jiang, S. Feng, R. Tian, G. Han, X. Jiang, M. Ye, H. Zou, Automation of nanoflow liquid chromatography-tandem mass spectrometry for proteome analysis by using a strong cation exchange trap column, PROTEOMICS. 7 (2007) 528–539. https://doi.org/10.1002/pmic.200600661.
[116] A.L. Capriotti, C. Cavaliere, A. Cavazzini, F. Gasparrini, G. Pierri, S. Piovesana, A. Laganà, A multidimensional liquid chromatography–tandem mass spectrometry platform to improve protein identification in high-throughput shotgun proteomics, J. Chromatogr. A. 1498 (2017) 176–182. https://doi.org/10.1016/j.chroma.2017.03.032.
[117] A. Bathke, D. Klemm, C. Gstóttner, C. Bell, R. Kopf, Rapid online reduction and characterization of protein modifications using fully automated two-dimensional high performance liquid chromatography-mass spectrometry, LC-GC Eur. 31 (2018) 10–21.
[118] M. Alvarez, G. Tremintin, J. Wang, M. Eng, Y.-H. Kao, J. Jeong, V.T. Ling, O.V. Borisov, On-line characterization of monoclonal antibody variants by liquid chromatography–mass spectrometry operating in a two-dimensional format, Anal. Biochem. 419 (2011) 17–25. https://doi.org/10.1016/j.ab.2011.07.033.
[119] B.Q. Tran, H. Loftheim, L. Reubsaet, E. Lundanes, T. Greibrokk, On-Line multitasking analytical proteomics: How to separate, reduce, alkylate and digest whole proteins in an on-Line multidimensional chromatography system coupled to MS, J. Sep. Sci. 31 (2008) 2913–2923. https://doi.org/10.1002/jssc.200800275.
[120] J. Ren, M.A. Beckner, K.B. Lynch, H. Chen, Z. Zhu, Y. Yang, A. Chen, Z. Qiao, S. Liu, J.J. Lu, Two-dimensional liquid chromatography consisting of twelve second-dimension columns for comprehensive analysis of intact proteins, Talanta. 182 (2018) 225–229. https://doi.org/10.1016/j.talanta.2018.01.072.
[121] P. Donato, F. Cacciola, E. Sommella, C. Fanali, L. Dugo, M. Dachà, P. Campiglia, E. Novellino, P. Dugo, L. Mondello, Online comprehensive RPLC × RPLC with mass spectrometry detection for the analysis of proteome samples, Anal. Chem. 83 (2011) 2485–2491. https://doi.org/10.1021/ac102656b.
[122] J.G. Dorsey, Editorial on “Fast, comprehensive two-dimensional liquid chromatography” by D.R. Stoll, X. Li, X. Wang, P.W. Carr, S.E.G. Porter, and S.C. Rutan, J. Chromatogr. A. 1168 (2007) 2. https://doi.org/10.1016/j.chroma.2007.08.029.
[123] P. Dugo, F. Cacciola, T. Kumm, G. Dugo, L. Mondello, Comprehensive multidimensional liquid chromatography: Theory and applications, J. Chromatogr. A. 1184 (2008) 353–368. https://doi.org/10.1016/j.chroma.2007.06.074.

[124] S.R. Groskreutz, M.M. Swenson, L.B. Secor, D.R. Stoll, Selective comprehensive multi-dimensional separation for resolution enhancement in high performance liquid chromatography. Part I: Principles and instrumentation, J. Chromatogr. A. 1228 (2012) 31–40. https://doi.org/10.1016/j.chroma.2011.06.035.
[125] E. Sommella, F. Cacciola, P. Donato, P. Dugo, P. Campiglia, L. Mondello, Development of an online capillary comprehensive 2D-LC system for the analysis of proteome samples, J. Sep. Sci. 35 (2012) 530–533. https://doi.org/10.1002/jssc.201100877.
[126] X. Hou, M. Sun, X. He, L. Chen, P. Zhang, L. He, Simultaneous quantification of five proteins and seven additives in dairy products with a heart-cutting two-dimensional liquid chromatography method: Liquid Chromatography, J. Sep. Sci. 38 (2015) 3832–3838. https://doi.org/10.1002/jssc.201500684.
[127] A.D. Catherman, O.S. Skinner, N.L. Kelleher, Top Down proteomics: Facts and perspectives, Biochem. Biophys. Res. Commun. 445 (2014) 683–693. https://doi.org/10.1016/j.bbrc.2014.02.041.
