

MASS SPECTROMETRY IN THE PROTEOME ANALYSIS OF MATURE CEREAL KERNELS

Vincenzo Cunsolo, Vera Muccilli, Rosaria Saletti, and Salvatore Foti*

*Dipartimento di Scienze Chimiche, Università degli Studi di Catania,
Viale A. Doria, 6, I-95125 Catania, Italy*

Received 22 March 2011; revised 7 July 2011; accepted 7 July 2011

Published online 22 August 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/mas.20347

In the last decade, the improved performance and versatility of the mass spectrometers together with the increasing availability of gene and genomic sequence database, led the mass spectrometry to become an indispensable tool for either protein and proteome analyses in cereals. Mass spectrometric works on prolamins have rapidly evolved from the determination of the molecular masses of proteins to the proteomic approaches aimed to a large-scale protein identification and study of functional and regulatory aspects of proteins. Mass spectrometry coupled with electrophoresis, chromatographic methods, and bioinformatics tools is currently making significant contributions to a better knowledge of the composition and structure of the cereal proteins and their structure–function relationships. Results obtained using mass spectrometry, including characterization of prolamins, investigation of the gluten toxicity for coeliac patients, identification of proteins responsible of cereal allergies, determination of the protein pattern and its modification under environmental or stress effects, investigation of genetically modified varieties by proteomic approaches, are summarized here, to illustrate current trends, analytical troubles and challenges, and suggest possible future perspectives. © 2011 Wiley Periodicals, Inc. Mass Spec Rev 31:448–465, 2012

Keywords: *cereal proteins; mass spectrometry; proteomics*

I. INTRODUCTION

Cereals, including rice, barley, and wheat, are the major crops for the global food supply, dominating world agriculture. Particularly, wheat, due to its adaptability to a wide range of environments and for the unique functional properties of its flour, represents the most widely grown, processed, and consumed cereal by humankind.

Functional and nutritional properties of cereals depend largely from their protein pattern, which represents the target of a considerable number of investigations. The recent research advances at the gene level have resulted in the complete sequence of the genome of a variety of organisms. The relatively small size of rice genome (*Oryza sativa* L., 450 Mb) and the importance of this crop as most important food for half of the world's population, led to establish great sequencing project, which resulted in the complete sequencing of the genome of rice (Goff et al., 2002; Yu et al., 2002). Due to the larger genome size, sequencing of the genomes of the other

cereals is still in progress. Therefore, for wheat (*Triticum aestivum*, 16.5 Gb, five times bigger than human genome), oat (*Avena sativa*, 11.4 Gb), rye (*Secale cereale*, 8.4 Gb), barley (*Hordeum vulgare*, 4.5 Gb), and maize (*Zea mays*, 2.5 Gb) efforts are being made to describe the cereals transcriptome by sequencing the cDNAs, generating expressed sequence tags (ESTs), instead of sequencing the whole genome (Matthews et al., 2003).

The improvements in analytical technology, together with the partial genome information and the availability of a large amount of ESTs gene indices, have determined a rapid advancing of plant proteomics studies, whose purpose has moved beyond the large-scale protein identification to the study of functional and regulatory aspects of proteins (i.e., comparative seed protein expression from different cultivars, post-translational modifications (PTMs), protein–protein interaction, etc.) (Skylas et al., 2001; Patterson & Aebersold, 2003; Mak et al., 2006; Agrawal, Jwa, & Rakwal, 2009; Finnie & Svensson, 2009; Finnie et al., 2009; Merlino et al., 2009; Nakagami et al., 2010). The reasonable interest in cereal proteomics goes far away from the elucidation of structure–functions relationships because of the strict connection with the needs for food, feed, and industry. These studies are supported by the ideas of getting insight into the different varieties and their quality related to stress, pathogens, yield, with the final aim of developing safe and high quality food and reducing the impact of agriculture on the environment.

Due to the high complexity and the dynamic nature of the cereals proteome, it is generally accepted that a single analytical approach for protein separation will not supply comprehensive exploration of the entire proteome. Therefore, it is often needed to use multiple technologies to improve proteome resolution and coverage, and to provide complementary results (Whitelegge, 2004). Currently there are two complementary proteomic approaches: the so-called “gel-based” approach, and the “gel-free” approach. The gel-based approach represents the cornerstone of proteome analysis and uses the unequalled resolving power for separation of complex protein mixtures provided by two-dimensional gel electrophoresis (2DE). Instead, recent developed “gel-free” approaches avoid the 2DE separation employing chromatographic technologies. Both approaches differ with respect to the extraction, separation, and detection of proteins and peptides. As a consequence of the complex chemical nature of proteins and to the large dynamic range of their concentrations, each approach is able to focus only a subfraction of the protein set. In any case, both gel-based and free-gel approaches are always followed by mass spectrometry analysis, which is the primary tool for

*Correspondence to: Salvatore Foti, Dipartimento di Scienze Chimiche, Università degli Studi di Catania, Viale A. Doria, 6, I-95125 Catania, Italy. E-mail: sfoti@dipchi.unict.it

identification and characterization of proteins and plays a fundamental role in proteomic investigation. Here, current applications of mass spectrometry-based methodologies in the proteome analysis of mature cereal kernels will be reported, with particular emphasis on the characterizations of prolamins. Examples of applications will include structural characterization of isolated proteins, characterization of composition of complex protein fractions, investigations of stress-related variation of protein expression, comparative analysis of proteome in genetically modified (GM) and parent cereals and investigations of allergens in wheat and other cereals. This overview will be preceded by a brief description of the classification of cereal prolamins and a brief review of the current mass spectrometry-based methodologies.

II. CLASSIFICATION OF CEREAL PROTEINS

A range of criteria has been used to define and classify cereal proteins. Pioneer works led to the concept of classifying cereal proteins according to their solubility, and this was formalized by T.B. Osborne working on the proteins from over 30 plant species in the late 19th/early 20th century (Osborne, 1924). Osborne classified proteins into four groups (often called “Osborne fractions”) based on their solubility in water (albumins), dilute salt solution (globulins), alcohol/water mixtures (prolamins), and dilute acid (glutenins). The terms “albumins” and “globulins” are also more widely used for proteins with similar solubility properties from other organisms while the prolamins are named specifically in different cereals: gliadins in wheat, hordeins in barley, secalins in rye, and zeins in maize. Although the Osborne classification is still used, the availability of amino acid sequences and three-dimensional structures of cereal proteins led most scientists to consider that it is more valid a classification of these proteins based on their functions, structure, or evolutionary relationships. On this respect, proteins can be grouped together into families if they have residue identities of 30% or greater or if they have lower sequence identities but their functions and structures are very similar. Families whose members have low sequence identities but whose structures and functional properties suggest a probable common evolutionary origin are placed together in superfamilies (Murzin et al., 1995; Lo Conte et al., 2002). Consequently, cereal proteins can be divided into three broad groups based on their function: structural and metabolic proteins; protective proteins; storage proteins. Structural and metabolic proteins include a wide number of proteins which contribute to cell, structure, and function, such as structural components of cell walls, enzymes, and proteins involved in transport, growth, signal transduction, gene transcription, and protein synthesis. This group comprises most of the proteins belonging to the two “Osborne fractions” called albumins and globulins. The group of protective proteins comprises a wide range of components involved in defence mechanism of the plant against pathogenic micro-organisms and/or insects (Shewry & Lucas, 1997). Finally, the functionally related group of storage proteins comprises components which are mostly deposited in seeds where they are accumulated during seed maturation. This group of proteins, which provides a store of amino acids and carbon skeletons for germination and seedling growth, represents more than 50% of the total protein

content in the mature cereal seeds, except for rice and oats, which have high proportion of glutenin and albumin storage proteins, respectively. However, today the most used classification subdivides cereal proteins into families/superfamilies. The most important are the superfamily of prolamins and the superfamily of cupins that comprise a number of important allergens (Tatham & Shewry, 2008).

A. The Prolamin Superfamily

The existence of the prolamin superfamily was first proposed in 1985 (Kreis et al., 1985) based on visual comparison of amino acid sequences of various proteins, which allowed the identification of conserved amino acid residues. Visual comparison, subsequently supported by more sophisticated computational methods, showed a characteristic conserved pattern of eight cysteine residues which can be defined by the formula: C-X_n-C-X_n-CC-X_n-CXC-X_n-C-X_n-C. Although some variation in this cysteine pattern does occur, the pattern is sufficiently well conserved to identify further members of this superfamily showing little or no other sequence identity.

The vast majority of the characterized members of the prolamin superfamily are seed proteins, which can be broadly divided into two types: the low-molecular-mass sulfur-rich seed proteins and the prolamins themselves.

The proteins belonging to the first type have molecular mass of about 7–16 kDa, eight or more conserved cysteine residues and show similar three-dimensional structures that are rich in α -helices. This group includes three major classes of cereal food allergens, 2S albumins, non-specific lipid transfer proteins (nsLTPs) and cereal α -amylase/trypsin inhibitors (Breiteneder & Radauer, 2004; Mills et al., 2004). The 2S albumins are a major group of storage proteins present in many dicotyledonous seeds (Shewry & Pandya, 1999). Typical 2S albumins are heterodimeric proteins that consist of two polypeptide chains of approximately 4 and 9 kDa linked together by four disulfide bonds (Shewry, Napier, & Tatham, 1995). They include major allergens from tree nuts and seeds such as Brazil nut, walnut, sesame, and mustard. The nsLTPs have a wide distribution in fruits, nuts, seeds, and vegetables and comprise 7–9 kDa monomeric proteins that are held together by four disulfide bonds to form a hydrophobic tunnel in their three-dimensional structures (Shin et al., 1995). Non-specific lipid transfer proteins play an important role in the defence mechanism of the plant against pathogenic micro-organisms and/or insects. Cereal α -amylase and protease inhibitors protect the plant against insect pests that feed on plant tissues (Carbonero & García-Olmedo, 1999). They have subunits with masses ranging from about 12 to 16 kDa, contain four disulfide bonds and can exist as monomers, dimers, or tetramers. Allergenic members within the cereals inhibitors include the most important wheat proteins contributing to bakers' asthma (Salcedo et al., 2004), which are often called chloroform/methanol or “CM” proteins as they were initially defined based on their selective extraction in chloroform/methanol mixtures.

The prolamins themselves are the second major type of proteins classified in the superfamily. They show an immense variability in their structures and properties, as a consequence of the taxonomic relationships of the cereals (Fig. 1). The prolamins of rice (*Bambusoideae*) are not closely related to the

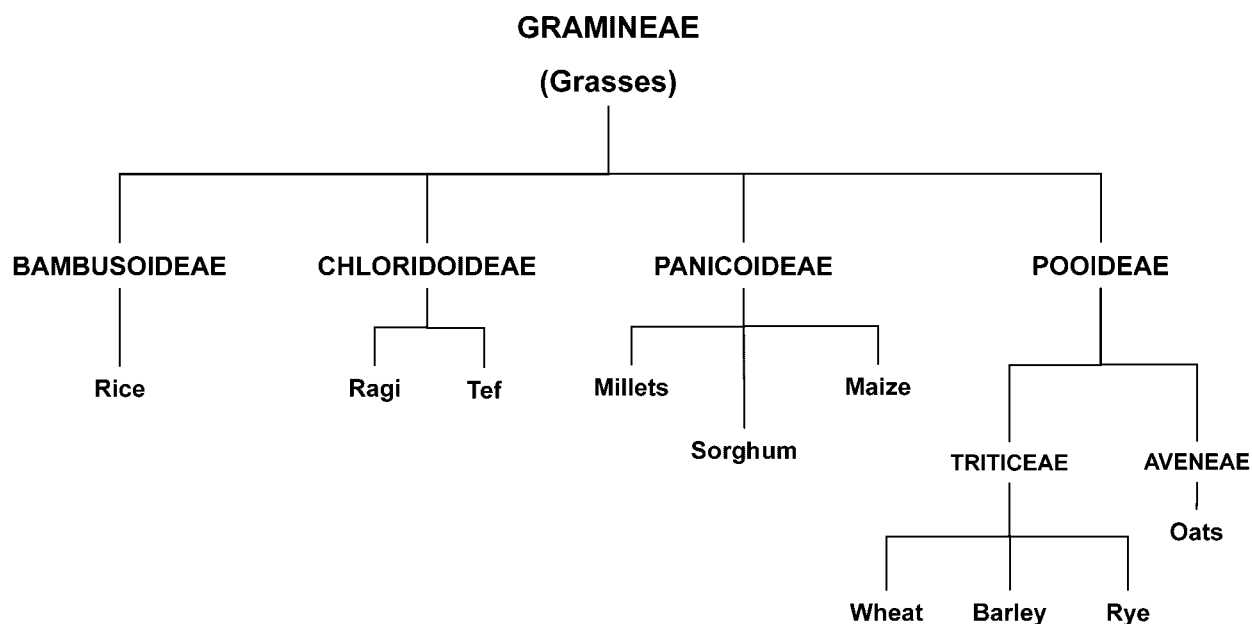


FIGURE 1. Taxonomic tree of Gramineae.

other cereals, whereas the prolamins of the *Panicoideae* and *Chloridoideae* appear to be closely related to each other. The prolamins of the *Triticeae* are closely related with homologous groups of proteins present in all three species and less to the prolamins of oats. Nevertheless, most prolamins share a common structural feature consisting in the presence of one or more domains formed by repeated blocks usually based on short (up to 10 amino acids) peptide motifs, which are rich of glutamine and other amino acids such as proline, glycine, and phenylalanine.

Classification based to structural and evolutionary relationships (Shewry & Tatham, 1990) assigns all of the prolamins of *Triticeae* (wheat, barley, and rye) to three broad groups: sulfur-rich (S-rich), sulfur-poor (S-poor), and High Molecular Weight (HMW) prolamins, with several subgroups within the S-rich group (Fig. 2).

In wheat, the S-rich group of prolamins comprises the monomeric α/β - and γ -type gliadins, and the polymeric B- and C-type Low Molecular Weight Glutenin Subunits (LMW-GSs). The α/β - and γ -type gliadins have masses ranging from about 30 to 40 kDa and present six or eight cysteine residues linked to form three or four intra-chain disulfide bonds, respectively (Shewry & Tatham, 1997).

