Analysis of food proteins and peptides by mass spectrometry-based techniques

Gianfranco Mamone, Gianluca Picariello, Simonetta Cairia, Francesco Addeo, Pasquale Ferranti

Abstract

Mass spectrometry has arguably become the core technology for the characterization of food proteins and peptides. The application of mass spectrometry-based techniques for the qualitative and quantitative analysis of the complex protein mixtures contained in most food preparations is playing a decisive role in the understanding of their nature, structure, functional properties and impact on human health. The application of mass spectrometry to protein analysis has been revolutionized in recent years by the development of soft ionization techniques such as electrospray ionization and matrix assisted laser desorption/ionization, and by the introduction of multi-stage and ‘hybrid’ analyzers able to generate de novo amino acid sequence information. The interfacing of mass spectrometry with protein databases has resulted in entirely new possibilities of protein characterization, including the high sensitivity mapping (femtomole to attomole levels) of post-translational and other chemical modifications, protein conformations and protein–protein and protein–ligand interactions, and in general for proteomic studies, building up the core platform of modern proteomic science. MS-based strategies to food and nutrition proteomics are now capable to address a wide range of analytical questions which include issues related to food quality and safety, certification and traceability of (typical) products, and to the definition of the structure/function relationship of food proteins and peptides. These different aspects are necessarily interconnected and can be effectively understood and elucidated only by use of integrated, up-to-date analytical approaches. In this review, the main aspects of current and perspective applications of mass spectrometry and proteomic technologies to the structural characterization of food proteins are presented, with focus on issues related to their detection, identification, and quantification, relevant for their biochemical, technological and toxicological aspects.

Keywords:
Proteomics
Peptidomics
Mass spectrometry
Food proteins and peptides
Food quality
Food safety

1. Introduction
2. Food quality: from raw materials to end-products
3. Food authenticity
4. Food functionality
   4.1. Integrated analytical strategies for the characterization of bioactive peptides
   4.2. Tracing the fate of food proteins after ingestion
   4.3. “Sensometabolomic” food analysis
   4.4. Interaction of proteins with other food components
5. Food safety
   5.1. Microbial contamination of food
   5.2. Proteomic strategies for isolation and characterization of food lectins
   5.3. Modifications induced by food processing
   5.4. MS in food allergy and intolerance
   5.5. MS and celiac disease (CD)
   5.6. Food from genetically modified organisms
6. Perspectives

© 2009 Elsevier B.V. All rights reserved.
1. Introduction

The recently developed analytical techniques applied to the study of living organisms are revealing a portrait in which the biomolecular assets and the interactive networks are much more intricate than what it was expected only a few years ago. Most foods are produced from living organisms and tissues and, thus, they entirely reflect the complexity of the biological systems from which they derive. Proteins and peptides are major constituents of foodstuffs and play a decisive role in determining their nutritional and functional properties. As part of extremely complex biochemical systems, food proteins are mixtures in which components arising from post-translational processing (phosphorylation, glycosylation, disulfide bridges), and alternative splicing of pre-messenger RNAs, simultaneously occur. The technological processes used in the preparation of food products further contribute to enhance the heterogeneity of the protein systems by inducing a number of chemical reactions such as proteolysis, non-enzymatic glycosylation and other relevant modifications. An additional source of complexity which hampers the complete characterization of the food proteomes is the wide dynamic range associated to the protein expression.

Similarly, to the biological systems from which it comes, the proteome of a food is a dynamic entity which reflects the environment and physiological state of the origin organisms and evolves in time. In fact, food proteins, for which the natural biological turnover is arrested, can particularly be subjected to further change during food storing and maturation. Therefore, efforts to fully characterize the food proteomes are often nothing but snapshots of an ever-changing system. For instance, proteolytic phenomena heavily modify the picture of the protein content in foods such as cheese, yogurt, and fermented meat products.

Recently, mass spectrometry (MS) has emerged as an indispensable and irreplaceable tool to analyze proteins and peptide mixtures arising from their proteolytic degradation. Among the toolkit of techniques developed to investigate proteins at the proteome-wide scale, MS has gained popularity especially because of its ability to handle the hierarchical complexity associated with the biological systems. For the same reason, the applications of MS in the study of food proteins and the range of resolvable analytical questions are rapidly increasing. MS is among the most sensitive techniques now available; it provides specificity, speed and reliability of the analytical response in a high-throughput fashion. The need of comprehensive proteomic analyses has prompted the development of progressively more advanced technologies that push the boundary of MS capabilities, creating a “virtuous circle” which in return has allowed MS to address ever-increasing analytical challenges in the study of biological systems.

The depth of informative proteome analysis achieved by MS is unreached by other techniques developed for proteomic purposes such as two-dimensional gel electrophoresis (2DE), two-hybrid analysis, and protein microarrays. The utility of these latter, however, is unquestionable, especially when combined with the more specific response provided by MS-based techniques.

MS analysis of biomolecules has been revolutionized by the development of electrospray (ESI) and matrix assisted laser desorption/ionization (MALDI) mass spectrometry, the two soft techniques of ionization. The extreme accuracy of ESI–MS for measuring protein molecular weight is complemented by the capability of MALDI–MS to analyze proteins with a mass greater than 100 kDa or polypeptides in complex mixtures. The technological development of mass analyzers, devices that separate ionized molecules according to their m/z ratio, has also been a fundamental issue in proteomic research [1]. Basically, four types of mass analyzers have found large application in proteomics: quadrupole (Q), ion trap (quadrupole ion trap, QIT, linear ion trap, LIT or LTQ), time-of-flight (TOF), and Fourier-transform ion cyclotron resonance (FTICR). They strongly differ in both the physical principles of ion separation and the analytical performances. Multi-stage and ‘hybrid’ instruments have been designed to combine the capabilities of different mass analyzers and include the Q–q–Q (triple quadrupole), Q–q–LIT, Q–TOF, TOF–TOF, and LTQ–FTICR. A new type of hybrid instrument referable to the LTQ–Fourier-transform technology is the commercially available LTQ–Orbitrap system [1]. The mass spectrometer can directly provide information on the mass of a particular peptide but can also be used to generate de novo amino acid sequence information, including post-translational modifications (PTMs), from tandem mass spectra (MS/MS) obtained either by post-source decay (PSD) or, especially in the hybrid arrays, collision induced dissociation (CID).

Very recent technological improvements further enlarge capabilities of tandem mass spectrometry strategies. Electron transfer dissociation (ETD), electron capture dissociation (ECD), photodissociation, are emerging as new tools in the analytical armamentarium to sequence peptides from complex samples and to characterize PTMs. ETD–MS/MS provides increased sequence coverage for small- to medium-sized peptides and perfectly complements conventional CID for “bottom–up” protein identifications. ECD–MS/MS yields richer fragmentation peptide spectra and, because of its non-ergodic nature, it preserves the labile PTMs. ECD is becoming a powerful means to carry out “top–down” identification of proteins. A detailed comparison of the characteristics and capabilities of the different instrumental settings is out of the purpose of the present review. However, many excellent surveys dealing with MS instrumentation are already available into the literature. Synoptic outlines of the MS instrumental properties and performances can be found in [1,2].