[128] U. Andjelković, M. Gavrović-Jankulović, T. Martinović, D. Josić, Omics methods as a tool for investigation of food allergies, TrAC - Trends Anal. Chem. 96 (2017) 107–115. https://doi.org/10.1016/j.trac.2017.07.011.
[129] R. Korte, S. Lepski, J. Brockmeyer, Comprehensive peptide marker identification for the detection of multiple nut allergens using a non-targeted LC-HRMS multi-method, Anal. Bioanal. Chem. 408 (2016) 3059–3069. https://doi.org/10.1007/s00216-016-9384-4.
[130] L. Monaci, A. Visconti, Mass spectrometry-based proteomics methods for analysis of food allergens, TrAC - Trends Anal. Chem. 28 (2009) 581–591. https://doi.org/10.1016/j.trac.2009.02.013.
[131] L. Monaci, R. Pilolli, A. De, G. Mamone, Mass spectrometry in food allergen research, Compr. Anal. Chem. 68 (2015) 359–393. https://doi.org/10.1016/B978-0-444-63340-8.00007-8.
[132] L. Switzar, M. Giera, W.M.A. Niessen, Protein digestion: An overview of the available techniques and recent developments, J. Proteome Res. 12 (2013) 1067–1077. https://doi.org/10.1021/pr301201x.
[133] X.-L. Ge, T. Shi, H. Wang, J. Zhang, Z.-Q. Zhang, Development of an Aqueous Polyethylene Glycol-Based Extraction and Recovery Method for Almond (Prunus armeniaca L.) Protein, Food Anal. Methods. 9 (2016) 3319–3326. https://doi.org/10.1007/s12161-016-0525-3.
[134] S.U. Kadam, C. Álvarez, B.K. Tiwari, C.P. O'Donnell, Extraction and characterization of protein from Irish brown seaweed Ascophyllum nodosum, Food Res. Int. 99 (2017) 1021–1027. https://doi.org/10.1016/j.foodres.2016.07.018.
[135] D. Nardiello, A. Natale, C. Palermo, M. Quinto, D. Centonze, Combined use of peptide ion and normalized delta scores to evaluate milk authenticity by ion-trap based proteomics coupled with error tolerant searching, Talanta. 164 (2017) 684–692. https://doi.org/10.1016/j.talanta.2016.10.102.
[136] E. Papaleo, P. Fantucci, L. De Gioia, Effects of calcium binding on structure and autolysis regulation in trypsins. A molecular dynamics investigation, J. Chem. Theory Comput. 1 (2005) 1286–1297.
[137] E. Khodadadi, B.A. Fakheri, S. Aharizad, A. Emamjomeh, M. Norouzi, S. Komatsu, Leaf proteomics of drought-sensitive and -tolerant genotypes of fennel, Biochim. Biophys. Acta - Proteins Proteomics. 1865 (2017) 1433–1444. https://doi.org/10.1016/j.bbapap.2017.08.012.

[138] S. Fekete, A. Beck, J.-L. Veuthey, D. Guillarme, Theory and practice of size exclusion chromatography for the analysis of protein aggregates, J. Pharm. Biomed. Anal. 101 (2014) 161–173. https://doi.org/10.1016/j.jpba.2014.04.011.
[139] E.S.P. Bouvier, S.M. Koza, Advances in size-exclusion separations of proteins and polymers by UHPLC, TrAC Trends Anal. Chem. 63 (2014) 85–94. https://doi.org/10.1016/j.trac.2014.08.002.
[140] H.G. Barth, G.D. Saunders, The state of the art and future trends of size-exclusion chromatography packings and columns, LCGC N. Am. 30 (2012).
[141] T. Arakawa, D. Ejima, T. Li, J.S. Philo, The critical role of mobile phase composition in size exclusion chromatography of protein pharmaceuticals, J. Pharm. Sci. 99 (2010) 1674–1692. https://doi.org/10.1002/jps.21974.
[142] V. González-Ruiz, A.I. Olives, M.A. Martín, Core-shell particles lead the way to renewing high-performance liquid chromatography, TrAC Trends Anal. Chem. 64 (2015) 17–28. https://doi.org/10.1016/j.trac.2014.08.008.
[143] T. Rabilloud, A.R. Vaezzadeh, N. Potier, C. Lelong, E. Leize-Wagner, M. Chevallet, Power and limitations of electrophoretic separations in proteomics strategies, Mass Spectrom. Rev. 28 (2009) 816–843. https://doi.org/10.1002/mas.20204.