The B- and C-type LMW-GSs are controlled by genes at the *Glu-A3* and *Glu-B3* loci on the short arms of chromosome 1A and 1B, respectively. The B- and C-type LMW-GSs, on the basis of their SDS-PAGE mobilities, have apparent molecular masses of 42–51 and 30–40 kDa, respectively (Payne et al., 1985). Direct *N*-terminal sequencing of isolated proteins and deduced sequences of several LMW-GS genes (Shewry & Tatham, 1997), showed that typical LMW-GS sequences can be classified as LMW-s, LMW-m, and LMW-i types, on the basis of the first amino acid of the mature polypeptide (serine, methionine, or isoleucine, respectively). The B-type LMW-GSs have an even number of cysteine residues, with two cysteine residues available for forming intermolecular

disulfide bonds (Lindsay & Skerritt, 1999) and therefore they are thought to act as chain extenders in the glutenin macropolymer. The *N*-terminal sequences of all except one of the C-type subunits are not closely related to the B-type subunits, but correspond to modified α/β - and γ -type gliadins, which may have been incorporated into the glutenin macropolymer because of mutations in the number and/or position of Cys residues (D'Ovidio & Masci, 2004). In fact, in addition to the cysteine residues present in the related α/β - and γ -gliadins, most of the C-type subunits contain an extra and unpaired cysteine residue that is presumably employed to form an inter-chain disulfide bond. Consequently, they should act as chain terminators in the gluten macropolymer.

The group of wheat S-poor proteins includes the monomeric ω -gliadins and the polymeric D-type LMW-GSs. The monomeric ω -gliadins, with masses ranging from about 55 to 80 kDa, consist almost entirely of repeats and, in contrast to the α/β - and γ -type gliadins, are characterized by a low content of sulfur-containing amino acid residues and a lack of cysteine residues.

The D-type LMW-GSs, encoded by *Glu-3* locus on the short arms of chromosome 1D, have an apparent molecular weight of 55–80 kDa and seem to be formed by mutation in one or more genes encoding ω -gliadins, which resulted in the presence of a single cysteine residue that allows inter-chain disulfide bond formation in the gluten macropolymer (Masci et al., 1993).

The third group of wheat grain prolamins is constituted by the HMW Glutenin Subunits which are encoded by genes (*Glu-1* loci) on the long arms of the group 1 chromosomes (1A, 1B, and 1D) (Payne & Lawrence, 1983; Shewry, Halford, & Tatham, 1989). Each locus comprises two genes linked together encoding two different types of HMW-GSs, x- and y-type (Payne, Corfield, & Blackman, 1981; Payne et al., 1987; Shewry, Halford, & Tatham, 1992). The x-type subunits generally have a slower electrophoretic mobility in

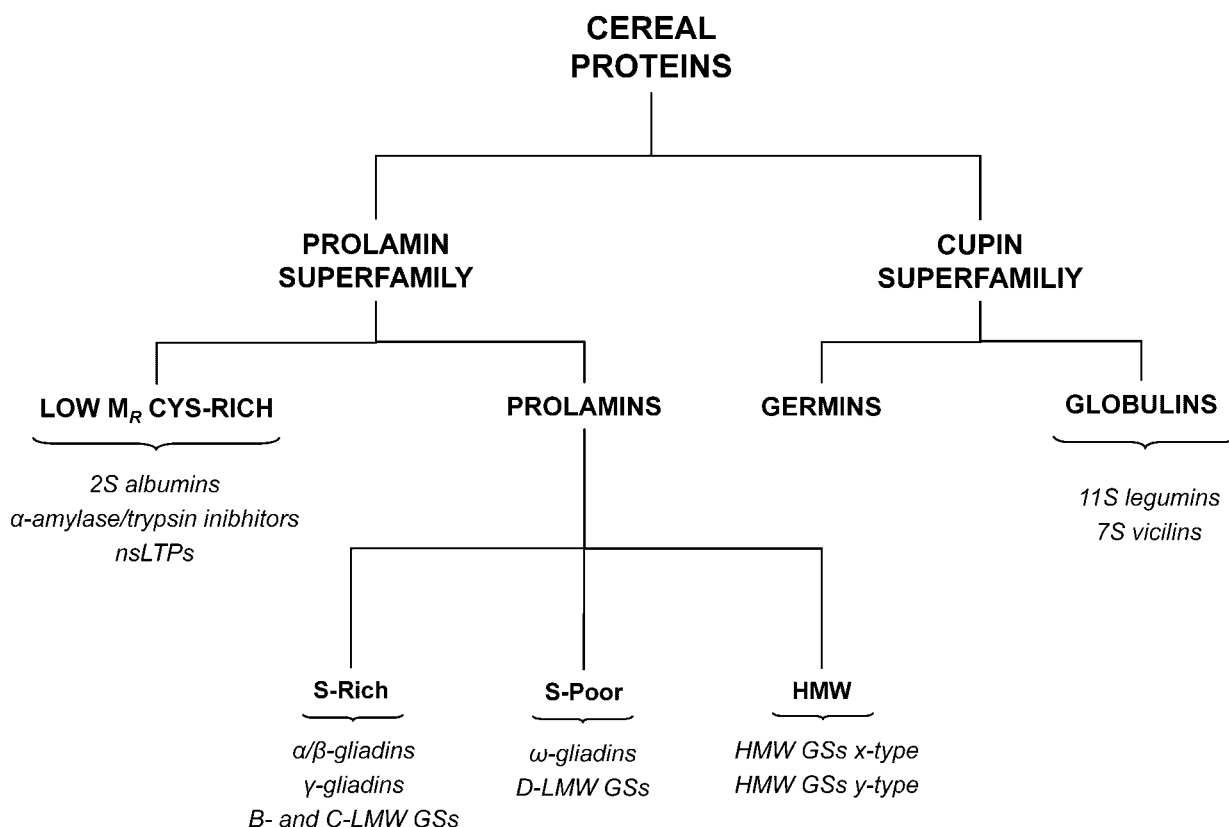


FIGURE 2. Classification of wheat proteins based to their structural and evolutionary relationships.

SDS-PAGE and higher molecular weight (83–88 kDa) than the y-type subunits (67–74 kDa). Bread wheats could, in theory, contain six different HMW-GSs but, due to the “silencing” of some of these genes, most common wheat cultivars possess three to five HMW-GSs. Thus, all hexaploid wheats contain at least the 1Bx, 1Dx, and 1Dy subunits, while some cultivars also contain a 1By subunit and a 1Ax subunit as well. As a consequence, it appears that the gene encoding the 1Ay subunit is always silent. Nevertheless, 1Ay subunits have been reported in the A-genome diploid species *T. monococcum* and *urartu* (Waines & Payne, 1987) and some bread wheats with six HMW-GSs have been reported (Johansson et al., 1993; Margiotta et al., 1996). The current nomenclature of HMW-GSs was developed by Payne & Lawrence (1983). It provides a chromosomal location of the genes, the subunit type (x- or y-type) and finally a number indicating the mobility of each subunit on SDS-PAGE gels (i.e., Bx7, By9).

HMW-GSs have a similar three domains structure, with non-repetitive domains at the N-terminus (81–105 residues) and C-terminus (42 residues), which are characterized by the presence of most or all of cysteines and flanking a large repetitive domain (640–830 residues) comprising tandem and interspersed repeats of two or three peptide motifs. Almost all x-type subunits have four cysteine residues, except subunit 1Dx5, which has five. All y-type subunits present seven cysteine residues, three of which are presumably available for inter-chain disulfide bonds (Shewry, Halford, & Tatham,

1992). Therefore, y-type HMW-GSs, together with subunits 1Dx5, are thought to have a branching effect in the gluten macropolymer (Payne & Corfield, 1979).

B. The Cupin Superfamily

The cupins are a functionally diverse superfamily of proteins having very low levels of sequence identities but with highly conserved structures. Cupins share two short conserved consensus sequence motifs and a common β -barrel structural core domain to which the term cupin (Latin *cupa*, a barrel) was given (Dunwell, 1998). A broad division can be made between proteins with one or two cupin domains, the latter being termed *bicupins*. The best-characterized cupins from plants are germins and storage globulins, which are single-chain cupins and bicupins, respectively.

Germins constitute a group of homologous proteins containing a characteristic sequence (PHIHPRATEI) known as the germin box (Lane et al., 1991). The name germin was given because the first described member of the family, wheat germin, was discovered in germinating wheat grains. However, it is now known that proteins from this family exist in all organs and developmental stages. Germins are involved in the response to various stress conditions, including resistance against infection, wilt, salt, or heat shock.

The second group of cupins is constituted by the storage globulins. On the basis of their sedimentation coefficient, the globulins can be divided into 7S vicilin-type globulins and the

11S legumin-type globulins (Lawrence, 1999). Mature 7S globulins are trimeric proteins (M_r 150–200 kDa) constituted by subunits with masses ranging from about 40 to 80 kDa and lacking of cysteine residues. Mature 11S globulins are often called legumins as they are characteristic of legume seeds. Typical legumins are hexameric proteins (M_r 300–400 kDa) constitute by six subunits of M_r about 60 kDa.

III. CURRENT MASS SPECTROMETRIC METHODOLOGIES ON PROTEINS ANALYSIS

The development in the late 1980s of the two “soft” desorption/ionization MS techniques ESI (Fenn et al., 1989) and MALDI (Karas & Hillenkamp, 1988; Tanaka et al., 1988) which are capable to produce molecular ions of intact biomolecules, has revolutionized the field of protein investigation, making polypeptides accessible to mass spectrometric analysis.

In the last decade, the improved performance and versatility of the mass spectrometers, together with the increasing availability of gene and genomic sequence database, led mass spectrometry to become an indispensable tool for either protein and proteome analyses. On this respect, the mass analyzer plays a fundamental role in the mass spectrometer technology, together with its key parameters such as sensitivity, resolution, mass accuracy, and the ability to generate MS/MS spectra. There are four basic kinds of mass analyzer currently used in protein analysis: ion trap (IT), time-of-flight (TOF), quadrupole (Q), and Fourier transform ion cyclotron resonance (FTICR) analyzer. These mass analyzers differ considerably in design and performance, each with its own strength and weakness, and can be stand alone or put together in tandem to take advantage of the strengths of each. Three-dimensional ion trap, frequently coupled with ESI source, represents the unique analyzer which alone allows to determine either the mass of a given peptide and its sequence. Three-dimensional ion trap is robust, sensitive, and moderately inexpensive, but shows a relatively low mass accuracy (Jonscher & Yates, 1997). The recent development of the “linear” or “two-dimensional ion trap” allowed to increase sensitivity, resolution, and mass accuracy (Hager, 2002; Schwartz, Senko, & Syka, 2002). Conversely, TOF analyzers, usually coupled with MALDI source, have high sensitivity, resolution, and mass accuracy, but do not permit to perform true MS/MS. However, the development of TOF/TOF (Medzihradszky et al., 2000; Bienvenut et al., 2002) and hybrid Q/TOF (Loboda et al., 2000; Chernushevich, Loboda, & Thomson, 2001) instruments, which can be used interchangeably with a MALDI and an ESI source, have overcome this disadvantage. The resulting fragment ion spectra of both TOF/TOF and Q/TOF instruments are often more extensive and informative than those generated by trapping instruments. The FTICR, like the ion trap, is a trapping analyzer, although it captures the ions under high vacuum in a high magnetic field. This type of analyzer is the most powerful currently available, particularly in terms of mass measurement accuracy and mass resolution. Mass measurement accuracy of these analyzers is approximately 1 ppm, surpassing ion traps by at least two orders of magnitude, and resolution can be routinely 100,000. Fourier transform ion cyclotron resonance analyzers are highly versatile and can be combined with both MALDI and ESI ion sources (Heeren

et al., 2004 and reference therein). However, in spite of the enormous potential, the expense, operational complexity, and low peptide-fragmentation efficiency of FT-MS instruments have limited their routine use in protein research. On the other hand, a recent technical development which joins a linear ion trap (LTQ) with a FTICR analyzer has overcome this problem. The hybrid LTQ-FT setup combine the tandem mass spectrometry capabilities and the fast scan times of the linear ion trap with the already reported capability of the FTICR (Syka et al., 2004a; Parks et al., 2007). Furthermore, the development of the new fragmentation techniques of electron-capture dissociation (ECD) (Zubarev, Kelleher, & McLafferty, 1998; Zubarev, 2006) and electron-transfer dissociation (ETD) (Syka et al., 2004b) has greatly improved the dissociation efficiency, resulting in more extensive fragmentation and better sequence coverage, including localizing PTMs, than the classical collision-induced dissociation (CID) technique (Shukla & Futrell, 2000).

On the light of the improving performance and versatility of the mass spectrometers, nowadays it is possible to measure the molecular mass of proteins and peptides with high accuracy and to determine additional structural features which include the primary structure or PTMs. Additionally, MS is frequently used in large-scale proteome research in order to provide high-throughput protein identification, protein expression, and protein–protein interaction. Protein identification is performed using the mass spectral data set via database searching with appropriate software programs such as Mascot, ProFound, MS-Fit (Aebersold & Mann, 2003; Lin, Tabb, & Yates, 2003), and similar. Moreover, proteomic experiments invoke the simultaneous analysis of several thousands of protein species and require a separation step (including gel electrophoresis and liquid chromatography-based methods) in order to simplify the complex mixtures of peptides and proteins that are delivered to the mass spectrometer. The choice of the separation methods used depends on the experimental question; so that, two complementary approaches are currently used in protein and proteomic researches: the so-called “bottom-up” and “top down” approaches (Kelleher et al., 1999; Reid & McLuckey, 2002). “Bottom-up” strategies, in which peptide detection is used to infer the protein presence, are the standard for large-scale or high-throughput analysis of highly complex samples. These are usually carried out by two different ways. In the first one, proteins in complex mixtures are firstly separated by 2D electrophoresis; then, the isolated proteins are subjected to enzymatic (or chemical) digestion followed by direct peptide mass fingerprinting-based acquisition or further peptide separation on-line coupled to tandem mass spectrometry. In the second way, the protein mixture is directly digested into a collection of peptides, which are separated by multidimensional chromatography on-line coupled to tandem mass spectrometric analysis (Fournier et al., 2007). The latter method, commonly referred to as “shotgun” approach or Multidimensional Protein Identification Technology (MudPIT) (Washburn, Wolters, & Yates, 2001) has gained popularity in the proteomics field (McDonald & Yates, 2003; Delahunty & Yates, 2007). However, although the “shotgun” approach is conceptually simple, it results in greatly increased complexity of generated peptide mixture. As a consequence, not all peptides resulting from the digestion of a protein can be observed or correctly identified by MS analysis. Moreover,

only the peptides present at high relative abundance are preferentially sampled, while information regarding the proteins represented as low abundance peptides in the complex mixture is commonly not obtained.