The mass spectrometric characterization of either 2D-separated proteins, or capillary HPLC-fractionated peptides, can monitor, in very complex samples, presence, expression level of single components, and map PTMs, which of the utmost importance in the study of food proteins, such as those of meat, cereals, or milk. Several combinations of electrophoretic and chromatographic strategies, developed to reduce the complexity of the biomolecules to be analyzed via MS, allows to investigate both “rough” or “deep” proteomes. In order to simplify sample preparation and to reduce times for an effective on-probe investigation, surface enhanced laser desorption/ionization (SELDI)–TOF–MS has revealed itself of utility in some applications [3].

The so-called multi-dimensional protein identification technology (MudPIT), pioneered by the laboratory of J. Yates III [4] has significantly increased the number of proteins detected by shotgun proteomics. MudPIT strategies are now substituting the role played until a few years ago by the 2DE–MS proteomic approach, overcoming many of the limitations suffered by 2DE. In particular, the hyphenation of MS and MS/MS with liquid chromatography has become the choice methodology in the different applications of quantitative proteomics which make use of isotope-coded protein label (reviews on this focus [5], while 2DE-based strategies of quantitation, such as differential DIGE, fail in their purposes in many cases [6]. However, at present, applications of quantification reagents in food science are still limited to a few cases [7].

In the achievement of the progresses in the study of proteins, it has not to be neglected the contemporary advancement of bioinformatics which allows to manage and store the crop of data generated at the genomic and proteomic level [8].

Given the large collection of methodologies available, it is clear that the use of MS in the study of proteins in general does not consist in the application of a single technique for all purposes. It is rather a series of tools, each of them best suited for a particular individual case. For any MS experiment, consideration should be given to the type of instrumentation, fragmentation method,
and overall strategy with respect to the contingent analytical inquiries.

In this review, through a forcibly not exhaustive series of recent successful research case studies taken as examples, we illustrate how mass spectrometry can be applied to the characterization of complex proteome and peptidome of food products. Description of MS-based studies on foods is categorized according to four groups of relevant issues in the field of food sciences: quality, authenticity, functionality, and safety.

2. Food quality: from raw materials to end-products

Proteins are largely responsible of the overall characteristics of many food preparations. Nutritional, rheological and sensory properties of milk-, cereal-, and meat-derived products depend on the composition of the protein fraction. This is in turn determined by the combination of genetic factors with those introduced by the technological processing. The most remarkable case is probably that of wheat flour-derived products, whose optimal characteristics are determined by the gluten proteins. They are conventionally subdivided into in two major groups: the monomeric gliadins and the polymeric glutenins. This difference in solubility largely reflects the ability of these proteins to form inter- or intra-molecular disulfide bonds.

Gliadins, traditionally subdivided into α/β, γ and ω fractions, according to their mobility in polyacrylamide gel electrophoresis at acid pH (A-PAGE), are alcohol-soluble components representing about half a gluten protein. Glutenin (GS) are subdivided into both LMW (in the molecular weight range 20,000–45,000) and HMW (in the molecular weight range 70,000–90,000), associated by inter-chain disulfide bonds. LMW–GS can be further subdivided on the basis of $p_I$ and molecular weight, into the B, C, and D groups. Southern blot analysis of wheat DNA has suggested that the number of genes encoding the α/β- and γ-type gliadin is extremely high (greater than 100 copies/haploid genome). In the light of this, conventional electrophoretic and chromatographic techniques alone, routinely used in prolamin analysis, must be considered inadequate when facing the problem of rationalizing the complexity.

The proteomic approach as applied to wheat gliadin analysis can complement studies at the gene level. The protein profiles of gliadin and LMW glutenin subunit components are very heterogeneous as evidenced by polyacrylamide gel electrophoresis in SDS (SDS-PAGE) and reverse phase HPLC (RP–HPLC) analysis, while a simpler protein profile are observed for the HMW glutenin proteins. The complexity of gluten protein can be attributed to the presence of two or three sets of homologous chromosomes in durum and bread wheat, respectively, and to additional polymorphism related to mutation of gluten protein genes into many allelic forms. ESI–MS and MALDI–TOF–MS have partly resolved the gluten heterogeneity, becoming an important tool for gliadins and glutenin characterization and providing a basis to classify wheat varieties at the molecular level.

Newly emerging technologies encompass MS and proteomics, as well as the rapidly expanding field of bioinformatics tools and interactive databases. These methods combine the analysis of wheat protein extracts by MS with statistical methods such as artificial neural network, partial least-squares regression, and principal component analysis in order to predict the variety or end-use quality of unknown wheat samples. MALDI–TOF–MS has been also used to study the alteration of gliadins during the baking process. The MALDI-based approach for the analysis of gliadin mixtures, although fast and specific, does not allow to identify the structural differences among individual gliadin components. The low accuracy at high mass and the complexity of the protein mixture, which gives rise to suppression phenomena in MALDI do not allow fine identification of the protein subtypes. Furthermore, the fact that gliadins can be modified by heating with a...
corresponding loss of resolution and a decrease in mass signal intensity suggests that MALDI–TOF–MS can be used only for semi-quantitative measurements of gliadins.

The most powerful approach in the field is obtained when the capability of MS in identifying protein structure is coupled to a high-resolution chromatographic or electrophoretic technique. LC/ESI–MS has been utilized for the characterization of B- and C-type LMW–GS in order to determine the number and relative molecular masses of the about forty LMW subunits and the number of cysteine residues present in the different subunits [20]. More recently a detailed characterization and identification of wheat gluten proteins has been carried out by RP–HPLC followed by high-resolution MALDI mass spectrometry (MS and MS/MS) of the resulting protein fractions and their tryptic peptide digests [21].

The aim of proteomic studies of gluten proteins is the identification of molecular markers to provide information about protein components linked to bread wheat quality and particularly to kernel hardness (Fig. 1). Though preliminary, this approach can further supply genome analysis with supplementary and complementary data since it comprises the study of functional rather than informational molecules. It is expected that proteomics combined with transcriptome analysis will become a very effective approach for controlling the genetic improvement of wheat. In this respect, an integrated approach involving heterologous expression, 2-DE, RP–HPLC and MS allowed to provide evidence that a naturally mutated α-gliadin with an extra cysteine residue was incorporated into the polymeric glutenin fraction because of mutations that affect cysteine number and distribution, thus influencing negatively the dough–making performance [22].

The technological processes required for industrial food preparations may induce chemical modifications in gluten proteins which may be relevant for the nutritional and toxicological characteristic of the products. These modifications include deamidation of glutamine residues [23] formation of pyroglutamyl peptides [24] and non-enzymatic glycosylation caused by heat treatments [25]. Furthermore, using FTICR–MS, the formation of tyrosine cross-links (other than the conventional disulfide bonds) during dough-making has been demonstrated to contribute to gluten development and to bread-making quality [16,26].