[144] I. François, A. de Villiers, B. Tienpont, F. David, P. Sandra, Comprehensive two-dimensional liquid chromatography applying two parallel columns in the second dimension, J. Chromatogr. A. 1178 (2008) 33–42. https://doi.org/10.1016/j.chroma.2007.11.032.
[145] P. Dugo, F. Cacciola, T. Kumm, G. Dugo, L. Mondello, Comprehensive multidimensional liquid chromatography: Theory and applications, J. Chromatogr. A. 1184 (2008) 353–368. https://doi.org/10.1016/j.chroma.2007.06.074.
[146] S.R. Groskreutz, M.M. Swenson, L.B. Secor, D.R. Stoll, Selective comprehensive multi-dimensional separation for resolution enhancement in high performance liquid chromatography. Part I: Principles and instrumentation, J. Chromatogr. A. 1228 (2012) 31–40. https://doi.org/10.1016/j.chroma.2011.06.035.
[147] K. Sandra, G. Vanhoenacker, I. Vandenheede, M. Steenbeke, M. Joseph, P. Sandra, Multiple heart-cutting and comprehensive two-dimensional liquid chromatography hyphenated to mass spectrometry for the characterization of the antibody-drug conjugate ado-trastuzumab emtansine, J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci. 1032 (2016) 119–130. https://doi.org/10.1016/j.jchromb.2016.04.040.
[148] X. Hou, M. Sun, X. He, L. Chen, P. Zhang, L. He, Simultaneous quantification of five proteins and seven additives in dairy products with a heart-cutting two-dimensional liquid chromatography method, J. Sep. Sci. 38 (2015) 3832–3838. https://doi.org/10.1002/jssc.201500684.
[149] A. Baglai, M.H. Blokland, H.G.J. Mol, A.F.G. Gargano, S. van der Wal, P.J. Schoenmakers, Enhancing detectability of anabolic-steroid residues in bovine urine by actively modulated online comprehensive two-dimensional liquid chromatography – high-resolution mass spectrometry, Anal. Chim. Acta. 1013 (2018) 87–97. https://doi.org/10.1016/j.aca.2017.12.043.
[150] B. Ji, B. Xia, J. Liu, Y. Gao, L. Ding, Y. Zhou, Application of fractionized sampling and stacking for construction of an interface for online heart-cutting two-dimensional liquid chromatography, J. Chromatogr. A. 1466 (2016) 199–204. https://doi.org/10.1016/j.chroma.2016.09.014.
[151] P. Petersson, K. Haselmann, S. Buckenmaier, Multiple heart-cutting two dimensional liquid chromatography mass spectrometry: Towards real time

determination of related impurities of bio-pharmaceuticals in salt based separation methods, J. Chromatogr. A. 1468 (2016) 95–101.
[152] K. Zhang, Y. Li, M. Tsang, N.P. Chetwyn, Analysis of pharmaceutical impurities using multi-heartcutting 2D LC coupled with UV-charged aerosol MS detection, J. Sep. Sci. 36 (2013) 2986–2992. https://doi.org/10.1002/jssc.201300493.
[153] M. Pursch, S. Buckenmaier, Loop-based multiple heart-cutting two-dimensional liquid chromatography for target analysis in complex matrices, Anal. Chem. 87 (2015) 5310–5317. https://doi.org/10.1021/acs.analchem.5b00492.
[154] S. Fekete, A. Beck, E. Wagner, K. Vuignier, D. Guillarme, Adsorption and recovery issues of recombinant monoclonal antibodies in reversed-phase liquid chromatography†: Liquid Chromatography, J. Sep. Sci. 38 (2015) 1–8. https://doi.org/10.1002/jssc.201400996.
[155] X. Ding, H. Zhang, H. Chen, L. Wang, H. Qian, X. Qi, Extraction, purification and identification of antifreeze proteins from cold acclimated malting barley (Hordeum vulgare L.), Food Chem. 175 (2015) 74–81. https://doi.org/10.1016/j.foodchem.2014.11.027.
[156] X. Zhao, F. Chen, J. Chen, G. Gai, W. Xue, L. Li, Effects of AOT reverse micelle on properties of soy globulins, Food Chem. 111 (2008) 599–605. https://doi.org/10.1016/j.foodchem.2008.04.026.