In the “top down” strategy, proteins in complex mixtures are firstly fractionated and separated into pure single proteins or less complex protein mixture. This step is followed by off-line static infusion of sample into the mass spectrometer for high-resolution mass measurement of intact protein ions and their direct fragmentation inside the instrument without prior digestion (McLafferty et al., 2007). In theory, this approach, thanks to the use of ECD fragmentation technique, places the entire sequence at least for small to medium-sized proteins (Ge et al., 2002; Zabrouskov & Whitelegge, 2007). Therefore, if the bottom-up method is widely accepted for the routine identification of proteins in complex mixture, when more extensive or specific data are needed, such as on polymorphisms or PTMs, the complementary “top-down” approach can often provide these in a very straightforward manner. The far higher masses measured with the top-down approach require correspondingly higher MS resolving power, so the instrument of choice has been the expensive Fourier transform mass spectrometer (FT MS) (Patrie et al., 2004 and reference therein). Of special promise for routine top-down applications is the recently developed Orbitrap mass spectrometer, equipped with a new type of mass analyzer, the Orbitrap, which was invented by Makarov in 1999 (Makarov, 2000) and was reported as a tool for proteomics research in 2005 (Hu et al., 2005). Moreover, the new hybrid LTQ-Orbitrap instrument has recently showed resolution and mass accuracy capabilities approaching those of FT MS, with very promising cost advantages (Makarov et al., 2006) and therefore it could open a new era for the mass spectrometry.

IV. CHARACTERIZATION OF PROLAMINS

A. Investigation of High Molecular Weight Glutenin Subunits (HMW-GSs)

HMW-GSs are minor components of gluten fraction in terms of quantity, but they are key factors in the process of bread-making because they are the major determinants of gluten elasticity, promoting the formation of larger glutenin polymers. The early mass spectrometric works on prolamins were focused on the determination of the molecular masses of HMW isolated as single subunits (Hickman et al., 1995) or as unfractionated HMW extracts (Garozzo et al., 1999). The studies on the identification of HMW-GSs masses by a single MALDI mass spectrum was motivated by the anomalous migration of these subunits in SDS-PAGE, attributed to conformational effects on mobility, which results in apparent molecular masses higher than those calculated from gene sequencing (Bunce, White, & Shewry, 1985). The MALDI mass spectrum of a crude extract with the complete HMW-GSs pattern profile has been proposed as a routine method in breeding programs for rapid identification of lines containing subunits related to quality and for varieties identification (Dworschak et al., 1998). Some of the HMW-GSs investigated were purified from the same cultivars that had been used to isolate and sequence the corresponding genes, whereas other subunits were obtained from cultivars different from those used to

sequence the genes (Hickman et al., 1995). The determined molecular masses were in most of the case consistent, within the range of analytical error, with those calculated from the gene sequences, with the exception of some subunits, which gave experimentally determined molecular masses differing than the expected values (Cunsolo, Foti, & Saletti, 2004b).

For a number of HMW-GSs, whose gene derived sequence were known, verification of the sequence correctness and assessment of the presence or absence of PTMs, has been obtained by tryptic digestion or cyanogen bromide cleavage of the proteins and subsequent analysis of the proteolytic mixtures by mass spectrometry. It should be highlighted that, the presence of large repetitive domain together with the low content of Lys and Arg amino acids residues along the sequence of the HMW-GSs determine the production of high molecular mass tryptic peptides. Consequently, a reliable sequence mapping of these subunits can be achieved coupling the RP-HPLC/ESI-MS/MS data with MALDI spectra acquired both at low- (m/z 600–4,000) and high-molecular range (m/z 4,000–30,000) (Foti et al., 2000; Cozzolino et al., 2001a,b; Cunsolo et al., 2002, 2003, 2004a). By combining data on the protein regions mapped by MALDI-TOF MS and LC-MS/MS, a coverage >95% was obtained for all the subunits investigated and the gene-derived sequences were generally confirmed, although some differences consisting of point modifications, insertion, or deletion of short sequences in the repetitive domain were detected. Moreover, the results obtained have highlighted the complete absence of glycosylation or other PTMs, thus clarifying a major controversial point in the structural characterization of these proteins. In a recent work, this analytical approach allowed the identification of a truncated form of the subunit By8 (Mamone et al., 2009). In particular, the comparison of the LC-MS/MS profile and both low- and high-mass range MALDI spectral profiles of the tryptic digest of this truncated By8 form with those obtained from the full-length subunit allowed to conclude that the truncated protein is lacking of the C-terminal region. The presence of this truncated form was explained with either endogenous proteolytic cleavage of the full-length protein or by gene deletion events occurring in the HMW-By8 allele.

On the light of the currently increasing knowledge of the primary structure of different HMW-GSs, which show a strong sequence similarity, an upgrade of their naming system, which could be no more related to their electrophoretic mobility, but on their amino acid sequence similarity, could be envisaged.

B. Investigation of Low Molecular Weight Glutenin Subunits (LMW-GSs)

LMW gluten subunits, which in a single cultivar account for a large number of components in contrast with the low number of HMW-GSs, have been less studied because of the complexity of this fraction. Direct MALDI-MS analysis of unfractionated LMW extracts usually shows a complex pattern of proteins in the 30–40 kDa range. The observed pattern may be suitable only for differentiation between wheat varieties, but the complexity of the mass spectra precludes the use of MALDI-MS as a stand-alone technique for the identification of individual components (Dworschak et al., 1998). On this respect, MALDI-TOF mass spectrometry has been often

employed after the chromatographic isolation of LMW-GSs extracts, as a tool for the determination of the molecular mass of new subunits (Masci et al., 1995; An et al., 2006) and of the number of Cys residues (Masci et al., 1998).

When proteomic techniques have been applied to the investigation of this peculiar class of storage proteins, it has become apparent that the characterization of LMW-GS extracts is very difficult and challenging. Indeed, these subunits are encoded by a high number of genes with similar structural characteristics. Most of these genes have not been isolated and characterized yet, making difficult the application of standard proteomic techniques and the determination of a direct correspondence between a gene sequence and its encoded protein. Furthermore, it should be noted that in LMW-GSs characterization, even if cleavage enzymes such as chymotrypsin and thermolysin have been tested (Vensel et al., 2007), the use of trypsin remains widespread. Therefore, as the amino acid sequences of LMW-GSs show a long repeating motif and a scarcity of cleavable tryptic sites, the number of the resulting tryptic peptides suitable for MS analysis is scant. As a consequence, MALDI mass spectra obtained from the tryptic digests of all LMW-GSs present very few signals and all the attempts to identify proteins by database search using the experimental m/z values are obviously unsuccessful. A possible way to identify LMW-GSs is to perform MS/MS experiments on the few available tryptic peptides and try to interpret the results obtained in the light of the peculiar primary structure of this class of proteins. Specifically, it should be considered that: (1) sequence coverage will obviously be scarce owing to the long repeating motif lacking tryptic cleavage sites; (2) the few peptides with masses suitable for MS/MS fragmentation (occurring in the *N*-terminal and *C*-terminal domains) would not be able to discriminate the single subunits because they are common to groups of LMW-GSs; (3) a single 2D-gel spot may frequently contain more proteins belonging to the same group of LMW-GS, differing by point substitutions.

Despite these intrinsic limitations, several studies on LMW-GSs by both gel-free and gel-based methodologies have been performed. A gel-free investigation of the B- and C-type LMW-GSs present in hexaploid wheats (Chinese Spring cv., *T. aestivum*) has demonstrated that HPLC coupled on-line with ESI MS or off-line with MALDI-TOF MS represents efficient tools for obtaining complementary data for the characterization of the gluten complex (Muccilli et al., 2005). Particularly, comparative analysis of the alkylated and unalkylated B- and C-type LMW-GSs fractions allowed the detection of 150 different components together with a reliable determination of the number of cysteine residues present in 42 proteins, which were tentatively classified based on their M_r and number of cysteine residues. An interesting information provided by these results was the existence of an extensive micro-heterogeneity of LMW-GSs, which was highlighted by the detection of several components with minor differences in the relative molecular masses co-eluting in the same HPLC peak, indicating a complexity higher than that resulting from the genomic analysis. Similarly, an off-line combination of HPLC, MALDI-MS, tryptic digestion, MS/MS analysis, and database searching was performed on the entire gluten extract of the Teal variety of Canadian hard red spring wheat (Qian et al., 2008). The data obtained by this approach demonstrated that

while the HMW-GSs were easily identified by means of their intact masses and their tryptic fragments, according to the published cDNA sequences, the identification of the LMW was limited to some group of subunits for two reasons: (1) several subunits have intact molecular masses too similar to be resolved by linear MALDI-TOF MS; (2) the DNA sequences of several LMW subunits show very strong similarities (Ikeda et al., 2002), and therefore most of the resulting tryptic peptides are identical. As a result, the data obtained do not provide definitive evidence that all of the expected subunits are effectively present.

On the other hand, all the proteomics studies performed on LMW-GSs by means of gel-based techniques encompass both the problems of achieving good electrophoretic separation for a class of proteins with almost similar electric charge and molecular masses, and the previously reported difficulties of obtaining protein identification. The aim of these works is not restricted to cultivars characterization (Mamone et al., 2009), but also includes efforts to better understand the role of LMW-GSs in gluten matrix (Lindsay & Skerritt, 1999) and seed maturation by means of comparison between common wheat cultivars and translocated or transgenic ones (Scossa et al., 2008).

Recent works on LMW-GSs have focused the attention on the not fully understood role of these ones in the gluten macropolymer and consequently on flour and semolina functionality by means of comparison of the extracts of a common wheat cultivar and a line carrying a 1BL.1RS translocation (Gobaa et al., 2007; Muccilli et al., 2010). Particularly, in *T. aestivum*, the replacement of the 1B short-arm chromosome by the 1R short-arm chromosome of rye was originally designed to achieve resistance to diseases such as leaf rust, stem rust, stripe rust, and powdery (Zeller & Hsam, 1983). This replacement has as consequence the substitution of the LMW and gliadin components associated at the *Glu-B3/Gli-B1* loci with monomeric ω - and 40 K γ -secalins, which are encoded by closely linked genes present at the *Sec-1* locus on the short arm of the chromosome 1R (Carrillo, Vazquez, & Orellana, 1990). On the contrary, no proteins homologous to the wheat LMW-GSs are present on this chromosome arm. Particularly, the comparative proteomic analysis, by means of 2D-PAGE, in-gel digestion, and MS/MS analysis, on the LMW fraction of the durum wheat cultivar Svevo and of its derived 1BL.1RS translocated line (Muccilli et al., 2010) allowed to establish the effects of the replacement of *Glu-B3* locus with the *Sec-1* locus on the short arm of the chromosome 1R. The peptide fragment fingerprint (PFF) approach, carried out by combining database searching of the MS/MS data and manual interpretation of tandem mass spectra, permitted to detect and characterize almost all the *N*-terminal tryptic peptides of the B-type LMW-GSs identified and therefore to conclude that all the three types of typical LMW-GS (LMW-s, LMW-m, and LMW-i) are present in Svevo whereas the latter group is the only one present when the *Glu-B3* locus is replaced by the short arm of the chromosome 1R. In addition, some α/β - and γ -gliadins were apparently identified by proteomic analysis, in both Svevo and its translocated line, although if they were not expected to be present in the fraction investigated, due to the extraction procedure employed which eliminates the monomeric gliadins in the preceding water/alcohol extraction step. It was concluded that it is reasonable to

assume that the proteins identified as gliadins were actually C-type LMW-GSs, corresponding to α/β - and γ -gliadins modified in the number of cysteine residues (Masci et al., 2002). On the other hand, a comparative proteomic analysis performed on 16 doubled haploid lines, with or without the 1BL.1RS translocation, revealed quantitative and qualitative protein variations in prolamins and other endosperm proteins (Gobaa et al., 2007). Of particular interest was the identification of a γ -gliadin carrying nine cysteine residues which was highly over-expressed in the 1BL.1RS translocated genotypes. This result suggests that the lack of LMW-GSs is counterbalanced by an over-expression of a relatively similar prolamin. Indeed, the finding of a gliadin with an odd number of cysteines suggests that at least one residue is not involved in an intra-chain disulfide bond, but it is available for covalent bonding to other polymeric glutenin subunits. The up-regulation of a γ -gliadin carrying nine cysteine residues upon the loss of an LMW-GS locus would mean that the glutenin polymerization is a regulated process. This process controls the amount of polymeric gluten subunits in order to maintain a certain level of polymerization in response to chromosomal rearrangement at their encoding loci. Better understanding of these mechanisms may lead to further improvements in the rheology of the dough and in the quality of wheat.

C. Gliadins, Hordeins, Secalins, and Celiac Disease

The molecular characterization of gliadins by mass spectrometry is a research topic of increasing interest since, due to their structural characteristics, these subunits take part in the development of the gluten macropolymer by means of inter-molecular hydrogen bonds (Gln residues) and hydrophobic interactions between non-polar amino acid side-chains. On the other hand, investigation of naturally mutated gliadins showed that they can be covalently involved in the gluten macropolymer. In fact, a proteomic investigation (Ferrante et al., 2006) demonstrated that an *in vivo* expressed novel γ -gliadin, containing an extra cysteine residue, was incorporated into the polymeric glutenin fraction, influencing negatively the dough-making performance.

On the other hand, proteomic approaches applied to the characterization of these monomeric subunits, and used as complement for studies at gene level (Dumur et al., 2004; DuPont et al., 2004), reveal the same limitations previously reported on the characterization of LMW glutenin subunits, including the complexity of the protein mixture and the very low number of the resulting tryptic peptides suitable for MS analysis. To overcome the latter limitation chymotrypsin was used as alternative enzyme (Mamone et al., 2005). Chymotrypsin having a broader specificity than trypsin, provided an efficient gliadin spot digestion with a greater number of medium-sized peptides suitable for MS and MS/MS analysis. However, it should be noted that, due to the intrinsic structural characteristics of these proteins, MALDI-TOF mass spectra of the chymotryptic peptides showed a complex profile and database search via peptide mass fingerprint (PMF) approach did not allowed any protein identification. Therefore, in order to overcome this drawback, the more exhaustive MS/MS approach was necessary. By coupling database searching and *de novo* sequencing of the MS/MS data of the chymotryptic peptides in each spot, sequences common to all isoforms belonging to

a given gliadin subfraction can be detected, as well as sequences differentiating the various isoforms of the same proteins. Moreover, *de novo* sequencing of different chymotryptic fragments allowed the characterization of sequence stretches carrying amino acid substitutions with respect to that present in the database, which could justify the slight differences in the coordinates of gliadins along the 2DE gel map. In this way, a substantial amount of information can be obtained on the complex heterogeneity of gliadins, providing an efficient tool to distinguish wheat varieties at molecular level.