Proteomic analysis of cereals has not been restricted to the study of wheat proteins and it is currently applied to both major and minor species. Barley (Hordeum vulgare) and rice (Oryza sativa L.) are among the most studied crops and their proteome has been deeply investigated at different physiological stages. Cultivar comparisons and genetic mapping of polymorphic protein spots in doubled haploid barley populations have provided a way to link the genome to the proteome and identify proteins that can influence grain quality [27–29]. Great efforts have been produced also in the proteomic characterization of rice (Oryza sativa L.). Rice is the most important cereal crop in Asia, but it is worldwide spread and, furthermore, it is considered as a model cereal plant for genetic and molecular studies. An immense progress had been made in rice genome sequence analysis during the last decade. This prompted the researchers to identify the functions, modifications, and regulations of every encoded protein. Thus, the analysis of rice proteome has recently yielded remarkable progress to systematically analyze and characterize the functional role of various tissues and organelles in rice. Major proteins involved in growth or stress responses have been identified and preliminary information about phosphoproteome and protein–protein interactions in rice has been achieved. The rice proteome database is under construction with the final aims of the molecular cloning of rice genes and of the prediction of the function of unknown proteins [30,31]. Similar progress has been achieved for maize [32–34] and soybean [35,36]. In these and other crops and tubers, whose production in several countries is based on genetically modified (GM) cultivars, a further source of complexity might rise from alterations in the normal proteome or from occurrence of foreign proteins [37–39].

Another field where the contribution of MS is being of primary interest for food technology resides in milk protein chemistry. The most relevant example concerns casein polymorphism analysis. The ability of mass spectrometry to analyze peptide mixtures generated by enzymatic or chemical cleavage of proteins can be used to perform protein fingerprints in which the occurrence of a genetic variant can be revealed by the mass shift of one or more peptides with respect to a reference sequence. In most cases the entire protein sequence can be screened in a single experiment. An example of one such study concerns the characterization of the genetic variants of ovine αS1-casein [for a complete review see [40]]. LC/ESI–MS was used to localize the differences between the common αS1-casein variant C and the less spread variant D, showing that the difference was a simple silent substitution, which, however, affected the degree to which the protein was phosphorylated. As a consequence of the structural variation, milks containing variant D showed bad aptitude to coagulate and therefore had a poor cheese yield. This example shows how the characterization of alleles at the protein level by mass spectrometry can provide the necessary information for developing typing procedures aimed to increase the efficiency of selection and breeding programs, which ultimately contribute to the improvement of the cheese-making quality of milk. Very recently a MS-based strategy based either on MALDI–TOF or LC/ESI–MS targeted detection of signature peptides has been developed to identify internally deleted goat αS1-casein in bulk milk. Through the utilization of opportunely designed peptide internal standards the deleted and full-length αS1-casein variants were also quantified. Due to strong influence of the casein variants on the quality of milk and on the clotting aptitude, the strategy provides a means to help producers to decide in advance the destination of milk for drinking or for cheesemaking [41]. Procedures to isolate and identify high, medium, and low molecular weight peptides in cheese, based on LC/ESI–MS or MALDI–TOF–MS have led to identify the casein degradation kinetics by lactic acid bacteria (LAB) and to obtain information about the technological characteristics of cheeses [42,43].

The large chemical, physical, and compositional differences among the different food matrices rise a series of technical problems, which require the application of specifically dedicated strategies. For example, in proteomic analysis of meat sarcoplasmic proteins, when electrophoretic separation is carried out on immobilized pH gradient (IPG) strips across the pH range 3–10, the most widely used in routine proteomic analysis, almost all components migrate towards the alkaline side of the gel and are poorly focused. Furthermore, the similarity in their molecular weights impairs resolution of the classical approach. To improve the resolution amongst components, acetic acid–urea–triton-polycrylamide gel electrophoresis (AUT-PAGE) analysis in the first dimension and standard sodium dodecyl sulfate polycrylamide gel (SDS–PAGE) in the second dimension have been developed [44]. The addition of Triton X-100, a non-ionic detergent, into the gel induces a differential electrophoretic mobility of proteins as a result of the formation of mixed micelles between the detergent and the hydrophobic moieties of polypeptides, separating basic proteins with a criterion similar to reversed phase chromatography based on their hydrophobicity. The acid pH induces positive net charges, increasing with the isoelectric point of proteins, thus allowing enhanced resolution in the separation. Similar separation criteria have to be applied for analysis of wheat gliadins, where an IEF separation across pH 6–11 range has to be adopted [45]. However, in this last case, further difficulty is represented by the low number and the high molecular weight of peptides obtained when trypsin is used as proteolytic agent for gliadin gel spots, due to the scarcity of lysine and arginine residues in the gliadin sequences. This occur-
renee makes sequence analysis complicated, as most peptides are outside the working range for carrying out efficient fragmentation spectra by MS/MS. The enzyme chymotrypsin, having a broader specificity than trypsin, provided an efficient gliadin spot digestion with a greater number of medium-sized peptides, better suited for ESI–MS/MS analysis.

Other approaches specifically developed to overcome the drawback represented by the huge dynamic concentration range of proteins in many food samples are based on enrichment procedures through the use of combinatorial peptide ligand libraries combined to MS-based techniques for detection of low-abundance proteins. They have been successfully applied to the analysis of several complex food matrices such as the chicken egg to reveal minor components [46].

Although these examples indicated that a case-to-case strategy has to be conceived depending from the different sample origin, they also proved that the general proteomic approach can be successfully adapted to face the analytical challenges raised by the diversity of the food matrices.

3. Food authenticity

The European Food Safety Authority has recently established a comprehensive system of authentication and traceability of food and feed in order to ensure food safety for human consumption at every stages of production. The EC regulation No. 509/2006 [47] protects consumers through a system of effective and impartial controls that define the safeguard of the ‘Protected Designation of Origin’ (PDO).

From the legislative point of view, quality standards have been established through the requirement of quality labels which specify the chemical composition of each product and in the most cases also geographical origin and production methods. This requirement has prompted an increased interest in developing analytical methods to assess quality and authenticity of food products. The sensitivity, specificity and speed in the analytical response candidate MS-based techniques as among the most accurate and versatile strategies to face the challenging tasks in this field. Moreover, proteins encrypt highly distinctive information about the identity and the history of a specific product. Thus, it is not surprising that targeted MS-based analyses of proteins have been largely developed for food authentication purposes, alone or in combination with high-resolution separative techniques [48,49].