[157] P. Das, G. Dutta, A. Barthakur, C.L. Mahanta, Tackling correlated responses during process optimisation of rapeseed meal protein extraction, Food Chem. 170 (2015) 62–73. https://doi.org/10.1016/j.foodchem.2014.08.053.
[158] M. Vergara-Barberán, M.J. Lerma-García, J.M. Herrero-Martínez, E.F. Simó-Alfonso, Use of an enzyme-assisted method to improve protein extraction from olive leaves, Food Chem. 169 (2015) 28–33. https://doi.org/10.1016/j.foodchem.2014.07.116.
[159] H. Li, K. Zhu, H. Zhou, W. Peng, X. Guo, Comparative study of four physical approaches about allergenicity of soybean protein isolate for infant formula, Food Agric. Immunol. 27 (2016) 604–623.
[160] H. Jia, T. Zhou, H. Zhu, L. Shen, P. He, Quantification of Gly m 5.0101 in Soybean and Soy Products by Liquid Chromatography-Tandem Mass Spectrometry, Molecules. 24 (2019) 68.
[161] A.B. Oyedeji, J.J. Mellem, O.A. Ijabadeniyi, Potential for enhanced soy storage protein breakdown and allergen reduction in soy-based foods produced with optimized sprouted soybeans, LWT. 98 (2018) 540–545.
[162] B. Hoffmann, S. Münch, F. Schwägele, C. Neusü\s s, W. Jira, A sensitive HPLC-MS/MS screening method for the simultaneous detection of lupine, pea, and soy proteins in meat products, Food Control. 71 (2017) 200–209.
[163] A. Amponsah, B. Nayak, Effects of Microwave and Ultrasound Assisted Extraction on the Recovery of Soy Proteins for Soy Allergen Detection, J. Food Sci. 81 (2016) T2876–T2885. https://doi.org/10.1111/1750-3841.13534.
[164] M. Perumalsamy, M.I. Batcha, Synergistic extraction of bovine serum albumin using polyethylene glycol based aqueous biphasic system, Process Biochem. 46 (2011) 494–497. https://doi.org/10.1016/j.procbio.2010.09.023.
[165] P. Sun, D. Li, Z. Li, B. Dong, F. Wang, Effects of glycinin on IgE-mediated increase of mast cell numbers and histamine release in the small intestine, J. Nutr. Biochem. 19 (2008) 627–633. https://doi.org/10.1016/j.jnutbio.2007.08.007.
[166] Y. Hao, Z. Zhan, P. Guo, X. Piao, D. Li, Soybean **β**-conglycinin-induced gut hypersensitivity reaction in a piglet model, Arch. Anim. Nutr. 63 (2009) 188–202. https://doi.org/10.1080/17450390902860026.

[167] S. Wilson, K. Blaschek, E. Gonzalez De Mejia, Allergenic proteins in soybean: Processing and reduction of P34 allergenicity, Nutr. Rev. 63 (2005) 47–58. https://doi.org/10.1301/nr.2005.feb.47-58.
[168] T.J. Puumalainen, S. Poikonen, A. Kotovuori, K. Vaali, N. Kalkkinen, T. Reunala, K. Turjanmaa, T. Palosuo, Napins, 2S albumins, are major allergens in oilseed rape and turnip rape, J. Allergy Clin. Immunol. 117 (2006) 426–432. https://doi.org/10.1016/j.jaci.2005.10.004.
[169] S.K. Sathe, H.H. Kshirsagar, K.H. Roux, Advances in seed protein research: A perspective on seed allergens, J. Food Sci. 70 (2005) R93–R120.
[170] M. Ebisawa, P. Brostedt, S. Sjölander, S. Sato, M.P. Borres, K. Ito, Gly m 2S albumin is a major allergen with a high diagnostic value in soybean-allergic children, J. Allergy Clin. Immunol. 132 (2013) 976–978.e5. https://doi.org/10.1016/j.jaci.2013.04.028.
[171] S. Quirce, M. Fernández-Nieto, F. Polo, J. Sastre, Soybean trypsin inhibitor is an occupational inhalant allergen [3], J. Allergy Clin. Immunol. 109 (2002) 178. https://doi.org/10.1067/mai.2002.120952.
[172] T. Zhou, S. Han, Z. Li, P. He, Purification and quantification of Kunitz trypsin inhibitor in soybean using two-dimensional liquid chromatography, Food Anal. Methods. 10 (2017) 3350–3360.