Recently, a comprehensive work on wheat gluten proteins identification was performed in order to evaluate the effects of the enzyme chosen for digestion, database searching procedure and decoy database (Vensel et al., 2011). The combination of Mascot and X!Tandem with the enzymatic cleavages performed by trypsin, chymotrypsin, and thermolysin, allows to increase the number of peptides and the number of identified proteins with respect to the use of a simple trypsin cleavage. Taking as example the α -gliadins, trypsin generated only the 1% of the identified peptides compared to the 59% and 40% of chymotrypsin and thermolysin, respectively. Moreover, both chymotrypsin and thermolysin produced peptides yielded complementary b- and y-series with clear and interpretable fragmentation patterns by n-ESI-Q-TOF MSMS analysis. Database search was performed on a “Super_Wheat” protein database, obtained combining assembled EST sequences from publicly available databases into cultivar specific contigs (Altenbach, Vensel, & DuPont, 2010a,b) with sequences from publicly available plant sequence repositories. The results, influenced by searching engine, database size, and the type of decoy database, emphasize how the searching criteria can determine the final results.

Another important aspect of gliadins is their involvement in triggering celiac disease (CD) (Ferranti et al., 2007; Kagnoff, 2007), a common enteropathy induced by ingestion of wheat gluten proteins and related prolamins from rye (secalin) and barley (hordein) in genetically susceptible individuals. In this context, the role of the MS-based approaches, aimed to the characterization of the epitopes and the explanation of the structural basis for gluten toxicity, is best established by the identification of peptides from α - and γ -gliadins, obtained by working on recombinant gliadins and synthetic peptides. LC-MS/MS/UV analysis of the enzymatic digested of a recombinant $\alpha 2$ -gliadin allowed the identification a relatively large Gln- and Pro-rich fragment, the “33-mer,” with high resistance to gastrointestinal breakdown (Shan et al., 2002). Moreover, sequence alignment searches using BLAST (Basic Local Alignment Search Tool) in protein databases evidenced the existence of several food grain-derived homologs of the 33-mer gliadin peptide. Noticeably, homolog peptides were only found in gliadins, hordeins, and secalins, belonging to toxic cereals in the celiac diet, whereas non-toxic cereal proteins, such as avenins (in oats), rice, and maize do not show homologous sequences to 33-mer gliadin. This result suggested that the 33-mer could represent the primary initiator of the inflammatory response to gluten in Celiac Sprue patients. On the other hand, it has been demonstrated that the enzyme tissue transglutaminase 2 (tTG or TG2), in a slightly acidic environments such as a typical of chronic inflammation, catalyzes the deamidation of Q-residues present in the gluten derived peptides to E-residues making the last ones better

T-cell antigens (Molberg et al., 1998). The high sensitivity, resolution, and mass accuracy of hybrid Q/TOF instruments allow differentiation of peptides where a single deamidated Q-residue is present. Several studies have pointed towards the role of TG2 in the selection of T-cell epitopes in CD, allowing the identification of several T-cell epitopes of wheat gluten (Dørum et al., 2010). A peptidomics approach was developed to identify transglutaminase-susceptible Q residues within a pepsin/trypsin gliadin digest from a winter wheat flour variety (Mamone et al., 2004). Particularly, tagging with a monodansyl cadaverine (MDC) fluorescent probe the TG2-susceptible glutamine residues, their localization within the peptide chain was obtained by nanospray tandem mass spectrometry, using the characteristic fragmentation by the 446.2 Da loss corresponding to Q-MDC instead of 128 Da of the native Q-residue. Moreover, it has been demonstrated that only a few gliadin peptides interact with TG2 and the preferentially modified glutamine residues are located at the +2 position with respect to a proline residue. In addition, the majority of gluten T-cell epitopes identified are originated from α - and γ -gliadins, and show different recognition pattern by patient derived T cells. α -Gliadin derived epitopes are recognized by T cells isolated from nearly all CD patients, whereas the responses against γ -epitopes are less frequently observed (Marti et al., 2005). This different behavior is related to the rate of TG2-mediated deamidation, which differs between these peptides. Particularly, α -gliadin derived epitopes showed a significant higher level of deamidation compared to the majority of epitopes from γ -gliadins that are less frequently recognized (Dørum et al., 2009).

On the other hand, the detection of the so-called “hidden allergens” in processed foods is needed for the food industry to ensure the correct labeling of their products and protect allergic consumers.

So that, gliadin extracts have been extensively studied by MALDI-TOF MS in order to identify their presence in flours, starches, and processed foods by the detection of characteristic mass patterns (Camafrita et al., 1998). Direct observation of MALDI-MS prolamins pattern represents a fast and specific method and is currently the only non-immunological analytical tool available for the detection of the presence of cereal proteins in gluten-free food. On the other hand, the alternative RP-HPLC/ESI-MS approach can be used for the identification of the γ -gliadin components, directly involved in CD, and therefore for selective monitoring of wheat-contaminated foods, even at minute levels. In addition, the characterization by LC-MS analysis of gliadin profiles from durum wheat may be used to differentiate the cultivars in terms of the nature and abundance of the protein components (Mamone et al., 2000). It is important to note that although different cultivars may show similar chromatographic profiles, the experimentally determined M_r of protein components with similar retention times are often different among the varieties, due to the occurrence of amino acid substitutions, insertions, or deletions in the gliadin sequences. Another characteristic feature of the LC-MS analyses of gliadins is that almost all the HPLC fractions contain several co-eluting components, which are easily detected only by ESI-MS analysis. Therefore, the number of protein species detected by LC-MS analyses in a single wheat variety is always higher than those detectable by gel electrophoretic or HPLC-UV analyses. Indeed, co-eluting proteins often differ for a single amino

acid substitutions or presence or absence of short peptide sequences and therefore they show minor differences in their molecular masses and may have similar electric charge. As a consequence they could comigrate in gel electrophoretic analysis leading to an underestimation of the number of protein species.

Similar consideration also applies to the differentiation hordein patterns. In a recent work, MALDI-TOF analysis, was successfully applied to identification of 13 malting barley varieties based on the characteristic mass patterns of hordeins (Salplachta & Bobálová, 2009). Mass spectra showed that all the hordein extracts were composed of ten of protein components belonging to the hordein subfraction groups: A hordeins ($M_r < 30$ kDa), B hordeins (30 kDa $< M_r < 50$ kDa), and C hordeins (50 kDa $< M_r < 75$ kDa). On the contrary, no signals of D-hordeins ($M_r > 90$ kDa) were detected. According to literature data, slight or no differences of A hordein signals were observed in different varieties. On the contrary, while C hordein signals were too weak to take these proteins into consideration, the profile of B hordeins showed remarkable differences from one variety to another and therefore could be considered to enable barley variety differentiation. However, it should be highlighted that, even if barley's storage proteins account for the 80% of the total protein content, the differences in hordein profile are considered of minor importance for malting qualities with respect to the salt-soluble protein fraction (Görg et al., 1992). Therefore, numerous studies on barley salt-soluble proteome were performed in order to determine quality composition correlations (Finnie et al., 2004a, 2009). As an example, an integration of the barley genetic data and seed proteome maps was used to screen 18 spring barley cultivars. This investigation allowed the identification of differences in seed protein profiles. Particularly, in some cases when a 2D spot in one cultivar was replaced in another one by a spot with similar molecular mass but different pI, MS peptide mass mapping allowed to demonstrate that spot variation was due to an amino acid substitution encoded by single nucleotide polymorphism (SNPs) (Finnie et al., 2009). Such proteome variations can be directly related to the genetic map and represent a link between cultivar traits, proteome and genome.

Another class of coeliac immunoreactive proteins is constituted by the secalins, the major storage proteins in rye (Gellrich, Schieberle, & Wieser, 2003). They are commonly classified on the basis of their N-terminal amino acid sequences and their molecular masses in ω -secalins, γ -75k and γ -40k secalins. Studies of amino acid composition, N-terminal amino acid sequences and molecular masses indicated that γ -40k secalins are closely related with γ -gliadins. However, secalins have been less studied with respect to the last ones. The few works reported on secalins are on the identification of major rye secalins as coeliac immunoreactive proteins (Rocher et al., 1996) and on the characterization of their primary structure and disulfide linkages in aggregated secalins. Particularly, studies carried out on the trypsin and thermolysin digests of γ -40k, coupling sequence analysis and mass spectrometric data (Gellrich, Schieberle, & Wieser, 2005), showed that the N-terminal domain of γ -40k secalins, like the corresponding region of γ -gliadins is poor of lysine and arginine residues and does not show cysteines and that their C-terminal domains share 85% of homology with the corresponding one of γ -gliadins. Similarly, the characterization of the disulfide

structure of aggregated γ -75k secalins (Gellrich, Schieberle, & Wieser, 2004a,b) allowed to demonstrate why these proteins, though partially homologous to gluten proteins, are not able to form comparably large gluten-like aggregates. Indeed, this study showed that γ -75k secalins, in their C-terminal domain, present four intramolecular disulfide bonds which are homologous to those of monomeric γ -gliadins, and in the N-terminal domain present, likely the LMW-GSs, an additional cysteine residue which is responsible for the aggregative nature of γ -75k secalins. In contrast to LMW subunits, γ -75k secalins do not possess an additional cysteine residue in the C-terminal region that could form a second intermolecular linkage and, therefore, the formation of large gluten-like aggregates in rye is hampered.

The corresponding prolamins in barley, consist of four major groups of polypeptides called A, B, C, and D hordeins, which account for 75–90%, 10–20%, and 2–4% of the total fraction, respectively. Determination of the molecular masses of nine C hordeins extracted from the barley mutant Risø 56 showed that, among nine isolated fractions, three contained a single component, whereas the others contained at least two components with minor differences in the molecular mass (Savage et al., 2002). The molecular mass values determined by MALDI-TOF MS (43,561–57,606 Da) were lower than those deduced from SDS-PAGE mobility (55–70 kDa) but higher than those reported based on gene sequences (28–39 kDa), suggesting that genes encoding typical C hordein proteins were not so far been isolated. As above reported for LMW-GSs and gliadins, the amino acid composition of hordeins, with low numbers of lysines and arginines residues, in combination with high content of proline and glutamine, complicates their identification in a conventional proteomic workflow. Therefore, in order to extend the sequence information on this kind of proteins, alternative approaches are requested. Recently, a new proteomic approach for hordeins characterization, which can also be used for the analysis of glutenin subunits and gliadins in wheat, or other cysteines containing proteins having a low number of lysines and arginines, has been reported (Rehulková et al., 2009). By this approach, in order to increase the number of cleavages sites, cysteines were converted into aminoethylcysteines by 2-bromoethylamine (BEA) prior to trypsin digestion. Aminoethylcysteines mimic lysine residues toward trypsin and therefore are recognized by this enzyme as potential cleavage sites. Thus, the protein chain is also cleaved at the C-terminal side of the modified cysteine, if not followed by a proline residue. Small extent of side reactions (towards histidine, protein N-terminus and methionine), could be observed after a longer exposition to BEA, but this did not hampered the analysis. The alkylation with BEA prior to trypsin digestion, which can be performed both in-gel and in-solution, significantly increases the obtained sequence information in comparison to a standard proteomic protocol involving tryptic digestion alone. Therefore, it represents a valid alternative approach in the characterization of proteins having a low number of lysine and arginine residues.

V. EFFECT OF STRESS, GENETICALLY MODIFIED CROPS, ALLERGENS IN CEREALS

Functional and nutritional properties of cereals are strictly related to their protein composition, which therefore represents

the target of a considerable number of proteomic researches. Cereal proteomics is mainly focused on the investigation of the different varieties and their characteristics related to stress, pathogens, yield, and to develop safe and high quality food. The major trouble of the proteomic studies on cereals is that the only completed, publicly available and well-annotated genome is that of rice, whereas the genomes of the other cereals are only partially sequenced. Major efforts are being made to describe the cereals transcriptome by sequencing the cDNAs, in order to generate ESTs, instead of sequencing the whole genome (Matthews et al., 2003). However, despite sequencing of the genomes of almost all cereals is still in progress, a large quantity of proteomic studies have been carried out on cereals such as wheat, barley, and maize, using homology-based searching of MS and MS/MS data for protein identification and ESTs data.

By this approach, which constitutes the only option whenever a genome is poorly characterized, proteins are identified by comparing peptides of the proteins of interest to orthologous proteins of species that are well characterized (Wilkins & Williams, 1997; Lester & Hubbard, 2002; Mathesius et al., 2002; Liska & Shevchenko, 2003; Junqueira et al., 2008). Many of the most representative examples of combination of protein and ESTs database search for protein identification on cereals were carried out from B. Svensson's group on barley proteome (Finnie & Svensson, 2003; Bak-Jensen et al., 2004; Østergaard et al., 2004). However, despite the limited sequence information available, proteomics approach has now increased in power thanks to the combination of MS with DIGE (differential in gel electrophoresis) followed by gel free techniques, such as isotope-coded affinity tags (ICAT), iTRAQ, and MudPIT (Afroz et al., 2011). Proteomics studies on cereal proteins have gone beyond simple cataloguing of proteins, being successful applied to the identification of individual proteins as markers of stress effects, GM varieties, and to the investigation of functional and regulatory aspects of proteins responsible of allergy.

A. Analysis of the Effects of Stresses in Crops

In nature, all crops suffer from environmental or other organism stresses. Environmental stresses, which include drought, salinity, temperatures extremes, and poor soil nutrition, are referred to as abiotic stresses. Stresses exerted on crops by other plants (weeds), animal pests, and pathogenic organisms are referred to as biotic stresses. Biotic and abiotic stresses can cause adverse effects on the growth of plants and cause loss of the productivity of crops, reducing average yields for most major crop plants. Plants respond to these stresses at molecular and cellular levels as well as at physiological and biochemical levels. This defence is brought about by alteration in the pattern of gene expression. Particularly, qualitative and quantitative changes in proteins are the result, leading to the modulation of certain metabolic and defensive mechanisms.

Therefore, a better understanding of these molecular adaptation mechanisms is crucial for continued development of rational breeding and also transgenic strategies to improve stress tolerance of crops (Skylas et al., 2002; Xiong, Schumaker, & Zhu, 2002; Afroz et al., 2011).