The wide-ranging contribution and potentials of MS to assess quality of milk-based products has recently been reviewed [50]. Several MS-based procedures have been developed to authenticate “raw materials” used in manufacturing of dairy products. For instance, adulteration of milk of higher commercial value with cheaper bovine milk is quite frequently. Taking advantage of the species-specific amino acid substitutions among the homologous protein chains which affect molecular weight of both caseins and whey proteins, it is possible to define fingerprinting profiles of milk from different species either by electrospray or MALDI–TOF–MS. Angeletti et al. [51] demonstrated the efficiency of MALDI–TOF–MS in the detection of bovine caseins in water buffalo mozzarella cheese. Analogously, adulteration by bovine milk was detected in ewe cheese [52]. On the basis of the identification of a water buffalo-specific γ-casein by capillary electrophoresis (CE) followed by off-line MS, an improved strategy to distinguish bovine and buffalo milk by monitoring the products of plasmin hydrolysis has recently been developed [53].

Selecting the whey proteins as biomarkers, Cozzolino et al. [54] identified the fraudulent addition of bovine milk to buffalo and ewe dairy products at levels as low as 2%. Similarly, the detection of bovine β-lactoglobulin (BLG) in goat milk by means of LC/ESI–MS has been reported [55]. Since the fraudulent contaminations by bovine milk are relatively easy to be discovered, an emerging deception in the manufacturing of dairy products is the use of curds deriving from countries out and even far from the geographical areas established by the disciplinary of production of PDO products. To establish the geographical origin of milk or curds is a challenging task, up to now for some aspects still unresolved. Following the draft delineated by the pioneering research of Caprioli and a co-worker [56] in the field of mass spectrometry profiling finalized to tissue imaging, developed for biological and biomedical purposes, the “molecular profiling” performed by MALDI–TOF–MS is emerging as a general tool for the discovery of biomarkers also applied to food authentication. In meat and fish authentication, and in bacterial identification MS-based techniques have the capability to substitute or complement other available strategies, such as the multiplex PCR assay.

The issue of food authentication, particularly urgent in the case of PDO food products, also requires efficient traceability systems, to ensure the correct application of the procedures throughout the productive processes and during distribution.

A relatively new application of proteomic/peptidomic technologies in the discovery of ‘process markers’, informative molecular markers indicative of specific raw ingredients, guide industrial processes or improve supply chain management. For example, the skim milk powders (MP) are largely utilized as feed or as ingredients for the manufacture of products ranging from industrial cheeses to fermented meat products, in which they are added as emulsifier, texturizers, and for their high nutritional value. Soybean proteins are also frequently added to a variety of food products not only for economic reasons, as in the case of MP, but also because of their interesting functional properties during the processing of food. The addition of plant proteins to MP is, however, illegal, practice; therefore the detection and identification of adulterations in MP, typically by addition of low levels of cheap plant proteins (legume, cereal proteins) is an important issue to assess genuineness of several widespread industrial food preparation categories. Because of the complexity of the matrices (both milk and vegetal protein sources) detection with traditional analytical methods (HPLC, electrophoresis) is unpracticable at the very low amounts of adulteration frequently used. Also, the official EU reference CZE method currently used to prove fraudulent addition of plant proteins to MP has limited reliability and accuracy. For these reasons, the application of proteomic methodologies, based on combined chromatographic and MS methods, opens up new possibilities in the routine quality control of MP. Two LC/ESI–MS/MS methods [57,58] were successful in detecting peptides originated from the major seed proteins of soy (glycinin, β-conglycinin) and pea (legumin, vicilin) in MP. They share the use of LC/ESI–Q–TOF–MS/MS aimed to sequence tryptic peptides from legume proteins. The first one included a pre-fractionation step of the (adulterated) MP samples. The second more general approach is based on an “untargeted” analysis combined with statistical PCA analysis. This method also allows to distinguish MP samples from different manufacturers or produced in different conditions (milk protein composition, heat treatment) of the production process.

4. Food functionality

4.1. Integrated analytical strategies for the characterization of bioactive peptides

Food proteins are source of biologically active peptides. The active peptide sequence is inactive within the parent protein until they are released during gastrointestinal digestion or processing of food. Although bioactive peptides derived from different ani-
Fig. 2. The food bioactive peptide analysis workflow illustrating integrated strategies for purification for the structural and functional characterization.

The bioactive peptides may act as regulatory compounds with hormone-like activity, exhibiting a wide range of biological functions. Recent analytical research on bioactive peptides in food has followed three main directions aimed at developing and validating sensitive and specific methods for (i) tracing the pathway of formation of bioactive peptides from the parent proteins; (ii) identifying the biological properties; (iii) improving the “positive” properties discovered in natural peptides by design of synthetic structural analogues or peptide mimetics. These studies are aimed to clarify the structure–activity relationship of peptides, essential information for the design of novel therapeutics or functional foods.

The newest analytical approaches combine mass spectrometry with cell biology, immunology, biochemistry, synthetic chemistry, and the use of combinatorial libraries to identify the patterns of peptide formation and the bioactivity of the peptides present in the sample. In the case of milk-derived protein/peptides, identification of bioactive components can be accomplished as illustrated in Fig. 2, following the general scheme: 1. Purification of peptides present in a food sample (milk, yogurt, cheese) using high-resolution separation techniques or selective methods for enrichment of some class of components (i.e. phosphopeptides). 2. Peptide structural characterization by MS analysis of intact or enzymatically digested peptides; as a general feature, this peptidomic approach combines a high-resolution separation technique (2-D gel electrophoresis, nanoHPLC) with MS or other micro-analytical techniques for single protein identification, characterization of post-translational modifications and database cataloguing. 3. In silico analysis for prediction of possible bioactive sequences. 4. Synthesis of pure peptides, or of peptide analogues or mimetics. 5. Confirmation of bioactivity through appropriate biological or cellular assays.

A recent illustrative example of the potentiality of these combined approaches is given by the characterization of the peptides naturally released in the whey during the production of buffalo Mozzarella cheese [67]. MS structural identification of peptides, coupled to the study of antioxidant properties and immunomodulatory effects exerted on intestinal cancer cells, identified two candidate peptides for bioactivity both deriving from β-casein fragments, and their mechanism of action. In this respect, it seems reasonable to predict that the integration of the new and complementary analytical approaches above illustrated will give a strong contribution to determine the structural features responsible of bioactivity, and the structure–function relationship of milk bioactive peptides.
Chabance et al. [71] identified the κ-casein-derived macropeptide and the N-terminal fragments of αs1-casein in the blood of an adult individual after ingestion of yogurt, by targeting selected sequences. Identification of the entire panel of food-derived peptides in blood plasma is complicated by the large occurrence of endogenous polypeptides but would be virtually faceable by using the available up-to-date strategies of high-resolution separation coupled to tandem MS. The stability of peptides to gastric digestion also has toxicological implications, as it is one of the criteria used to assess the allergenic potential of food proteins. Therefore, the MS identification of food peptides surviving gastrointestinal digestion provides a means to help in localizing the antigenic determinants (epitopes) along the allergenic protein chains.

In order to identify the peptides generated from protein digestion, several model systems, reproducing the gastrointestinal digestion have been developed. Proteolytic systems simulating the physiological digestive enzyme pools including pepsin and pancreatin have been used [72]. Agudelo et al. [73] designed a pilot plant to perform the continuous removal of digestion products during in vitro proteolysis, in order to mimic the in vivo process and follow the fate of potentially bioactive peptides. Finally, a model of the intestinal digestion allowed to identify potential epitopes involved in eliciting the immunological response to gluten proteins in celiac patients [74].

