UNIVERSITÁ DEGLI STUDI DI PARMA

Dottorato di ricerca in SCIENZE E TECNOLOGIE ALIMENTARI

Ciclo XXVI

WHEAT ALLERGIES: A PEPTIDOMIC APPROACH

Coordinatore:

Chiar.mo Prof. FURIO BRIGHENTI

Tutor:

Chiar.mo Prof. STEFANO SFORZA

Chiar.mo Prof. GIANNI GALAVERNA

Dottorando: BARBARA PRANDI

SUMMARY

1 ABSTRACT

This thesis deals with the characterization of the peptides obtained after simulated gastrointestinal digestion of wheat proteins, with specific focus on those known to be involved in celiac disease. This autoimmune entheropathy, besides the genetic predisposition, is triggered by gluten ingestion. Several studies correlated the increased prevalence of celiac disease with different factors, including the amount and quality of dietary gluten. In this thesis, different in vitro digestion models were applied to wheat samples and the resulting peptides where identified and quantified. The results highlighted that different wheat varieties, even with similar total gluten content, can lead to huge differences in terms of immunotoxic peptides generated after digestion. The genotypic variability of wheat gluten could thus be used in varietal selections aimed to reduce people exposure to immunotoxic peptides. Other allergenic wheat proteins, such as α-amylase/trypsin inhibitors, instead, were found to be more affected by environmental factors (such as growing area). Beside the immunological point of view, in this thesis several analytical methods applied to wheat proteins were also developed using LC-MS and LC-MS/MS techniques. In particular, methods for detection and quantification of an α-amylase/trypsin inhibitor and gluten were developed, , as well as a method to detect common wheat adulteration in durum wheat samples.

RIASSUNTO

Questa tesi riguarda la caratterizzazione dei peptidi ottenuti dopo digestione gastrointestinale simulata delle proteine del frumento, con particolare attenzione rivolta a quelli noti per essere coinvolti nella celiachia. Questa enteropatia autoimmune, oltre alla predisposizione genetica, è scatenata dall'ingestione di glutine. Diversi studi hanno messo in relazione l'aumento della prevalenza della celiachia con alcuni fattori, tra cui la quantità e la tipologia del glutine ingerito con la dieta. In questa tesi, diversi modelli di digestione in vitro sono stati applicati a campioni di frumento e i peptidi generati sono stati identificati e quantificati. I risultati hanno evidenziato che diverse linee di frumento, anche con un contenuto di glutine molto simile, possono portare a drastiche differenze in termini di peptidi immunotossici generati dopo digestione. La variabilità genotipica del glutine di frumento potrebbe essere sfruttata quindi per effettuare una selezione varietale volta a diminuire l'esposizione della popolazione ai peptidi immunotossici. Altre proteine allergeniche del frumento, come gli inibitori della tripsina/ α-amilasi, invece, si sono dimostrate essere più influenzate da fattori ambientali (come l'area di coltivazione). Oltre al punto di vista immunologico, in questa tesi sono stati anche sviluppati alcuni metodi analitici per le proteine del frumento, utilizzando tecniche LC-MS e LC-MS/MS. In particolare, sono stati sviluppati metodi per la rivelazione e quantificazione di un inibitore della tripsina/α-amilasi e del glutine, così come è stato sviluppato un metodo per la rivelazione di adulterazioni da grano tenero in farine di grano duro.

2 GENERAL INTRODUCTION

2.1 WHEAT AND WHEAT PROTEINS

2.1.1 TAXONOMY

Wheat is a cereal of the genus *Triticum*, belonging to the Poaceae family.[1] The genome of the genus Triticum is composed by seven chromosomes that can be present in two, four or six copies, based on the ploidy level of the species.[2] Among the six different *Triticum* species, three are relevant for human consumption:

- ✓ *Triticum monococcum*, also called einkorn, is a diploid species (AA genome) thought to be the first wheat species to be domesticated by humans around 7500 a.C., in the Middle East.[3]
- ✓ *Triticum turgidum* is a tetraploid species (genome AABB) widely harvested for flour production. The subspecies of food interest are *Triticum turgidum* spp *dicoccum* (commercialized with the name spelt), *Trititicum turgidum* spp *turanicum* (comprising the variety present in the market as Kamut®) and, most of all, *Triticum turgidum* spp *durum* (widely used in the Mediterrean area for the production of pasta, bread and cous cous).
- ✓ *Triticum aestivum*, also called common wheat, has a hexaploid genome (AABBDD) and it is the most important wheat species worldwide, used in a huge variety of baked products, pasta but also as thickening agent in sauces and creams.