In the last decade, numerous comparative proteomic investigations of plants before and after specific stresses have been performed, in order to analyze biochemical pathways and the complex response of cereal plants (Qureshi, Qadir, & Zolla, 2007). Almost all these investigations are carried out by gel-based approaches. Differentially expressed spots from two series of 2D-maps are investigated for both protein identification and characterization. The differentially expressed proteins identified by this approach are usually grouped according their putative physiological functions. Metabolic changes in maize plants under phosphorus deficit, which represent a major limitation for plant growth and development, were investigated by this approach (Li et al., 2007). Authors identified 106 (20% of all the detected proteins) differentially expressed proteins under phosphate stress by MALDI-TOF MS. The proteins identified were mainly involved in biosynthesis of phytohormones, cell cycle, signal transduction, and transcriptional regulation. On the other hand, plants respond to pathogenic attacks modifying their cellular metabolism and producing antimicrobial compounds that act directly to prevent pathogen invasion. As an example, a classical proteomic approach was used to investigate the defense response of germinating maize embryos against fungal infection (Campo et al., 2004). By coupling homology-based searching of MALDI-TOF data with nESI-IT MS/MS, the proteins induced in response to fungal infection were identified. Some of them, were enzymes involved in the protection of plant tissues against oxidative damage and detoxification of cytotoxic products. However, a general trend observed in many different studies is that, besides over-expressing specific proteins as consequence of each stress, plants respond to diverse stresses in a similar manner. In fact, together a wider range of metabolic proteins, the largest type of differentially expressed proteins is constituted by heat shock proteins (HSPs) (Timperio, Egidi, & Zolla, 2008), a class of functionally related proteins usually named according to their molecular weight (i.e., HSP 60, HSP 70, and HSP 90). The most prominent heat shock induced proteins in plant are constituted by small HSP (smHSPs, M_r 15–40 kDa) which are usually undetectable in vegetative tissues under physiological conditions but increase in abundance by developmental stimuli or by environmental stresses (Sun, Van Montagu, & Verbruggen, 2002). Heat shock proteins are induced or accumulated in response to a broad spectrum of stresses such as low and high temperatures, drought, salinity, and act as molecular chaperons. A recent investigation of abundant and low-abundant proteins of rice leaves upon heat stress (Lee et al., 2007) showed a total of 48 differentially expressed proteins in stressed samples compared to controls. More than 37% of the proteins identified belong to the functional class of HSPs or smHSPs, suggesting again that they may play a pivotal role in protecting the cellular damage under heat stress (Majoul et al., 2003, 2004; Sule et al., 2004; Lin et al., 2005). In another study, the consequences of heat stress on the accumulation of metabolic proteins in mature durum wheat kernels were investigated. This work showed that heat stress alters the durum wheat seed proteome, revealing 132 differentially expressed proteins with respect to the control seeds. Forty-seven components were identified by MS analysis and included many HSPs (Laino et al., 2010).

Changes in cereal proteome in response to drought, a major environmental parameter adversely affecting development

and crop yield, have been investigated (Ali & Komatsu, 2006; Hajheidari et al., 2007). Three wheat genotypes under normal and drought conditions were investigated and many of the differentially expressed proteins were related to the response to water stress. Most of drought-related proteins in cereal grain are redox-related and are known to undergo a pronounced change in sulfhydryl status during development and germination (Buchanan & Balmer, 2005). The redox-linked proteins, target of thioredoxins (Trxs), could serve as genetic markers for this type of stress.

The proteomics approaches used to investigate the target proteins of different Trx involved two different strategies, (i) labeling of target proteins with thiol-specific probes and (ii) trapping of target proteins by binding to active-site mutants of Trxs. These approaches have been highly effective since several hundred putative target proteins have been identified (Wong et al., 2004; Montrichard et al., 2009).

The production and scavenging of chemically reactive species, such as reactive oxygen species (ROS)/reactive nitrogen-oxide species (RNS), are central to a broad range of biotic and abiotic stress and physiological responses in plants (Rinalducci, Murgiano, & Zolla, 2008; Witzel et al., 2009).

The global identification of PTMs related to protein oxido-reduction and characterization of redox control systems is becoming an interesting field of plant proteomics (Navrot et al., 2011). Consequently, the latest researches in the field of plant redox proteomics would deserve a more exhaustive overview. Difficulties in redox studies deal with sample preparation which could interfere with the dynamic and labile nature of redox PTMs (e.g., methionine and tryptophan oxidation) (Cohen, 2006; Perdivara et al., 2010) and with the limitation occurred in the identification of modified proteins. Targets of present and future attempt to monitor changes in redox PTMs on a proteome scale are cysteine, methionine, and tryptophan oxidation, and nitrosylation of cysteine and tyrosine residues.

B. Analysis of Genetically Modified Crops

Despite the controversies surrounding GM organisms, the production of GM crops is increasing. Therefore, the detection of genetically modified organisms (GMOs) is one of the most important consumer concerns regarding food safety and quality. Several strategies have been developed to identify unintended alterations in the composition of GM food crops that may occur as a result of the genetic modification process (Cellini et al., 2004; Chassy et al., 2004). The concept of substantial equivalence was formulated as a comparative tool to identify similarities and differences between the GM food crop and its isogenic counterpart (OECD, 1993). Introduction of exogenous DNA sequences into the plant genome for GMO production may result, in addition to the intended effects, in unexpected events such as possible inactivation of endogenous genes (Somers & Makarevitch, 2004) or activation of silent genes and formation of fusion proteins (Conner & Jacobs, 1999). In addition, a possible modification of metabolic pathways in GMOs could lead to alterations in concentrations of nutrients or, in theory, even to the production of new protein allergens or toxins, that could compromise the product safety (Kuiper et al., 2001; Cellini et al., 2004). Identification of unintended effects resulting from genetic modification are carried out by two different approaches: the targeted and

non-targeted approach. Normally a targeted approach, such as measurements of single known micro- or macro-nutrients, toxins and anti-nutrients, is used for characterize spotting alterations in the composition of a GM organism compared with the parent organism. However, the targeted approach has its limitations with respect to a restricted selection of compounds, and cannot be used for the detection of unknown toxins or anti-nutrients. To increase the chances of detecting unintended effects, profiling investigations by non-targeted approaches such as DNA/RNA microarray technology, metabolomics, or proteomics are employed. A proteomic approach, involving electrophoretic, chromatographic, and mass spectrometric methods, allowed the identification and characterization of the transgenic protein in Roundup ReadyTM soya and maize. Particularly, the identification was accomplished in 0.9% GM soya seeds, thus demonstrating the applicability of MS with respect to the current European Union (EU) threshold for required GMO food-labeling (Ocaña et al., 2007). Proteomics was also applied for a comparative investigation between the environmental effects and gene insertion by particle bombardment (event MON 810) on protein expression of a maize variety (Zolla et al., 2008). This study showed that the environment affects protein expression (in term of total number of protein spots) more than gene insertion by particle bombardment. Furthermore, transgenic plants reacted differentially to the same environmental conditions with respect to their wild-type lines because they had a strongly modified genome. On the light of these results, the proteomic approach could represent a powerful tool to detect GM varieties, to evaluate the safety of both GM and conventionally bred crops and to choose the transgenic seeds having the less negative alterations. Proteomics can be also applied to determine similarities and differences between the existing traditional crops and the new transgenic ones. On this respect, protein spot patterns from five transgenic barley lines were compared with those from greenhouse-grown wild-type Golden Promise seeds, an old barley cultivar with high salt tolerance but susceptible to several fungal pathogens (Finnie et al., 2004b). In-gel digestion, MALDI-TOF MS and peptide mass mapping of differentially expressed 2D spots allowed to evidence that 2D-gel patterns of transgenic seeds presented spots containing the herbicide resistance marker phosphinothricin acetyltransferase, which was instead absent in the wild-type Golden Promise seeds.

In another study, a parallel transcriptional and proteomic comparison of seeds from a transformed bread wheat line that over-expresses a transgenic low molecular weight glutenin subunit relative to the corresponding non-transformed genotype was performed (Scossa et al., 2008). Results showed that as a response to the over-accumulation of the LMW-GS, a decrease gene expression and synthesis of most of the endosperm proteins belonging to the storage proteins was observed in the transgenic seeds. This differential accumulation was detected by proteomics analysis and was also observed in the corresponding changes in transcript levels by microarray experiments. This response might be a mechanism of compensation, which seems to be proportional to the expression level of the introduced glutenin transgene. In fact, plants might react to a strong increase in the production of nutrients or storage compounds by activating a silencing/degradation pathway avoiding the over-accumulation of storage reserves.

As mentioned previously, the genetic modification of plants could raise the production of new proteins that could potentially be human allergens. On the other hand, the converse was demonstrated using genetic engineering to remove a pollen allergen from ryegrass (Bhalla, Swoboda, & Singh, 2001). Another study shows that it is possible to eliminate a major or immunodominant soybean seed allergen, the Gly m Bd 30K protein, by gene-silencing techniques (Herman et al., 2003). These results also provided an evidence for substantial equivalence of composition of transgenic and non-transgenic seeds, after elimination of the dominant allergenic protein.

However, it should be noted that one of the major difficulties in proteomics approaches involving 2DE and MS techniques is the quantification of proteins. Indeed, only the most abundant proteins are detected by 2DE and the dynamic range of quantification is limited (Görg, Weiss, & Dunn, 2004; Issaq & Veenstra, 2008). Alternatives may be provided by quantification on the basis of ICAT (Chen et al., 2007). Moreover, the detection of differences related to the genetic modification is often very difficult, taking into account the great amount of proteins not connected to such changes and given the natural variability of plant protein patterns related to the environmental conditions. On this respect, the development of validated interconnected databases containing information on variations in profiles associated with developmental stages and environmental conditions, together with the combination of immunoblotting and protein microarrays may offer interesting possibilities.

C. Analysis of Cereal Allergens

Cereals represent one of the eight foods which are responsible for more than 90% of food allergies. Processing of wheat and related cereals may lead to decreased allergenicity, while, as above reported, engineering technology offers opportunities to eliminate allergens by suppressing gene expression. Symptoms of cereal allergy include atopic eczema/dermatitis, urticaria, and anaphylaxis which appear to be related to several cereal proteins, mainly belonging to the cupin superfamily or prolamin superfamily with storage and/or protective functions. Two types of allergy are particularly well-characterized: the bakers' asthma (Houba, Doekes, & Heederick, 1998), resulting from inhalation of flour and dust during grain processing, and the wheat-dependent exercise-induced anaphylaxis (WDEIA), a form of allergic reaction induced by the ingestion of a causative allergen and subsequent physical exercise (Lee et al., 2008). Most of the recent studies show that the cereal grain α -amylase inhibitors (also known as chloroform methanol, or CM, soluble proteins) represent the major group of proteins responsible for bakers' asthma (Salcedo et al., 2004), while the ω_5 -gliadins (part of gluten protein fraction) appear to be more important in WDEIA (Morita et al., 2003). Recently, proteomic strategies, mainly based on 2DE, IgE immunoblotting, and MS analyses, have been applied for rapid and easy identification of allergenic proteins, and the new strategy is now referred to as "allergenomics" (Yagami et al., 2004).

It should be noted that three main problems hamper the identification of cereals food allergens. The first one concerns the inability of classical proteomic approaches to analyze all protein components. This inability is directly related to the wide dynamic range of protein concentrations, biological

variance, protein extraction efficiency, and proteome complexity. Improvements in proteome coverage have come from sub-cellular proteomics and multidimensional HPLC studies. The second trouble is the absence of double-blind, placebo-controlled food challenge studies that compare the allergenic profile of Osborne's fractions in subjects with real cereals allergy. Finally, the last problem is related to the absence of data concerning the differences in IgE-binding capacity between raw and cooked cereals (Pastorello et al., 2007). However, not always the results obtained by "allergenomics" approaches are in agreement with those from classical immunological studies (i.e., ELISA, Western blot) on cereal food allergy. A comprehensive characterization of wheat allergens by means of 2D-PAGE, immunoblot analysis, and mass spectrometry was carried out on whole wheat flour proteins (Akagawa et al., 2007). IgE-binding proteins detected by immunoblotting with sera of patients with a food allergy to wheat were identified by coupling in-gel trypsin digestion of the corresponding 2D-gel spots, MALDI-TOF MS and database searching. This approach allowed the identification of four well-known wheat allergens (serpins, α -amylase inhibitors, γ -gliadins, and LMW-GSs), whereas other known allergens such as HMW-GSs, α/β -, and ω -gliadins were not recognized. Such discrepancies may be ascribed to differences in wheat cultivar, patient population, and method used, or may be assigned to allergens which, extracted from different flours, differ slightly in amino acid composition and/or PTMs and, therefore, could exhibit different epitopes. Moreover, the observed differences in the IgE-response could be related to the currently applied methodologies. Indeed, the reduction and alkylation steps performed before 2DE analysis alter the structure of proteins and may determine the antibody binding sites disappearance. In order to investigate the impact of protein treatments before electrophoresis, a comparison of the detectable number IgE-binding components of wheat salt-soluble extracts has been performed both by 1DE under non-reducing and reducing conditions (Sotkovský et al., 2008). 1DE under non-reducing conditions and immunoblotting revealed IgE-binding proteins of molecular mass ranging 10–100 kDa, whereas separation under reducing conditions showed predominantly low molecular mass components ranging 10–30 kDa recognized by the majority of patient's sera. These changes in the IgE-binding pattern could be a consequence of conformational changes, reduction of inter-chain disulfide bridges, and uncovering or generation of new antigenic structures (epitopes). To mimic the digestion of wheat containing food, the effect of pepsin on the immunogenicity of the extracted wheat proteins was also tested. Two-dimensional gel electrophoresis and IgE-antibodies (Abs) analyses were performed on both intact and pepsin digested extracts, revealing the presence of numerous wheat allergens. Proteins, frequently recognized by IgE, were analyzed by MALDI-TOF, nLC-MS/MS, and MAS-COT searching. By this approach 19 potential wheat allergens, including well-known (CM16 and CM17 α -amylase inhibitors, profilin, Tri a Bd 27K protein, etc.) and some new IgE-binding proteins (β -amylase, β -D-glucan exohydrolase, serpin, peroxidase 1, etc.) were identified.

Recently, six allergens (vicilin, globulin-2, 50 kDa γ -zein, endo-chitinase, thioredoxin, and trypsin-inhibitor) already known in several food stuff, were found, for the first time, in maize flour, a system that has been thought to be a

non-allergenic food par excellence (Fasoli et al., 2009; Pastorello et al., 2009). This investigation adds a non-negligible number of allergens to the scant list of maize allergens reported so far, constituted only by nsLTP, and therefore lays out the foundation for screening GM maize in search of potential new allergens.

In routine settings, natural allergen extracts are used for *in vivo* (skin-prick) and *in vitro* (specific IgE Abs) testing. As was demonstrated by several authors, most patients do not raise IgE Abs in response to all extracted cereal components, but only to some. The studies currently performed contribute to the identification of frequently recognized cereal allergens and to the analysis of factors affecting immunogenicity of these proteins. This can be considered only a starting point and a necessary step in the development of diagnostic test with higher sensitivity and specificity (Jahn-Schmid et al., 2003; Wöhrle et al., 2006). The characterization of well-defined cereal allergens could be either applied in specific IgE Abs testing, cell activation assays and also in complex allergen-microarrays (biochip technology).