4.3. “Sensometabolomic” food analysis

One of the main aspects of research in food technology is the characterization of those compounds which give a food its peculiar taste, odor, flavor and in general the sensory characteristics highly appreciated by consumers. Although multiple studies have been performed in the past years to characterize the key taste compounds of several foods, the data available in the literature are rather contradictory, particularly in the case of fermented products (meat and dairy products, wine and beer) where the flavor compounds imparted to them are generated by a complex series of metabolic and technological processes. The taste-active molecules which differentiate the many different types of cheese are, for instance, the products of a cascade of lipolytic and proteolytic pathways which produce the large variety of compounds imparting them the typical sensory characteristics. The comprehensive spectrum of sensory active, low molecular weight compounds, coined “sensometabolome” [75], reflects the sensory phenotype and triggers the typical smell and taste of food products. The goal of sensometabolomics is to catalogue, quantify, and evaluate the sensory activity of metabolites that are present in raw materials and/or are produced upon food processing such as thermal or high-hydrostatic pressure treatment and fermentation, respectively. The “sensometabolomic” approach consists in a combination of techniques including gel permeation chromatography, ultrafiltration, solid phase extraction, preparative RP–HPLC, and hydrophilic interaction chromatography (HILIC), and has its core in MS analysis combined with analytical sensory tools. Very recently, it enabled the comprehensive mapping of bitter-tasting metabolites of Gouda cheese [76], while LC/ESI–MS and LC/ESI–MS/MS, independent synthesis, and sensory analysis identified peptides responsible of bitter, umami and kokumi taste formed by proteolysis of caseins. The same approach has been applied to analysis of the characteristic olfactory sulfur compounds and their precursors which are produced by detoxification mechanisms in grapes and by yeast fermentation during wine–making of some typical wines, such as Sauvignon Blanc [77]. Among the objectives of sensometabolomics are the identifications of the most active compounds, and to determine their human threshold concentrations, to improve the olfactory and sensory characteristics of the derived food products.

4.4. Interaction of proteins with other food components

For their structural properties, proteins are able to establish either covalent or non-covalent interactions with other proteins or with different food components, including water, which results in changes in protein hydration, solubility, viscosity of solutions, film formation, gelling, and adsorption at the interface between aqueous and lipid phases. Examples of important food protein aggregations include the formation of cheese curd (aggregated, destabilized casein micelles) from acidification of milk or from chymosin-catalyzed hydrolysis of κ-casein, interaction and cross-linking of myosin chains causing loss of succulence and texture in meat and fish, aggregation of albumins in heated eggs, and development of gluten during dough-making.

Plant phenols bind dietary proteins and gut enzymes in either reversible or non-reversible way. This irreversible complexation may lead to loss in nutritional quality and to anti-nutritional and toxic effects [78]. Also, glucosinolates, occurring in a great number of Cruciferae vegetables, upon hydrolysis yield isothiocyanates (ITCs). ITCs react even under mild conditions with protein ε-amino or sulfhydryl groups to form thiourea or dithiocarbamate derivatives, respectively. Binding to ITCs affect the structural and functional properties of food proteins. Structural modifications in milk whey proteins, chosen as a model, induced by phenols have been characterized by MALDI–TOF–MS [79], whereas MS, also coupled to separative techniques, has been proposed as the elective analytical technique to monitor covalent changes induced in proteins by ITCs [80].

In the food industry, polyphenols are also involved in the formation of precipitates in beverages and in modifying the foaming and organoleptic properties. In a recent study [81], the interactions between plant polyphenols with milk β-casein and their effects on the surface properties at the air/liquid interface were monitored by combining rheological data with the information provided by ESI–MS and light scattering on the formation of non-covalent polyphenol–protein complexes. This integrated approach could probably be used as a basic methodology in the future, also applying the novel emerging MS techniques such as ion mobility MS for the studies of supramolecular complexes [82], to evaluate the changes in food functional properties (odor binding, nutrient transport) induced by the formation of protein complexes but also to use the binding protein of milk proteins such as β-casein or BLG to design new drug carrier for the pharmaceutical industry.

5. Food safety

To date, the use of MS in food safety has concentrated on two main areas, the detection of micro-organisms which may cause food spoilage or be hazardous to human health and the safety evaluation of food components.

5.1. Microbial contamination of food

In the first case, mass spectrometry is applied to the control and detection of food borne micro-organisms. Traditional means of controlling microbial spoilage and safety hazards in foods include freezing, blanching, sterilization, curing and use of preservatives. However, the developing consumer trend for ‘naturality’, as indicated by the strong growth in sales of organic and chilled food products, has resulted in a move towards milder food preservation techniques. This raises new challenges for the food industry. Proteomic approaches have been directed to the development of methods for bacterial profiling through MALDI–TOF–MS.
and ESI–MS/MS fingerprinting of bacterial proteins in order to distinguish among different species and, in some cases, among strains [83,84]. Through this profiling methods, it was possible fast and sensitive detection of pathogens or spoilage micro-organisms affecting food quality and safety during processing and storage. More accurate description of the contaminating micro-organisms have been achieved by integration of proteomics with peptidomic and metabolomic methodologies able to provide either structural or quantitative identification of specific metabolites produced by the various spoilage micro-organisms [85]. It can be foreseen that these methods are being integrated to design sensitive sensors on a microchip surface for automated detection. The “-omics” technologies can also help scientists to derive better understanding of the life cycles of bacteria. Defining the mode of action of food borne bacteria and the mechanisms that confer ‘stress resistance’ should enable more rational design of food preservation techniques. In addition, this information can also be used to pinpoint areas of the food chain that are most susceptible to microbial contamination.

In this respect, the analysis of pathogenic micro-organism deserves particular caution, as the risks associated to their contamination are not limited to their living presence and capacity of infectivity, but they can generally release protein/peptide toxins able to survive for long time even in foods after bacterial cell contamination has been removed, as it happens for many of the microbes which cause food borne diseases, including *Listeria monocytogenes, Staphylococcus aureus, Escherichia coli, Clostridium botulinum* A and various *Salmonella* species. All these pathogenic bacteria excrete a variety of virulence factors into extracellular medium and to the cell surface which have essential roles in the colonization and insurrection of the host cells, and thus reflect the degree of bacterial pathogenicity. These toxins, being heat-stable and resistant to proteases, can be a danger for the consumer health. For the exploration of virulence factors expressed in the secreted proteome fraction, in a very recent study different *Staphylococcus aureus* strains were analyzed using gel-based bottom–up proteomic approach [86]. A still more complex situation is expected for analysis of foods, mostly constituted by a highly complex matrix of proteins, lipids, carbohydrates and many other molecular species which interfere with detection of the predictable toxin amounts (in the order of ppb). For this reason, the combination of MS methodologies with advanced immunochemical, chromatographic and electrophoretic isolation procedures has to been applied. One such study has been carried out to define the toxin contamination levels of two ripened PDO Italian cheeses, Grana Padano and Pecorino Romano. A procedure combining proteomic approach with immunochromatographic, chromatographic and electrophoretic techniques and Tandem MS analyses was developed to monitor production and levels of Enterotoxin A (SEA) and B (SEB) of *S. aureus* and Shiga-like toxins produced by *E. coli* O157:H7, as outlined in Fig. 3. By producing cheese samples using milk willingly contaminated with bacteria it was possible to monitor 10–100 ppb contamination level, and analysis of randomly collected market samples allowed to exclude toxin contamination in the two cheese types [87].