[173] C.-H. (W. . Chen, Review of a current role of mass spectrometry for proteome research, Anal. Chim. Acta. 624 (2008) 16–36. https://doi.org/10.1016/j.aca.2008.06.017.
[174] C. Piras, P. Roncada, P.M. Rodrigues, L. Bonizzi, A. Soggiu, Proteomics in food: Quality, safety, microbes, and allergens, Proteomics. 16 (2016) 799–815. https://doi.org/10.1002/pmic.201500369.
[175] A.G. Marshall, C.L. Hendrickson, G.S. Jackson, Fourier transform ion cyclotron resonance mass spectrometry: A primer, Mass Spectrom. Rev. 17 (1998) 1–35. https://doi.org/10.1002/(SICI)1098-2787(1998)17:1<1::AID-MAS1>3.0.CO;2-K.
[176] J.S. Page, C.D. Masselon, R.D. Smith, FTICR mass spectrometry for qualitative and quantitative bioanalyses, Curr. Opin. Biotechnol. 15 (2004) 3–11. https://doi.org/10.1016/j.copbio.2004.01.002.
[177] J. Bergquist, M. Palmblad, M. Wetterhall, P. Håkansson, K.E. Markides, Peptide mapping of proteins in human body fluids using electrospray ionization fourier transform ion cyclotron resonance mass spectrometry, Mass Spectrom. Rev. 21 (2002) 2–15. https://doi.org/10.1002/mas.10016.
[178] M. Witt, J. Fuchser, G. Baykut, Fourier transform ion cyclotron resonance mass spectrometry with NanolC/microelectrospray ionization and matrix-assisted laser desorption/ionization: Analytical performance in peptide mass fingerprint analysis, J. Am. Soc. Mass Spectrom. 14 (2003) 553–561. https://doi.org/10.1016/S1044-0305(03)00138-7.
[179] T.L. Quenzer, M.R. Emmett, C.L. Hendrickson, P.H. Kelly, A.G. Marshall, High sensitivity fourier transform ion cyclotron resonance mass spectrometry for biological analysis with nano-LC and microelectrospray ionization, Anal. Chem. 73 (2001) 1721–1725. https://doi.org/10.1021/ac001095q.
[180] M. Mann, P. Højrup, P. Roepstorff, Use of mass spectrometric molecular weight information to identify proteins in sequence databases, Biol. Mass Spectrom. 22 (1993) 338–345. https://doi.org/10.1002/bms.1200220605.
[181] J.R. Yates III, Database searching using mass spectrometry data, Electrophoresis. 19 (1998) 893–900. https://doi.org/10.1002/elps.1150190604.

[182] E.D. Dodds, B.H. Clowers, P.J. Hagerman, C.B. Lebrilla, Systematic characterization of high mass accuracy influence on false discovery and probability scoring in peptide mass fingerprinting, Anal. Biochem. 372 (2008) 156–166. https://doi.org/10.1016/j.ab.2007.10.009.
[183] D.M. Horn, E.C. Peters, H. Klock, A. Meyers, A. Brock, Improved protein identification using automated high mass measurement accuracy MALDI FT-ICR MS peptide mass fingerprinting, Int. J. Mass Spectrom. 238 (2004) 189–196. https://doi.org/10.1016/j.ijms.2004.03.016.
[184] S.B. Badgujar, V.V. Patel, A.H. Bandivdekar, Foeniculum vulgare Mill: A review of its botany, phytochemistry, pharmacology, contemporary application, and toxicology, BioMed Res. Int. 2014 (2014). https://doi.org/10.1155/2014/842674.
[185] F. Borghesan, G. Mistrello, S. Amato, M.G. Giuffrida, D. Villalta, R. Asero, Mugwort-fennel-allergy-syndrome associated with sensitization to an allergen homologous to Api g 5, Eur. Ann. Allergy Clin. Immunol. 45 (2013) 130–137.
[186] J.A. Karty, M.M.E. Ireland, Y.V. Brun, J.P. Reilly, Defining absolute confidence limits in the identification of Caulobacter proteins by peptide mass mapping, J. Proteome Res. 1 (2002) 325–335. https://doi.org/10.1021/pr025518b.
[187] R. Asero, G. Mistrello, D. Roncarolo, S. Amato, Detection of some safe plant-derived foods for LTP-allergic patients, Int. Arch. Allergy Immunol. 144 (2007) 57–63. https://doi.org/10.1159/000102615.

[i] ExPaSy Bioinformatics Resource Portal, https://web.expasy.org/peptide_mass/.