VI. CONCLUSIONS AND PERSPECTIVES

Mass spectrometry, virtually unemployed until 15 years ago for the investigation of cereal proteins, has made rapid progress as an analytical technique with many new types of platforms being introduced. Improvements of sensitivity, selectivity and mass measurement accuracy have transformed mass spectrometry, establishing itself as an indispensable tool for protein characterization in cereals. The role of mass spectrometry rapidly evolved from simple measurements of the molecular mass of isolated proteins to a fine determination of structural features such as amino acid sequences and the site of attachment and type of PTMs of proteins present in a complex mixtures. However, it should be noted that the investigation of prolamins by current methodologies is often very difficult and challenging due to the peculiar characteristics of these proteins. Indeed, most prolamins, encoded by a high number of genes with similar structural characteristics, have amino acid sequences with a long repeating motif and a scarcity of cleavable tryptic sites, that lead to a pool of tryptic peptides which are frequently not suitable for MS/MS analysis. In order to overcome these difficulties, classical approaches for protein analysis should be adapted to the peculiarity of these proteins by introducing appropriate modifications. On the other hand, the genomes of cereals, with exception of rice, are only partially known and major efforts are being made to describe their transcriptome by sequencing the cDNAs, generating ESTs, instead of sequencing the whole genome. With the aim of improving the identification rate of cereal proteins, high-throughput database dependent searching (cross-species and EST-based) and databases independent *de novo* sequencing, combined with error tolerant BLAST searching are usually performed.

Proteomic studies on cereals have rapidly increased in the last years moving beyond the large-scale protein identification to the study of functional and regulatory aspects of proteins such as variation in protein composition during grain filling and seed maturation or differential protein expression under different environmental conditions or specific stresses. On the other hand, current MS-based approaches for the detection of

GM varieties, allergenic proteins in food, and identification of lines containing prolamins associated with superior quality can provide practical and reliable analytical protocols for the development of breeding programs.

REFERENCES

- Aebersold R, Mann M. 2003. Mass spectrometry-based proteomics. *Nature* 422:198–207.
- Afroz M, Ali GM, Mir A, Komatsu S. 2011. Application of proteomics to investigate stress-induced proteins for improvement in crop protection. *Plant Cell Rep* 30:745–763.
- Agrawal GK, Jwa NS, Rakwal R. 2009. Rice proteomics: Ending phase I and the beginning of phase II. *Proteomics* 9:935–963.
- Akagawa M, Handoyo T, Ishii T, Kumazawa S, Morita N, Suyama K. 2007. Proteomic analysis of wheat flour allergens. *J Agric Food Chem* 55:6863–6870.
- Ali GM, Komatsu S. 2006. Proteomic analysis of rice leaf sheath during drought stress. *J Proteome Res* 5:396–403.
- Altenbach SB, Vensel WH, DuPont FM. 2010a. Analysis of expressed sequence tags from a single wheat cultivar facilitates interpretation of tandem mass spectrometry data and discrimination of gamma gliadin proteins that may play different functional roles in flour. *BMC Plant Biol* 10:7.
- Altenbach SB, Vensel WH, DuPont FM. 2010b. Integration of transcriptomic and proteomic data from a single wheat cultivar provides new tools for understanding the roles of individual alpha gliadin proteins in flour quality and celiac disease. *J Cereal Sci* 52:143–151.
- An X, Zhang Q, Yan Y, Li Q, Zhang Y, Wang A, Pei Y, Tian J, Wang H, Hsam SLK, Zeller FJ. 2006. Cloning and molecular characterization of three novel LMW-i glutenin subunit genes from cultivated einkorn (*Triticum monococcum* L.). *Theor Appl Genet* 113:383–395.
- Bak-Jensen KS, Laugesen S, Roepstorff P, Svensson B. 2004. Two-dimensional gel electrophoresis pattern (pH 6–11) and identification of water-soluble barley seed and malt proteins by mass spectrometry. *Proteomics* 4:728–742.
- Bhalla PL, Swoboda I, Singh MB. 2001. Reduction in allergenicity of grass pollen by genetic engineering. *Int Arch Allergy Appl Immunol* 124:51–54.
- Bienvenut WV, Deon C, Pasquarello C, Campbell JM, Sanchez JC, Vestal ML, Hochstrasser DF. 2002. Matrix-assisted laser desorption/ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins. *Proteomics* 2:868–876.
- Breiteneder H, Radauer C. 2004. A classification of plant food allergens. *J Allergy Clin Immunol* 113:821–830.
- Buchanan BB, Balmer Y. 2005. Redox regulation: A broadening horizon. *Annu Rev Plant Biol* 56:187–220.
- Bunce NAC, White RP, Shewry PR. 1985. Variation in estimates of molecular weights of cereal prolamins by SDS–PAGE. *J Cereal Sci* 3:131–142.
- Camafeita E, Solís J, Alfonso P, López JA, Sorell L, Méndez E. 1998. Selective identification by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of different types of gluten in foods made with cereal mixtures. *J Chromatogr A* 823:299–306.
- Campo S, Carrascal M, Coca M, Abián J, San Segundo B. 2004. The defense response of germinating maize embryos against fungal infection: A proteomics approach. *Proteomics* 4:383–396.
- Carbonero P, García-Olmedo F. 1999. A multigene family of trypsin/α-amylase inhibitors from cereals. In: Shewry PR, Cases R, editors. *Seeds proteins*. Dordrecht: Kluwer Academic Publisher. pp. 617–633.
- Carrillo JM, Vazquez JF, Orellana J. 1990. Linkage relationships between the loci *Sec 1* and *Sec 3* in rye (*Secale cereale* L.). *Heredity* 64:125–130.
- Cellini F, Chesson A, Colquhoun I, Constable A, Davies HV, Engel KH, Gatehouse AMR, Karenlampi S, Kok EJ, Leguay JJ, Lehesranta S, Nøtborn HP, Pedersen J, Smith M. 2004. Unintended effects and their detection in genetically modified crops. *Food Chem Toxicol* 42:1089–1125.
- Chassy B, Hlywka JJ, Kleter GA, Kok EJ, Kuiper HA, Mc Gloughlin M, Munro IC, Phipps RH, Reid JE. 2004. Nutritional and safety assessments of foods and feeds nutritionally improved through biotechnology: An executive summary. *Compr Rev Food Sci Food Saf* 3: 38–104.
- Chen X, Sun L, Yu Y, Xue Y, Yang P. 2007. Amino acid-coded tagging approaches in quantitative proteomics. *Expert Rev Proteomics* 4: 25–37.
- Chernushevich IV, Loboda AV, Thomson BA. 2001. An introduction to quadrupole/time-of-flight mass spectrometry. *J Mass Spectrom* 36: 849–865.
- Cohen SL. 2006. Ozone in ambient air as a source of adventitious oxidation. A mass spectrometric study. *Anal Chem* 78:4352–4362.
- Conner AJ, Jacobs JM. 1999. Genetic engineering of crops as potential source of genetic hazard in the human diet. *Mutat Res* 443:223–234.
- Cozzolino R, Di Giorgi S, Fisichella S, Garozzo D, Lafiandra D, Palermo A. 2001a. Matrix-assisted laser desorption/ionization mass spectrometric peptide mapping of high molecular weight glutenin subunits 1Bx7 and 1Dy10 in Cheyenne cultivar. *Rapid Commun Mass Spectrom* 15:778–787.
- Cozzolino R, Di Giorgi S, Fisichella S, Garozzo D, Lafiandra D, Palermo A. 2001b. Proteomics of gluten: Mapping of subunit 1Ax2* in Cheyenne cultivar by matrix-assisted laser desorption/ionization. *Rapid Commun Mass Spectrom* 15:1129–1135.
- Cunsolo V, Foti S, Saletti R, Gilbert S, Tatham AS, Shewry PR. 2002. Investigation and correction of the gene-derived sequence of glutenin subunit 1Dx2 by matrix-assisted laser desorption/ionisation mass spectrometry. *Rapid Commun Mass Spectrom* 16:1911–1918.
- Cunsolo V, Foti S, Saletti R, Gilbert S, Tatham AS, Shewry PR. 2003. Structural studies of glutenin subunits 1Dy10 and 1Dy12 by matrix-assisted laser desorption/ionisation mass spectrometry and high-performance liquid chromatography/electrospray ionisation mass spectrometry. *Rapid Commun Mass Spectrom* 17:442–454.
- Cunsolo V, Foti S, Saletti R, Gilbert S, Tatham AS, Shewry PR. 2004a. Structural studies of the allelic wheat glutenin subunits 1Bx7 and 1Bx20 by matrix-assisted laser desorption/ionisation mass spectrometry and high-performance liquid chromatography/electrospray ionisation mass spectrometry. *J Mass Spectrom* 39:66–78.
- Cunsolo V, Foti S, Saletti R. 2004b. Mass spectrometry in the characterization of cereal seed proteins. *Eur J Mass Spectrom* 10:359–370.
- D'Ovidio R, Masci S. 2004. The wheat low-molecular-weight glutenin subunits. *J Cereal Sci* 39:321–339.
- Delahunty CM, Yates JR III. 2007. MudPIT: Multidimensional protein identification technology. *BioTechniques* 43:563–569.
- Dørum S, Qiao SW, Sollid LM, Fleckenstein B. 2009. A quantitative analysis of transglutaminase 2-mediated deamidation of gluten peptides: Implications for the T-cell response in celiac disease. *J Proteome Res* 8(4):1748–1755.
- Dørum S, Arntzen MØ, Qiao SW, Holm A, Koehler CJ, Thiede B, Sollid LM, Fleckenstein B. 2010. The preferred substrates for transglutaminase 2 in a complex wheat gluten digest are Peptide fragments harboring celiac disease T-cell epitopes. *PLoS ONE* 5(11):e14056, and references therein.
- Dumur J, Jahier J, Bancel E, Laurière M, Bernard M, Branlard G. 2004. Proteomic analysis of aneuploid lines in the homeologous group 1 of the hexaploid wheat cultivar Courtot. *Proteomics* 4:2685–2695.
- Dunwell JM. 1998. Cupins: A new superfamily of functionally diverse proteins that include germins and plant storage proteins. *Biotechnol Genet Eng Rev* 15:1–32.
- DuPont FM, Vensel W, Encarnacao T, Chan R, Kasarda DD. 2004. Similarities of omega gliadins from *Triticum urartu* to those encoded on

- chromosome 1A of hexaploid wheat and evidence for their post-translational processing. *Theor Appl Genet* 108:1299–1308.
- Dworschak RG, Ens W, Standing KG, Preston KR, Marchylo BA, Nightingale MJ, Stevenson SG, Hatcher DW. 1998. Analysis of wheat gluten proteins by matrix-assisted laser desorption/ionisation mass spectrometry. *J Mass Spectrom* 33:429–435.
- Fasoli E, Pastorello EA, Farioli L, Scibilia J, Aldini G, Carini M, Marocco A, Boschetti E, Righetti PG. 2009. Searching for allergens in maize kernels via proteomic tools. *J Proteomics* 72:501–510.
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246:64–71.
- Ferrante P, Masci S, D'Ovidio R, Lafiandra D, Volpi C, Mattei B. 2006. A proteomic approach to verify in vivo expression of a novel gamma-gliadin containing an extra cysteine residue. *Proteomics* 6:1908–1914.
- Ferranti P, Mamone G, Picariello G, Addeo F. 2007. Mass spectrometry analysis of gliadins in celiac disease. *J Mass Spectrom* 42:1531–1548.
- Finnie C, Svensson B. 2003. Barley Feasibility study of a tissue-specific approach to barley proteome analysis: Aleurone layer, endosperm, embryo and single seeds. *J Cereal Sci* 2:217–227.
- Finnie C, Svensson B. 2009. Barley seed proteomics from spots to structures. *J Proteomics* 13:315–324.
- Finnie C, Maeda K, Østergaard O, Bak-Jensen KS, Larsen J, Svensson B. 2004a. Aspects of the barley seed proteome during development and germination. *Biochem Soc Trans* 32: Part 3.
- Finnie C, Steenholdt T, Roda Noguera O, Knudsen S, Larsen J, Brinch-Pedersen H, Bach Holm P, Olsen O, Svensson B. 2004b. Environmental and transgene expression effects on the barley seed proteome. *Phytochemistry* 65:1619–1627.
- Finnie C, Bagge M, Steenholdt T, Østergaard O, Bak-Jensen KS, Backes G, Jensen A, Giese H, Larsen J, Roepstorff P, Svensson B. 2009. Integration of the barley genetic and seed proteome maps for chromosome 1H, 2H, 3H, 5H and 7H. *Funct Integr Genomics* 9:135–143.
- Foti S, Maccarrone G, Saletti R, Roepstorff P, Gilbert S, Tatham AS, Shewry PR. 2000. Verification of the cDNA deduced sequence of glutenin subunit 1Dx5 and an M_r 58000 repetitive peptide by matrix-assisted laser desorption/ionisation mass spectrometry (MALDI-MS). *J Cereal Sci* 31:173–183.
- Fournier ML, Gilmore JM, Martin-Brown SA, Washburn MP. 2007. Multidimensional separations-based shotgun proteomics. *Chem Rev* 107:3654–3686.
- Garozzo D, Cozzolino R, Di Giorgi S, Fisichella S, Lafiandra D. 1999. Use of hydroxyacetophenones as matrices for the analysis of high molecular weight glutenin mixtures by matrix-assisted laser desorption/ionisation mass spectrometry. *Rapid Commun Mass Spectrom* 13:2084–2089.
- Ge Y, Lawhorn BG, ElNaggar M, Strauss E, Park JH, Begley TP, McLafferty FW. 2002. Top down characterization of larger proteins (45kDa) by electron capture dissociation mass spectrometry. *J Am Chem Soc* 124:672–678.
- Gellrich C, Schieberle P, Wieser H. 2003. Biochemical characterization and quantification of the storage protein (secalin) types in rye flour. *Cereal Chem* 80:102–109.
- Gellrich C, Schieberle P, Wieser H. 2004a. Biochemical characterization of γ -75k secalins of rye. I. Amino acid sequences. *Cereal Chem* 81:290–295.
- Gellrich C, Schieberle P, Wieser H. 2004b. Biochemical characterization of γ -75K secalins of rye. II. Disulphide bonds. *Cereal Chem* 81:296–299.
- Gellrich C, Schieberle P, Wieser H. 2005. Studies of partial amino acid sequences of γ -40k secalins of rye. *Cereal Chem* 82:541–545.
- Gobaa S, Bancel E, Kleijer G, Stamp P, Branlard G. 2007. Effect of the 1BL.1RS translocation on the wheat endosperm, as revealed by proteomic analysis. *Proteomics* 7:4349–4357.
- Goff SA, Ricke D, Lan TH, Presting G, Wang RL, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchinson D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong JP, Miguel T, Paszkowski U, Zhang SP, Colbert M, Sun WL, Chen LL, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu YS, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalima T, Oliphant A, Briggs S. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp japonica). *Science* 296:92–100.
- Görg A, Postel W, Baumer M, Weiss W. 1992. Two-dimensional polyacrylamide gel electrophoresis, with immobilized pH gradients in the first dimension, of barley seed proteins: Discrimination of cultivars with different malting grades. *Electrophoresis* 13:192–203.
- Görg A, Weiss W, Dunn MJ. 2004. Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 4:3665–3685.
- Hager JW. 2002. A new linear ion trap mass spectrometer. *Rapid Commun Mass Spectrom* 16:512–526.
- Hajheidari M, Eivazi A, Buchanan BB, Won JH, Majidi I, Hosseini Salekdeh G. 2007. Proteomics uncovers a role for redox in drought tolerance in wheat. *J Proteome Res* 6:1451–1460.
- Heeren RM, Kleinnijenhuis AJ, McDonnell LA, Mize TH. 2004. A mini-review of mass spectrometry using high-performance FTICR-MS methods. *Anal Bioanalytical Chem* 378:1048–1058.
- Herman EM, Helm RM, Jung R, Kinney AJ. 2003. Genetic modification removes an immunodominant allergen from soybean. *Plant Physiol* 132:36–43.
- Hickman DR, Roepstorff P, Shewry PR, Tatham AS. 1995. Molecular weights of high molecular weight subunits of glutenin determined by mass spectrometry. *J Cereal Sci* 22:99–103.
- Houba R, Doekes G, Heederick DJJ. 1998. Occupational respiratory allergy in bakery workers: A review of the literature. *Am J Ind Med* 34:529–546.
- Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Cooks RG. 2005. The Orbitrap: A new mass spectrometer. *J Mass Spectrom* 40:430–443.
- Ikeda TM, Nagamine T, Fukuoka H, Yano H. 2002. Identification of new low-molecular-weight glutenin subunit genes in wheat. *Theor Appl Genet* 104:680–687.
- Issaq H, Veenstra T. 2008. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): Advances and perspectives. *Biotechniques* 44: 697–700.
- Jahn-Schmid B, Harwanegg C, Hiller R, Bohle B, Ebner C, Scheiner O, Mueller MW. 2003. Allergen microarray: Comparison of microarray using recombinant allergens with conventional diagnostic methods to detect allergen-specific serum immunoglobulin E. *Clin Exp Allergy* 33:1443–1449.
- Johansson E, Henriksson P, Svensson G, Heneen WK. 1993. Detection, chromosomal location and evaluation of the functional value of a novel high M_r glutenin subunit found in Swedish wheats. *J Cereal Sci* 17:237–245.
- Jonscher KR, Yates JR III. 1997. The quadrupole ion trap mass spectrometer—A small solution to a big challenge. *Anal Biochem* 244: 1–15.
- Junqueira M, Spirin V, Balbuena TS, Thomas H, Adzhubei I, Sunyaev S, Shevchenko A. 2008. Protein identification pipeline for the homology-driven proteomics. *J Proteomics* 71:346–356.
- Kagnoff MF. 2007. Celiac disease: Pathogenesis of a model immunogenetic disease. *J Clin Invest* 117:41–49.
- Karas M, Hillenkamp F. 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60:2299–2301.
- Kelleher NL, Lin HY, Valaskovic GA, Aaserud DJ, Fridriksson EK, McLafferty FW. 1999. Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *J Am Chem Soc* 121:806–812.