2.1.2 DIFFUSION AND USES

Wheat is the third cereal for diffusion and production for human consumption, after corn and rice, occupying about 22% of the total cultivated area in the world. The major producers of common wheat are the United States and Canada in America, Pakistan, India, Kazakhistan and Russia in Asia and Australia. Smaller crops are also found in Europe, South America, Africa and China.[4] Most of the wheat production (about two third) is intended for food production for a wide variety of products. Rarely wheat kernels are used as such for food production (only in some breakfast cereals), in the majority of cases kernels are ground to obtain flours. On the basis of their protein content flours are used for different purposes: weak flours (with low protein level) are used for cakes, biscuits and tarts; medium flours are used for most foods (crackers and white bread) and strong flours (with high protein content) are mainly used for hearth breads and whole wheat breads.[5]

Durum wheat is less diffuse worldwide than common wheat, but it is very important for some economies of the Mediterranean basin (Italy, Spain and Greece in Europe; Turkey in the Middle East) and Canada and United States in America.[6] Doubtless, the intended use of durum wheat is the pasta production (especially dried pasta). The flour obtained from durum wheat is called semolina and, beside pasta production, it is also used for different North African dishes such as cous cous and bulghur. Traditional breads made from durum wheat are very diffused in Southern Italy and in the Middle East.

2.1.3 DURUM WHEAT COMPOSITION

The wide diffusion of wheat for human consumption is mainly due to its high carbohydrate content (Table 2.1) that constitute about 70% of the weight. Quite all the carbohydrate fraction of durum wheat is constituted of starch (68% on the total weight),[7] with only a low percentage of free sugars, that shows a slight increase during

flour storage due to amylases action. Starch is located in the endosperm of the kernel, where the starchy granule are surrounded from a protein matrix. Durum wheat starch is very digestible because it shows a very low percentage of resistant starch (around 0.6%),[8] thus being an important source of glucose for human diet.

In wheat, lipids form 1-2% of the endosperm, 8-15% of the germ, and about 6% of the bran with an average of 2-4% of the whole kernel.[9] Usually the lipid fraction in flour is negligible, because the germ is removed during milling in order to avoid oxidative processes that lead to flour rancidity during storage.

Table 2.1 Durum wheat composition divided in proximate, minerals, vitamins and lipids.[10]

NUTRIENT	UNIT	PER 100 g	NUTRIENT	UNIT	PER 100 g
Proximates			**Vitamins**		
Water	g	10.94	Thiamin	mg	0.419
Energy	kcal	339	Riboflavin	mg	0.121
Protein	g	13.68	Niacin	mg	6.738
Lipids	g	2.47	Pyridoxal phosphate	mg	0.419
Carbohydrates	g	71.13	Folates	µg	43
Minerals			**Lipids**		
Calcium	mg	34	Saturated fatty acids	g	0.454
Iron	mg	3.52	Monounsaturated fatty acids	g	0.344
Magnesium	mg	144	Polyunsaturated fatty acids	g	0.978
Phosphorus	mg	508	Saturated fatty acids	g	0.454
Potassium	mg	431	Monounsaturated fatty acids	g	0.344
Sodium	mg	2	Polyunsaturated fatty acids	g	0.978
Zinc	mg	4.16	Saturated fatty acids	g	0.454

The protein fraction of wheat is mainly localized in the endosperm and in the aleuronic layer of the kernel. Wheat proteins can be classified on the basis of their role in the plant in:

- ✓ Functional proteins, namely those that have an active role in the cell, such as enzymes, transporters and defence proteins. They can be further divided on the basis of their solubility in albumins (water soluble) and globulins (soluble in saline solutions). They constitute about 20% of total wheat proteins, with a ratio albumins/globulins of 2:1.
- ✓ Storage proteins, also called gluten proteins, that have a role as nitrogen reserve in the seeds, that will be used during the germination. Also gluten proteins can be further classified for their solubility properties in gliadins (alcohol soluble) and glutenins (soluble in presence of chaotropic and reducing agents). They constitute about 80% of wheat proteins, with a slight prevalence of glutenins.