The second main issue, safety evaluation of food components is concerned with the presence of toxic compounds which may be originally present in the raw material (and therefore need to be eliminated by the manufacturing process, as in the case of legume lectins) or conversely may be generated during the production process.

### 5.2. Proteomic strategies for isolation and characterization of food lectins

In humans, the consumption of raw kidney beans has been shown to cause gastroenteritis, nausea and diarrhea [88] and several severe but non-fatal poisoning outbreaks caused by ingestion of the abnormal levels of the lectin phytohemagglutinin (PHA) present in foods have been reported during years [89]. On the other hand, commercial protein concentrates of common beans, the so-called “starch-blockers” are more frequently used as dietary supplements to control body overweight in therapy of obesity [90], notwithstanding the absence of reliable scientific evidence for their efficacy [91]. The increasing success of these supplements among patients and consumers call the need for appropriate methods to determine lectin levels, which are currently based on colorimetric assays, high performance anion exchange chromatography with pulsed amperometric detector (HPAEC–PAD) analysis, or toxicity tests [92], all of which are of necessity unspecified.

The normal range of lectin levels in *Phaseolus vulgaris* is 1–10 g/kg (and up to 5% seed dry weight), which also implies a risk for health in processed food [93]. The most common method of destruction is heat treatment in boiling water or by extrusion [94]. Considering the possible risk for health in processed food, it is important to introduce efficient heat-based or other kinds of treatments in the processing of bean derived food as well as

---

**Fig. 3.** Analytical MS approach for detection of *Staphylococcal* Enterotoxin B (SEB) in long-ripened cheese.
techniques to ensure the absence of toxicity in these products. The development of specific and sensitive methods needs to rely on the lectin structural and functional definition, which can now be achieved by integrated proteomic and glycomic methodologies (Fig. 4).

For the recovery of lectins from crude leguminous extracts, legume seed proteins are usually suspended in appropriate aqueous buffers; solutions are clarified by centrifugation and the proteins in the supernatant are then precipitated by ammonium sulfate. Lectins can be purified by means of electrophoresis and reversed phase HPLC, or also by size exclusion (SEC) and ion-exchange (IEC) chromatography; these latter two techniques also preserve the tertiary and quaternary structure of the lectins, allowing to assay their biological properties (agglutinating profiles). Taking advantage of the specific affinity for carbohydrates (or glycoproteins), plant lectins are more effectively isolated using sugar-derived or sugar-immobilized stationary phases. Lectins with binding specificities for beta-D-galactosides, alpha-D-mannosides, alpha-L-fucosides, alpha-D-glucosides and for N-acetyl-D-glucosamine and N-acetyl-D-galactosamine have been isolated on agarose or sephadex resins with the proper carbohydrates immobilized as affinity ligands [95] (Fig. 4). After isolation, MS sequence confirmation or structural characterization of lectins can be carried out according to the gel electrophoresis-based or to the electrophoresis-free proteomic strategies. In the case of possibly glycosylated lectins, identification of glycosylation site(s) can be accomplished as illustrated in Fig. 4, following the general scheme: 1. Proteolytic digestion. 2. Capture of the glycopeptides choosing among several possible strategies including HIVIC, affinity enrichment by lectins and hydrazide resin glycol-capture [96]. 3. Chemical or enzymatic de-glycosylation (by peptide N-glycosidase F in the case of the N-glycosylation), preferably in H218O. 4. LC/ESI–MS/MS sequencing of deglycosylated peptides for identification of the linker amino acid. Characterization of sugar moieties can be performed after oligosaccharide isolation and MS analysis of native or derivatized glycans.

These integrated proteomic and glycomic strategies are able to define the structural changes (proteolysis, oxidation, sugar changes) which can occur in raw and in industrially treated products. In this way, they constitute the basis to allow for either the structural or quantitative analysis of PHAs in different origin samples. An illustrative example is the control of bean flour deriving from the industrial dry thermal treatment normally carried out to inactivate lectins [97]. A step forward will be the scaling-down of this procedure to an automated miniaturized scale through immobilization of the glycoprotein on a solid support to obtain an affinity sensor capable of detecting low-ppm amounts of native lectins in legumes and in the derived foods. This approach reverses the idea of "glycan-capture lectin array" to a "lectin-capture glycan array", as it uses an immobilized glycoprotein to detect and quantify PHA lectins.

5.3. Modifications induced by food processing

The most remarkable example of modifications induced by technological treatments which may affect the overall food quality concerns the products of the heat treatment on materials such as milk, meat and cereal products. The main modifications induced by heat treatment on milk proteins are the denaturation of whey proteins and the complex series of covalent reactions known as Maillard reactions [98], which produce a decrease in nutritional quality and the formation of possible toxic compounds. The covalent adducts between whey proteins and lactose in milk have been evidenced by MS-based techniques and used as qualitative markers to trace the thermal history of bovine milk [99–101]. On the other hand, well-controlled Maillard reaction can also be voluntarily induced to achieve specific benefits like aroma generation in bread and baked products or to improve the physico–chemical properties of whey proteins. The novel compounds formed have also been proposed as useful markers to demonstrate either uncorrected heat treatments or the presence of heated milks/milk powders added to fresh milks. Some markers of heated milks have been identified while the most interesting are the heat-susceptible phosphorylated serine and threonine residues of caseins yielding dehydroalanine and methyl–dehydroalanine, respectively [102]. The conjugated carbon–carbon double band reacting with nucleophilic amino– (Lys, His), or thiol–(Cys) group forms lysinoalanine (LAL), histidinoalanine and lanthionine residues, respectively [103]. Presence of LAL-containing casein has motivated researchers to develop procedures for LAL detection in milk products. Infant formulas based on different casein/whey ratio values present an upper limit of 1000 and 300 mg LAL kg−1 crude protein for liquid and dried milk products respectively. Two HPLC analytical methods separating 9-fluorenyl-methylchloroformate (FMOC) [104] and dansyl chloride–LAL derivative [105] are routinely applied for LAL quantification. The main difficulty lies in detecting trace amounts of LAL in presence of other dominating free amino acids. An analytical method involving detection by LC/ESI–MS in positive selected ion monitoring (SIM) of FMOC–LAL-derivative without any sample pre-fractionation step has been developed recently [106]. The procedure differs from others previously developed for dairy products in that confirmation and direct quantification of the FMOC–LAL identity is concurrently carried out. It must be highlighted, however, that the most recent studies indicate that LAL-containing proteins in heated foods does not have toxic effects; rather, LAL represents a very useful marker for detection of heat treated products in a variety of products ranging from drinking milk to MP to dairy products. A very different case is that of molecules whose production in foods may represent a serious health concern. In several heated foods, high levels of acylamide (AA) have been found as the product of the Maillard reaction between amino acids (mainly Asn, but also Gln and Met) and reducing sugars (p-fructose, p-galactose, lactose, glucose) in consequence of roasting, toasting, frying processes [107] MS-based methods find application in analysis of raw and treated materials, for instance hazelnuts and almonds, on which roasting is usually carried out either to generate the typical flavors or to allow for storing and further transformation processes. These methods were also used very recently to optimize the effects of cultivars and environmental growth conditions on the level of free amino acids in wheat grain, and, consequently, on the final AA levels during flour processing for the preparation of baked goods [108]. The high AA reactivity towards nucleophiles makes also this compound able to form covalent adducts with macromolecules such as proteins and DNA, which is responsible of its neurotoxic, clastogenic, carcinogenic and toxicant effects. On the other side the quantitative MS-based analysis of covalent adducts of AA with proteins (as well as of similar electrophilic toxicants) is at the basis of modern methods for evaluating the level of human exposure [109]. These data are then used to determine the acceptable daily intake of a particular food or food chemical. Although hazard analysis is clearly important, gathering appropriate data can be costly and time consuming, requiring detailed toxicological experimentation in animals, often on an empirical basis. Mass spectrometry-based technologies can offer a number of benefits when conducting toxicological evaluation as their high-throughput nature means that it is possible to analyze multiple samples in a timely and cost-effective manner.