- Kreis M, Forde BG, Rahman S, Mifflin BJ, Shewry PR. 1985. Molecular evolution of the seed storage proteins of barley, rye and wheat. *J Mol Biol* 183:499–502.
- Kuiper HA, Kleter GA, Noterborn HPJM, Kok EJ. 2001. Assessment of the food safety issues related to genetically modified foods. *Plant J* 27:503–528.
- Laino P, Shelton D, Finnie C, De Leonardis AM, Mastrangelo AM, Svensson B, Lafiandra D, Masci S. 2010. Comparative proteome analysis of metabolic proteins from seeds of durum wheat (cv. Svevo) subjected to heat stress. *Proteomics* 10:2359–2368.
- Lane B, Bernier F, Dratewka-Kos E, Shafai R, Kennedy T, Pyne C, Munro JR, Vaughan T, Walters D, Altomare F. 1991. Homologies between members of the germin gene family in hexaploid wheat and similarities between these wheat germinals and certain *Physarum* spherulins. *J Biol Chem* 266:10461–10469.
- Lawrence MC. 1999. Structural relationships of 7S and 11S globulins. In: Shewry PR, Cases R, editors. *Seeds proteins*. Dordrecht: Kluwer Academic Publisher. pp. 517–541.
- Lee DG, Ahsan N, Lee SH, Kang KY, Bahk JD, Lee IJ, Lee BH. 2007. A proteomic approach in analyzing heat-responsive proteins in rice leaves. *Proteomics* 7:3369–3383.
- Lee JY, Yoon S, Ye YM, Hur GY, Kim S, Park HS. 2008. Gliadin-specific IgE in wheat-dependent exercise-induced anaphylaxis. *Allergy Asthma Proc* 29:614–621, and references therein.
- Lester PJ, Hubbard SJ. 2002. Comparative bioinformatic analysis of complete proteomes and protein parameters for cross-species identification in proteomics. *Proteomics* 2:1392–1405.
- Li K, Xu C, Zhang K, Yang A, Zhang J. 2007. Proteomic analysis of roots growth and metabolic changes under phosphorus deficit in maize (*Zea mays* L.) plants. *Proteomics* 7:1501–1512.
- Lin D, Tabb DL, Yates JR III. 2003. Large-scale protein identification using mass spectrometry. *Biochim Biophys Acta* 1646:1–10.
- Lin SK, Chang MC, Tsai YG, Lur HS. 2005. Proteomic analysis of the expression of proteins related to rice quality during caryopsis development and the effect of high temperature on expression. *Proteomics* 5:2140–2156.
- Lindsay MP, Skeritt JH. 1999. The glutenin macropolymer of wheat flour doughs: Structure–function perspectives. *Trends Food Sci Technol* 10:247–253.
- Liska AJ, Shevchenko A. 2003. Expanding the organismal scope of proteomics: Cross-species protein identification by mass spectrometry and its implications. *Proteomics* 3:19–28.
- Lo Conte L, Brenner SE, Hubbard T, Chothia C, Murzin AG. 2002. SCOP database in 2002: Refinements accommodate structural genomics. *Nucleic Acids Res* 30:264–267.
- Loboda AV, Krutchinsky AN, Bromirski M, Ens W, Standing KG. 2000. A tandem quadrupole/time-of-flight mass spectrometer with a matrix-assisted laser desorption/ionization source: Design and performance. *Rapid Commun Mass Spectrom* 14:1047–1057.
- Majoul T, Bancel E, Tribou E, Ben Hamida J, Branlard G. 2003. Proteomic analysis of the effect of heat stress on hexaploid wheat grain: Characterization of heat-responsive proteins from total endosperm. *Proteomics* 3:175–183.
- Majoul T, Bancel E, Tribou E, Ben Hamida J, Branlard G. 2004. Proteomic analysis of the effect of heat stress on hexaploid wheat grain: Characterization of heat-responsive proteins from non-prolamins fraction. *Proteomics* 4:505–513.
- Mak Y, Skylas DJ, Willows R, Connolly A, Cordwell SJ, Wrigley CW, Sharp PJ, Copeland L. 2006. A proteomic approach to the identification and characterisation of protein composition in wheat germ. *Integr Genomics* 6:322–337.
- Makarov A. 2000. Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. *Anal Chem* 72:1156–1162.
- Makarov A, Denisov E, Kholomeev A, Balschun W, Lange O, Strupat K, Horning S. 2006. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal Chem* 78:2113–2120.
- Mamone G, Ferranti P, Chianese L, Scafuri L, Addeo F. 2000. Qualitative and quantitative analysis of wheat gluten proteins by liquid chromatography and electrospray mass spectrometry. *Rapid Commun Mass Spectrom* 14:897–904.
- Mamone G, Ferranti P, Melck D, Tafuro F, Longobardo L, Chianese L, Addeo F. 2004. Susceptibility to transglutaminase of gliadin peptides predicted by a mass spectrometry-based assay. *FEBS Lett* 562:177–182.
- Mamone G, Addeo F, Chianese L, Di Luccia A, De Martino A, Nappo A, Formisano A, De Vivo P, Ferranti P. 2005. Characterization of wheat gliadin proteins by combined two-dimensional gel electrophoresis and tandem mass spectrometry. *Proteomics* 5:2859–2865.
- Mamone G, De Caro S, Di Luccia A, Addeo F, Ferranti P. 2009. Proteomic-based analytical approach for the characterization of glutenin subunits in durum wheat. *J Mass Spectrom* 44(12):1709–1723.
- Margiotta B, Urbano M, Colaprico G, Johansson E, Buonocore F, D'Ovidio R, Lafiandra D. 1996. Detection of γ -type subunit at the *Glu-A1* locus in some Swedish bread wheat lines. *J Cereal Sci* 23:203–211.
- Marti T, Molberg O, Li Q, Gray GM, Khosla C, Sollid LM. 2005. Prolyl endopeptidase-mediated destruction of T cell epitopes in whole gluten: Chemical and immunological characterization. *J Pharmacol Exp Ther* 312:19–26.
- Masci S, Lafiandra D, Porceddu E, Lew EJ-L, Tao HP, Kasarda DD. 1993. D-glutenin subunits: N-terminal sequence and evidence for the presence of cysteine. *Cereal Chem* 70:581–585.
- Masci S, Lew EJL, Lafiandra D, Porceddu E, Kasarda DD. 1995. Characterization of low-molecular-weight glutenin subunits in durum wheat by RP-HPLC and N-terminal sequencing. *Cereal Chem* 72:100–104.
- Masci S, D'Ovidio R, Lafiandra D, Kasarda DD. 1998. Characterization of a low-molecular-weight glutenin subunit gene from bread wheat and the corresponding protein that represents a major subunit of the glutenin polymer. *Plant Physiol* 118:1147–1158.
- Masci S, Rovelli L, Kasarda DD, Vensel WH, Lafiandra D. 2002. Characterisation and chromosomal localisation of C-type low-molecular-weight glutenin subunits in the bread wheat cultivar Chinese Spring. *Theor Appl Genet* 104:422–428.
- Mathesius U, Imin N, Chen H, Djordjevic MA, Weinman JJ, Natera SH, Morris AC, Kerim T, Paul S, Menzel C, Weiller GF, Rolfe BG. 2002. Evaluation of proteome reference maps for cross-species identification of proteins by peptide mass fingerprinting. *Proteomics* 2:1288–1303.
- Matthews DE, Croll VL, Lazo GR, Anderson OD. 2003. GrainGenes, the genome database for small-grain crops. *Nucleic Acids Res* 31:183–186.
- McDonald WH, Yates JR III. 2003. Shotgun proteomics: Integrating technologies to answer biological questions. *Curr Opin Mol Ther* 5:302–309.
- McLafferty FW, Breuker K, Jin M, Han X, Infusini G, Jiang H, Kong X, Begley TP. 2007. Top-down MS, a powerful complement to the high capabilities of proteolysis proteomics. *FEBS J* 274:6256–6268.
- Medzhradszky KF, Campbell JM, Baldwin MA, Falick AM, Juhasz P, Vestal ML, Burlingame AL. 2000. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal Chem* 72:552–558.
- Merlino M, Leroy P, Chambon C, Branlard G. 2009. Mapping and proteomic analysis of albumin and globulin proteins in hexaploid wheat kernels (*Triticum aestivum* L.). *Theor Appl Genet* 118:1321–1337.
- Mills EN, Jenkins JA, Alcocer MJ, Shewry PR. 2004. Structural, biological, and evolutionary relationships of plant food allergens sensitizing via the gastrointestinal tract. *Crit Rev Food Sci Nutr* 44:379–407.
- Molberg O, Mcadam SN, Körner R, Quarsten H, Kristiansen C, Madsen L, Fugger L, Scott H, Norén O, Roepstorff P, Lundin KE, Sjöström H, Sollid LM. 1998. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nat Med* 4(6):713–717.