2.1.3.1 GLUTEN PROTEINS

Wheat proteins have a poor biological value, due to the low presence of lysine and methionine. Much more important is the technological function of gluten proteins that need a separate discussion for their importance in

the rheological properties of wheat derived products. In fact, after kneading with water gluten, proteins can form a polymeric network (due to intermolecular disulphide bridges) that confers viscoelasticity and cohesiveness to the dough. In this way, the product can retain the air, formed for the kneading and/or leavening processes, during cooking, thus obtaining a porous structure, typical for bread and baked products.

As previously anticipated, gluten is a complex of proteins that can be classified in gliadin and glutenins. Both classes show a high content in the amino acids proline and glutamine (from which the name prolamines). The high presence of these two amino acids strongly contributes to the elasticity and toughness of the gluten: the proline residues induce a β-spiral conformation of the protein, stabilized by the hydrogen bonds generated among the glutamine lateral chains and by the intra and intermolecular disulphide bridges generated by the abundant cysteine residues.[11]

Gliadins are alcohol soluble monomeric proteins with a low molecular weight (ranging from 25 to 75 KDa) and can be further divided on the basis of their electrophoretical mobility in α-gliadin, γ-gliadin and ω-gliadins (from the fastest to the slowest).[12] α-gliadins and γ-gliadins contains different cysteine residues that form intramolecular disulphide bridges, while ω-gliadins are sulphur poor proteins.[13]

Glutenins are soluble only in presence of chaotropic and reducing agents, due to the presence of intermolecular disulphide bridges that confers them a polymeric structure, with a very high molecular weight (>100 KDa). Glutenins can be divided on the basis of their molecular weight in Low Molecular Weight (LMW) and High Molecular Weight (HMW).

2.2 FOOD ALLERGIES

Food allergies are non-toxic adverse reaction of the immune system against a certain food or food component (in most of the cases proteins). They do not occur in all people but only in some previously sensitized subjects. As recently endorsed by the European Academy of Allergy and Clinical Immunology, food allergies can be IgE mediated or not IgE associated and the difference between them lies in the type of antibody involved in the immune reaction:[14] while in IgE-mediated food allergies only Immunoglobulins E are involved, in not IgE associated food allergies IgA and/or IgG antibodies or by other cells of the immune system are involved.[15]

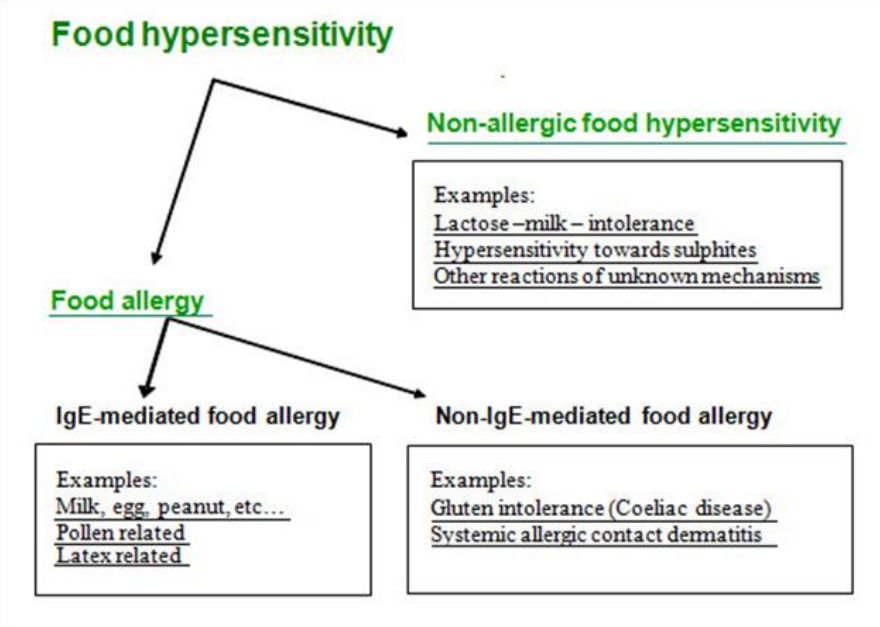

Figure 2.1 Classification of the different adverse reactions to food.[16]

The difference between food allergies and food intolerance lies in the involvement of the immune system. While this is always involved in food allergies, food intolerance follows other mechanism, such as enzymes failure (as happens in lactose intolerance) or hypersensitivity to certain food rich in biogenic amines such as histamine.