5.4. MS in food allergy and intolerance

Food allergy is an increasingly important issue for the food science. Although the various diseases are limited to only cate-
gories of subjects, the number of patients who are interested is steadily increasing in the last ten years, making this field of primary importance for food industry. Mass spectrometry-driven identification of either genes for allergic diseases or allergenic proteins is being successfully carried out [110]. Systematic analysis of cereal and legume species such as wheat, rice, pea, soy, peanut using high-resolution separation techniques in combination with MS and multi-dimensional protein identification is leading to detection and identifcation of several previously uncharacterized allergenic proteins in seeds samples, thus demonstrating the potentiality of proteomic approaches to survey food samples with regard to the occurrence of allergens. By these studies, the basis for the development of LC/ESI–MS/MS-based methods with triple quadrupole mass analyzer has been founded, which allowed good detection limits (1 ppm) for almonds, pecan nuts, hazelnuts, walnuts antigens in food ingredients such as soy, milk, chocolate, cornflakes, and rice crisps.

Allergy to cow milk is one of the more prevailing food allergies in infants, and in several cases is due to intolerance for the main milk proteins β-casein and BLG. Despite enzymatic hydrolysis of major milk proteins is often employed for increasing food tolerance, residual antigens in hydrolyzed milk formulas have been reported. Recent studies have been directed to obtain high tolerable dairy products by limited proteolysis of milk proteins to remove sequences involved in allergic or intolerance phenomena. Limited proteolysis was carried out on transient conformers produced by heating at sub-denaturing temperatures and the resulting peptides were immediately removed by a simultaneous ultrafiltration step [111]. In these conditions, the peptides permeating from the ultrafiltration membrane are not further degraded to small peptides or amino acids, as occurs in current hydrolysis processes, thus preserving most of their functional and nutritional properties. The information collected by combined immunochemical and MS analysis in this phase made it straightforward to scale up the process at first to pilot-plant scale and, later, to pre-industrial scale [111].

5.5. MS and celiac disease (CD)

Gliadins are considered as one of the main factors triggering CD [112], a common enteropathy induced by ingestion of wheat gliadin and related prolams from oat (avenin), rye (secalin) and barley (hordein) in genetically susceptible individuals. The structural basis for gliadin toxicity in CD is not completely clear neither are understood the molecular basis of the toxicity, also because of the structural complexity of gliadins.

The high percentage of proline residues makes gliadin resistant to gastric–pancreatic and intestinal digestive proteases, so that long gliadin fragments can reach high concentration levels in the gut epithelium [113]. For this reasons, during endoluminal digestion, a family of Pro- and Gln-rich polypeptides that are responsible for the inappropriate T-cell-mediated immune response are released. Several fragments are recognized as toxic [114,115], whereas other peptides are shown to be immunostimulating such as peptides 31–43 of α2-gliadin and 31–49 of α-A-gliadin [117] and 33-mer epitope, corresponding to fragment 57–89 of α2-gliadin [116]. Interestingly, deamidation of gliadin peptides mediated by tissue transglutaminase (TG) increases peptide toxicity [118].

Peptidomics being an efficient tool for the rapid identification of peptides in complex mixtures, can facilitate identification of TG-mediated modifications of peptides. The resolution and specificity achieved when using the new generation of hybrid quadrupole-time-of-flight instruments allows differentiation of peptides where a single deamidated Q residues is present, even in complex mixtures, such as those occurring in the enzymatic digests of gluten proteins. The newest analytical approaches combine MS with cell biology, immunology, synthetic chemistry, and the use of combinatorial library to identify the HLA-binding motifs and T-cell recognition patterns in gliadin-derived peptide sequences. In order to detect gliadin peptides derived from gastric and pancreatic (PT) digestion possibly modified by TG, a peptidomic analytical approach capable of selective probing of susceptible Q was developed [74] which used a fluorescent chemical label to identify the TG-susceptible peptides in the complex PT digests.

At present, new therapeutic approaches are sought which would permit celiacs to “peacefully” coexist with gluten, including search for genetically modified wheat lacking toxic gluten peptides. Furthermore, two approaches are being tested to prevent or reduce gluten toxicity in a wheat flour: either masking of gliadin epitopes (possibly by use of TG) or their proteolytic degradation, all fields where MS is playing an important role in validating results. On the side of degradation of toxic epitopes in CD, the main difficulty is that the gluten-derived T-cell epitopes are highly resistant to proteolytic degradation within the gastrointestinal tract which does not possess the enzymatic equipment to efficiently cleave proline-rich peptides driving the abnormal immune intestinal response in CD patients. For this reason, oral supplementation with exogenous prolyl-endopeptidases produced for instance by from Aspergillus niger able to digest gluten has been proposed as an alternative treatment to the gluten-free diet. In vitro and in vivo studies driven by MS [119] confirmed in part the effectiveness of prolyl-endopeptidase an oral supplement to reduce gluten intake in patients. Alternative approaches for gluten degradation are based on gluten fermentation with different microbial media, including probiotic preparations [120] or sourdough Lactobacilli [121] and using 2DE, MALDI–TOF–MS and LC/MS analysis to monitor the effective disappearance of the toxic epitopes.