- Montrichard F, Alkhalfioui F, Yano H, Vensel WH, Hurkman WJ, Buchanan BB. 2009. Thioredoxin targets in plants: The first 30 years. *J Proteomics* 72:452–474.
- Morita E, Matsuo H, Mihara S, Morimoto K, Savage AW, Tatham AS. 2003. Fast omega-gliadin is a major allergen in wheat-dependent exercise-induced anaphylaxis. *J Dermatol Sci* 33:99–104.
- Muccilli V, Cunsolo V, Saletti R, Foti S, Masci S, Lafiandra D. 2005. Characterization of B- and C-type low molecular weight glutenin subunits by electrospray ionization mass spectrometry and matrix-assisted laser desorption/ionization mass spectrometry. *Proteomics* 5:719–728.
- Muccilli V, Cunsolo V, Saletti R, Foti S, Margiotta B, Scossa F, Masci S, Lafiandra D. 2010. Characterization of a specific class of typical low molecular weight glutenin subunits of durum wheat by a proteomic approach. *J Cereal Sci* 51:134–139.
- Murzin AG, Brenner SE, Hubbard T, Chothia C. 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* 247:536–540.
- Nakagami H, Sugiyama N, Mochida K, Daudi A, Yoshida Y, Toyoda T, Tomita M, Ishihama Y, Shirasu K. 2010. Large-scale comparative phosphoproteomics identifies conserved phosphorylation sites in plants. *Plant Physiol* 153:1161–1174.
- Navrot N, Finnie C, Svensson B, Häggglund P. 2011. Plant redox proteomics. *J Proteomics* DOI: 10.1016/j.jprot.2011.03.008.
- Ocaña MF, Fraser PD, Patel RK, Halket JM, Bramley PM. 2007. Mass spectrometric detection of CP4 EPSPS in genetically modified soya and maize. *Rapid Commun Mass Spectrom* 21:319–328.
- OECD safety evaluation of foods derived by modern biotechnology: Concepts and principles; 1993; <http://www.oecd.org/dataoecd/57/53/1946129.pdf>.
- Osborne TB. 1924. The vegetable proteins, 2nd edition. London: Longmans Green and Co. pp. 153–154.
- Østergaard O, Finnie C, Laugesen S, Roepstorff P, Svensson B. Proteome analysis of barley seeds: Identification of major proteins from two-dimensional gels (pI 4–7). *Proteomics* 2004. 4:2437–2447.
- Parks BA, Jiang L, Thomas PM, Wenger CD, Roth MJ, Boyne II MT, Burke PV, Kwast KE, Kelleher NL. 2007. Top-down proteomics on a chromatographic time scale using linear ion trap Fourier transform hybrid mass spectrometers. *Anal Chem* 79:7984–7991.
- Pastorello EA, Farioli L, Conti A, Pravettoni V, Bonomi S, Iametti S, Fortunato D, Scibilia J, Bindsløv-Jensen C, Ballmer-Weber B, Robino AM, Ortolani C. 2007. Wheat IgE-mediated food allergy in European patients: α -Amylase inhibitors, lipid transfer proteins and low-molecular-weight glutenins. *Int Arch Allergy Immunol* 144:10–22.
- Pastorello EA, Farioli L, Pravettoni V, Scibilia J, Conti A, Fortunato D, Borgonovo L, Bonomi S, Primavesi L, Ballmer-Weber B. 2009. Maize food allergy: Lipid-transfer proteins, endochitinases, and α -zein precursor are relevant maize allergens in double-blind placebo-controlled maize-challenge-positive patients. *Anal Bioanalytical Chem* 395:93–102.
- Patrie SM, Charlebois JP, Whipple D, Kelleher NL, Hendrickson CL, Quinn JP, Marshall AG, Mukhopadhyay B. 2004. Construction of a hybrid quadrupole/Fourier transform ion cyclotron resonance mass spectrometer for versatile MS/MS above 10 kDa. *J Am Soc Mass Spectrom* 15:1099–1108.
- Patterson SD, Aebersold RH. 2003. Proteomics: The first decade and beyond. *Nat Genet* 33:311–323.
- Payne PI, Corfield KG. 1979. Subunit composition of wheat glutenin proteins, isolated by gel filtration in a dissociating medium. *Planta* 145:83–88.
- Payne PI, Lawrence GD. 1983. Catalogue of alleles for the complex gene loci Glu-A1, Glu-B1 and Glu-D1 which code for high molecular weight subunits of glutenin in hexaploid wheat. *Cereal Res Commun* 11:29–35.
- Payne PI, Corfield KG, Blackman JA. 1981. Correlation between the inheritance of certain high-molecular-weight subunits of glutenin and bread-making quality in progenies of six crosses of bread. 32: 51–60.
- Payne PI, Holt LM, Jarvis MG, Jackson EA. 1985. Two-dimensional fractionation of the endosperm proteins of bread wheat (*Triticum aestivum*): Biochemical and genetic studies. *Cereal Chem* 62:319–326.
- Payne PI, Nightingale MA, Krattiger AF, Holt LM. 1987. The relationship between HMW glutenin subunit composition and the bread-making quality of British-grown wheat varieties. *J Sci Food Agric* 40:51–65.
- Perdivara I, Deterding LJ, Przybylski M, Tomer KB. Mass spectrometric identification of oxidative modifications of tryptophan residues in proteins: Chemical artifact or post-translational modification? *J Am Soc Mass Spectrom* 2010. 21:1114–1117.
- Qian Y, Preston K, Krokhin O, Mellish J, Ens W. 2008. Characterization of wheat gluten proteins by HPLC and MALDI TOF mass spectrometry. *J Am Soc Mass Spectrom* 19:1542–1550.
- Qureshi MI, Qadir S, Zolla L. 2007. Proteomics-based dissection of stress-responsive pathways in plants. *J Plant Physiol* 164:1239–1260.
- Rehulková H, Marchetti-Deschmann M, Pittenauer E, Allmaier G, Rehulka P. 2009. Improved identification of hordeins by cysteine alkylation with 2-bromoethylamine, SDS-PAGE and subsequent in-gel tryptic digestion. *J Mass Spectrom* 44(11):1613–1621.
- Reid GE, McLuckey SA. 2002. ‘Top down’ protein characterization via tandem mass spectrometry. *J Mass Spectrom* 37:663–675.
- Rinalducci S, Murgiano L, Zolla L. 2008. Redox proteomics: Basic principles and future perspectives for the detection of protein oxidation in plants. *J Exp Botany* 59(14):3781–3801.
- Rocher A, Calero M, Soriano F, Mendez E. 1996. Identification of major rye secalins as coeliac immunoreactive proteins. *Biochim Biophys Acta* 1295:13–22.
- Salcedo G, Sánchez-Monge G, García-Casado G, Armentia A, Gomez L, Barber D. 2004. The cereal α -amylase/trypsin inhibitors family associated with bakers’ asthma and food allergy. In: Mills ENC, Shewry PR, editors. *Plant food allergens*. Oxford: Blackwell Science Ltd. pp. 70–86.
- Salplachta J, Bobálová J. 2009. MALDI-TOF mass spectrometry of hordeins rapid approach for identification of malting barley varieties. *J Mass Spectrom* 44:1287–1292.
- Savage AWJ, Saletti R, Foti S, Shewry PR, Tatham AS. 2002. Determination of molecular weights of C hordeins by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). *Cereal Chem* 79:768–770.
- Schwartz JC, Senko MW, Syka JE. 2002. A two-dimensional quadrupole ion trap mass spectrometer. *J Am Soc Mass Spectrom* 13:659–669.
- Scossa F, Laudencia-Chinguanco D, Anderson OD, Vensel WH, Lafiandra D, D’Ovidio R, Masci R. 2008. Comparative proteomic and transcriptional profiling of a bread wheat cultivar and its derived transgenic line overexpressing a low molecular weight glutenin subunit gene in the endosperm. *Proteomics* 8:2948–2966.
- Shan L, Molberg Ø, Parrot I, Hausch F, Filiz F, Gray GM, Sollid LM, Khosla C. 2002. Structural basis for gluten intolerance in celiac sprue. *Science* 297(5590):2275–2279.
- Shewry PR, Lucas JA. 1997. Plant proteins that confer resistance to pests and pathogens. In: Callow J, editor. *Advances in botanical research*. Vol. 26. USA: Academic Press. pp. 135–192.
- Shewry PR, Pandya MJ. 1999. The 2S albumins storage proteins. In: Shewry PR, Cases R, editors. *Seeds proteins*. Dordrecht: Kluwer Academic Publisher. pp. 563–586.
- Shewry PR, Tatham AS. 1990. The prolamin storage proteins of cereal seeds structure and evolution. *Biochem J* 267:1–12.
- Shewry PR, Tatham AS. 1997. Disulphide bonds in wheat gluten proteins. *J Cereal Sci* 25:207–227, and references therein.
- Shewry PR, Halford NG, Tatham AS. 1989. The high molecular weight subunits of wheat, barley and rye: Genetics, molecular biology, chemistry and role in wheat gluten structure and functionality. In: Mifflin BJ, editor. *Oxford surveys of plant molecular and cell biology*. Vol. 6. UK: Oxford University Press. pp. 163–219.

- Shewry PR, Halford NG, Tatham AS. 1992. High molecular weight subunits of wheat glutenin. *J Cereal Sci* 15:105–120.
- Shewry PR, Napier JA, Tatham AS. 1995. Seed storage proteins: Structures and biosynthesis. *Plant Cell* 7:945–956.
- Shin DH, Lee JY, Hwang KY, Kim KK, Sun SW. 1995. High-resolution crystal structure of the non-specific lipid-transfer protein from maize seedling. *Structure* 3:189–199.
- Shukla AK, Futrell JH. 2000. Tandem mass spectrometry: Dissociation of ions by collisional activation. *J Mass Spectrom* 35:1069–1090.
- Skylas DJ, Copeland L, Rathmell WG, Wrigley CW. 2001. The wheat-grain proteome as a basis for more efficient cultivar identification. *Proteomics* 1:1542–1546.
- Skylas DJ, Cordwell SJ, Hains PG, Larsen MR, Basseal DJ, Walsh BJ, Blumenthal C, Rathmell W, Copeland L, Wrigley CV. 2002. Heat shock of wheat during grain filling: Proteins associated with heat-tolerance. *J Cereal Sci* 35:175–188.
- Somers DA, Makarevitch I. 2004. Transgene integration in plants: Poking or patching holes in promiscuous genome. *Curr Opin Biotechnol* 15:126–131.
- Sotkovský P, Hubálek M, Hernychová L, Novák P, Havranová M, Setinová I, Kitanovicová A, Fuchs M, Stulík J, Tucková L. 2008. Proteomic analysis of wheat proteins recognized by IgE antibodies of allergic patients. *Proteomics* 8:1677–1691.
- Sule A, Vanrobayeys F, Hajo G, Van beumen J, Devreese B. 2004. Proteomic analysis of small heat shock protein isoforms in barley shoots. *Phytochemistry* 65:1853–1863.
- Sun W, Van Montagu M, Verbruggen N. 2002. Small heat shock proteins and stress tolerance in plants. *Biochim Biophys Acta* 1557:1–9.
- Syka JE, Marto JA, Bai DL, Horning S, Senko MW, Schwartz JC, Ueberheide B, Garcia B, Busby S, Muratore T, Shabanowitz J, Hunt DF. 2004a. Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications. *J Proteome Res* 3:621–626.
- Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF. 2004b. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci USA* 101:9528–9533.
- Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T. 1988. Protein and polymer analyses up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 2:151–154.
- Tatham AS, Shewry PR. 2008. Allergens in wheat and related cereals. *Clin Exp Allergy* 38:1712–1726.
- Timperio AM, Egidi MG, Zolla L. 2008. Proteomics applied on plant abiotic stresses: Role of heat shock proteins (HSP). *J Proteomics* 71:391–411.
- Vensel WH, Dupont FM, Chan R, Hurkman WJ. 2007. Mass spectrometry based identifications of LMW glutenin subunits. In: Lookhart GL, Ng PKW, editors. *Gluten proteins 2006*. Minneapolis, MN: AACCI. pp. 347–351.
- Vensel WH, Dupont FM, Sloane S, Altenbach SB. 2011. Effect of cleavage enzyme, search algorithm and decoy database on mass spectrometric identification of wheat gluten proteins. *Phytochemistry* 72:1154–1161.
- Waines JG, Payne PI. 1987. Electrophoretic analysis of the highmolecular-weight glutenin subunits of *Triticum monococcum*, *T. urartu*, and the A genome of bread wheat (*T. aestivum*). *Theor Appl Genet* 74:71–76.
- Washburn MP, Wolters D, Yates JR. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19:242–247.
- Whitelegge JP. 2004. Mass spectrometry for high throughput quantitative proteomics in plant research: Lessons from thylakoid membranes. *Plant Physiol Biochem* 42:919–927.
- Wilkins MR, Williams KL. 1997. Cross-species protein identification using amino acid composition, peptide mass fingerprinting, isoelectric point and molecular mass: A theoretical evaluation. *J Theor Biol* 186:7–15.
- Witzel K, Weidner A, Surabhi GK, Borner A, Mock HP. 2009. Salt stress-induced alterations in the root proteome of barley genotypes with contrasting response towards salinity. *J Exp Botany* 60:3545–3557.
- Wöhrl S, Vigl K, Zehetmayer S, Hiller R, Jarisch R, Prinz M, Stingl G, Kopp T. 2006. The performance of a component-based allergen-microarray in clinical practice. *Allergy* 61:633–639.
- Wong JH, Cai N, Balmer Y, Tanaka CK, Vensel WH, Hurkman WJ, Buchanan BB. 2004. Thioredoxin targets of developing wheat seeds identified by complementary proteomic approaches. *Phytochemistry* 65:1629–1640.
- Xiong L, Schumaker KS, Zhu JK. 2002. Cell signaling for cold, drought, and salt stresses. *Plant Cell* 14 (Suppl): S165–S183.
- Yagami T, Haishima Y, Tsuchiya T, Tomitaka-Yagami A, Kano H, Matsunaga K. 2004. Proteomic analysis of putative latex allergens. *Int Arch Allergy Appl Immunol* 135:3–11.
- Yu J, Hu SN, Wang J, Wong GKS, Li SG, Liu B, Deng YJ, Dai L, Zhou Y, Zhang XQ, Cao ML, Liu J, Sun JD, Tang JB, Chen YJ, Huang XB, Lin W, Ye C, Tong W, Cong LJ, Geng JN, Han YJ, Li L, Li W, Hu GQ, Huang XG, Li WJ, Li J, Liu ZW, Li L, Liu JP, Qi QH, Liu JS, Li L, Li T, Wang XG, Lu H, Wu TT, Zhu M, Ni PX, Han H, Dong W, Ren XY, Feng XL, Cui P, Li XR, Wang H, Xu X, Zhai WX, Xu Z, Zhang JS, He SJ, Zhang JG, Xu JC, Zhang KL, Zheng XW, Dong JH, Zeng WY, Tao L, Ye J, Tan J, Ren XD, Chen XW, He J, Liu DF, Tian W, Tian CG, Xia HG, Bao QY, Li G, Gao H, Cao T, Wang J, Zhao WM, Li P, Chen W, Wang XD, Zhang Y, Hu JF, Wang J, Liu S, Yang J, Zhang GY, Xiong YQ, Li ZJ, Mao L, Zhou CS, Zhu Z, Chen RS, Hao BL, Zheng WM, Chen SY, Guo W, Li GJ, Liu SQ, Tao M, Wang J, Zhu LH, Yuan LP, Yang HM. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp indica). *Science* 296:79–92.
- Zabrouskov V, Whitelegge JP. 2007. Increased coverage in the transmembrane domain with activated-ion electron capture dissociation for top-down Fourier-transform mass spectrometry of integral membrane proteins. *J Proteome Res* 6:2205–2210.
- Zeller FJ, Hsam SLK. In: Sakamoto, S., editors. *Proceedings of the 6th International Wheat Genetics Symposium*, Kyoto University, Kyoto 1983, pp. 161–173.
- Zolla L, Rinalducci S, Antonioli P, Righetti PG. 2008. Proteomics as a complementary tool for identifying unintended side effects occurring in transgenic maize seeds as a result of genetic modifications. *J Proteome Res* 7:1850–1861.
- Zubarev R. 2006. Protein primary structure using orthogonal fragmentation techniques in Fourier transform mass spectrometry. *Expert Rev Proteomics* 3:251–261.
- Zubarev RA, Kelleher NL, McLafferty FW. 1998. Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J Am Chem Soc* 120:3265–3266.