Although a toxic gluten-free diet is the only treatment for CD patients, a reliable in vitro or in vivo system to analyze toxicity in food samples to control diet is not available yet. Mass spectrometry is the only non-immunological method presently available to detect with high specificity gliadins and related prolams in flours and in food samples. A first approach is based on the possibility of obtaining characteristic MALDI–TOF–MS profiles of unfraccionated gliadin, hordein, secalin and avenin extracts [122] (see also the scheme of Fig. 1). Based on these four distinguishable mass patterns, prolams from different cereals can be differentiated and also identified when simultaneously present in foods [13]. A remarkable application is in the analysis of the products of the starch hydrolysis, such as glucose syrup, crystalline dextrose and maltodextrins (largely used as sweeteners, anti-crystallizers, and stabilizing agents), obtained industrially through chemical and/or enzymatic methods. In this products gluten determination by immunological tests is made unreliable by a series of factors, including the low amount of gluten to be detected being dispersed in a very high amount of substances (low and high mass sugars, other by-products of the process) which interfere with determination. Gluten semi-quantitative measurement in these products by combining procedures of extraction and isolation with MALDI–TOF–MS analysis made it possible to detect and identify low quantities of protein (estimated sensitivity 1–10 ppm), thus allowing to verify whether these products exceeded the 20 ppm limit required for foods “rendered” gluten-free [123]. These results could add useful information for developing diets and therapy in CD patients. Also, the pattern of proteins/peptides present in the samples was found to vary either qualitatively or quantitatively, depending on the sample type. This also meant that the MS approach may allow to identify the differences and quantify the protein/peptide level in different industrial products of the same category.
5.6. Food from genetically modified organisms

The safety assessment of GM plants and derived food and feed follows a “comparative” approach, i.e. the biomolecular expression pattern of food and feed is compared with their non-GM counterparts in order to predict intended and unintended effects.

One of the pitfalls in the safety assessment of GM foods is the concept of “substantial equivalence” formulated by Organisation for Economic Co-operation and Development in 1993, based on the idea that existing foods could serve as a baseline for comparing the properties of a GM food with its conventional counterpart (for a critical review about the topics see [38]. Effects of GM foods cannot be predicted exclusively from information about the insertion site of the genetic construct which does not take into account gene regulation, gene–gene interactions, and possible interferences in metabolic pathways. The modified metabolism, the possibility of novel fusion proteins, or other pleiotropic effects due to the particular site of transgene integration (e.g. interruption of important open reading frames or regulatory sequences) could importantly compromise food of feed safety [39].

To identify possible unintended effects due to the use of GM crops, targeted analysis of specific compounds, which represent the key of important metabolic pathways in the plant like macro- and micro-nutrients, known allergens, anti-nutrients and toxins, has to be carried out in parallel with the comparative phenotypic analysis of the GM plant and of its near isogenic counterpart. Furthermore, information about natural ranges of variation of the compounds included in comparative analyses is essential for both health and economical reasons. Differences in expression traits are subsequently assessed with respect to their potential impact on the environment, safety for humans and animals, and nutritional quality. Although a case-by-case analytical strategy should be applied, the MS-based broad scale profiling techniques of the new “omics” sciences offer a tool to envisage unintended effects in the biomolecular expression pattern [39]. Recently, proteomics has successfully complemented other existing safety assessment techniques for identifying side effects occurring in transgenic seeds of a variety of maize [32]. The protein expression profile in transgenic maize was compared to the corresponding isogenic control for two subsequent generations. The changes in protein expression of isogenic plants as response to the altered environmental conditions were also considered, revealing that the up- or down-regulation of several proteins in transgenic seeds was imputable to the insertion of a single gene into the maize genome. Similarly, analysis of the protein profiles of non-GM potato (Solanum tuberosum) genotypes compared to several GM lines, clearly indicated that the variation between the non-GM cultivars/genotypes was much greater than the differences of the GM lines with the respective natural counterpart [124]. These results were also corroborated by the comparisons of the metabolic patterns [125] and suggested that changes in the metabolic traits of GM lines can be strictly controlled and limited.

Proiling strategies based on the capillary electrophoresis (CE), alone [34] or opportunely combined with MS detection (CE–MS) [33], have also been applied to distinguish transgenic and non-transgenic soybean varieties and to compare zein fingerprints of GM and non-GM isogenic lines of maize, respectively.

In general, proteomic profiling, especially in combination with other profiling methods, is capable to reduce uncertainty in the prediction of unintended effects, by providing much more information about crop composition than targeted analysis alone does. In this respect, accurate quantification of each protein expression level made possible by specific isotope-coded protein label reagents are assuming increasing importance in differentiation of GM from natural profiles [5,7]. These factors make proteomics increasingly important when developing second generation GM crops with multiple genes, engineered metabolic pathways, or producing edible pharmaceuticals.

![Fig. 4. The analytical workflow for qualitative and quantitative characterization of legume lectins.](image-url)
6. Perspectives

Nutrition plays a crucial role in health as well as disease. Despite its youth, food proteomics is already influencing multiple aspects of the food chain (agriculture, food production, food safety and quality assurance) and is delivering economic benefits and improving aspects of human nutrition and health. By dissecting the proteome and peptidome in their constituents, using the new technologies, specific markers can be identified to trace the complete productive processes, from raw materials to end-products, and even food products originating from specific geographical areas can be characterized at the molecular level.

Over time, the integrated approach based on the various “omics” for food peptide and protein characterization, all of which rely on MS analysis, should accelerate the development of functional food products as well as increase knowledge to develop novel foods and ingredients for various applications, not limited to food industry. For example, the capacity of MS techniques to analyze supramolecular complexes is opening new scenarios in the understanding of the protein–protein, protein–carbohydrate, and protein–lipid interactions, which in turn determine the rheological properties of complex matrices; this is the case, for example, of wheat doughs and of their industrially designed non-conventional substitutes directed to intolerant patients, where tG-modified non-gluten proteins are being tested [126]. At the same time, the definition of the lipid-binding capacity of food derived proteins such as BLG is helping to develop totally new protein carriers for efficient and targeted drug delivery which are also highly resistant to gastric digestion [127].

Food allergy and intolerance are other fields in which MS has proved to be essential at various levels, including the controversial feature of resistance to digestion of allergenic proteins or the efficiency of removal of epitopes from a food destined to patients. The proper MS techniques, selected from time to time within the large range of possible applications, can provide means to identify the immunodominant epitope(s) in the target food, as starting point for designing antibodies against specific epitopes, and to standardize the quantitative detection of trace amounts. Therefore, large cost- and health-saving results are being achieved by application of the MS-based efficient monitoring procedures. Also the advances and future developments of protein array technology, which will be largely driven by MS-derived structural information, will ensure more rapid and accurate detection of food safety for the consumer. For all these reasons, employment of the novel MS techniques in food proteomics is expected to increase during the next years, due to their emerging potentiality in the molecular characterization of food products, aimed at the certification of novel as well as traditional foods.

References