2.2.1 IgE-MEDIATED FOOD ALLERGIES

Food allergies are adverse immunological reaction that develops after the ingestion, by a sensitized subject, of a food containing the allergen, namely the antigen that will be recognized by the IgE antibodies. Usually the antigen is a protein or a class of proteins.

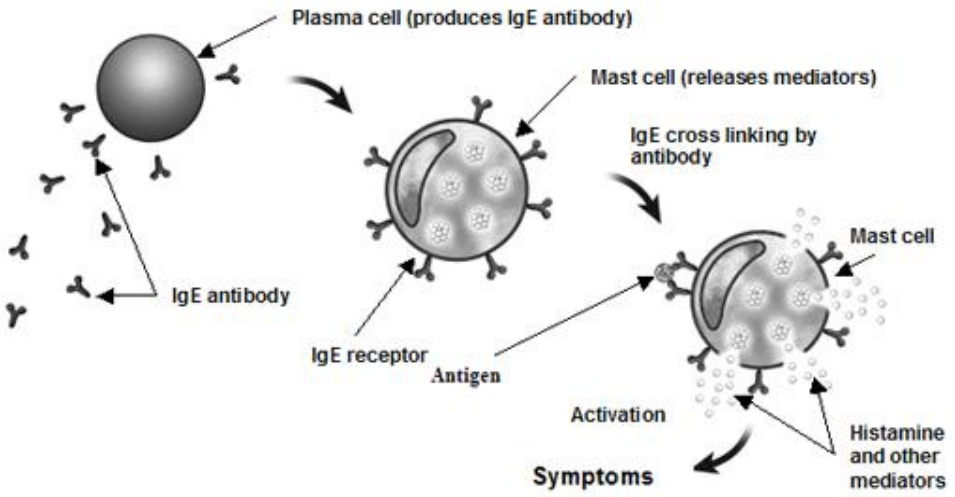

Figure 2.2 Schematic mechanism of an IgE mediated immunological reaction to a food allergen.[17]

IgE antibodies are usually produced by the organism as defence against parasites, but in some predisposed subjects they react towards some proteins, even present in foods.

For the predisposed subject it is necessary a previous sensitization phase in which the immune system is exposed to the antigen. The food allergen passes the intestinal barrier and is then recognized by the Antigen Presenting Cells (APC), generating a multiproteic complex called Major Hystocompatibility Complex (MHC). This causes the induction of cytokines production that stimulate B lymphocytes to produce specific IgE towards the target allergenic protein (Figure 2.2). The specific IgE will bind to the mast cells that are thus sensitized against the allergen.

Once the allergic subject is sensitized, a subsequent exposure to the allergen (reaction phase) will cause a strong immunological response: the allergen will be recognized directly by the sensitized mast cells that will degranulate releasing pro-inflammatory mediators such as histamine, heparin and cytokines that will cause the allergy symptoms. [18] The severity of symptoms can vary greatly among different patients and allergies; skin, gastrointestinal tract and/or respiratory system can be involved, until anaphylactic shock, which is the worst case.[19]

It is important to underline that not all the allergenic protein is involved in this immunological reaction but only some specific region (called epitopes), that are specifically recognized and bound by the IgE antibodies. Usually an allergen shows different epitopes within their sequence, which can be recognized by different antibodies. Two type of epitope are known:[20]

- ✓ Linear epitopes: the IgE recognize a specific amino acid sequence (usually 8-11 amino acids length). They are the most stable epitopes since only a proteolytic action can destroy them (thus leading to loss of allergenicity). To be able to trigger an allergic reaction, linear epitope must be stable to gastrointestinal digestion, meaning that they must be resistant to cleavage by the digestive proteases.
- ✓ Conformational epitopes: the recognition by the IgE antibodies can occur due to the tertiary structure of the epitope. Denaturation processes can lead to a loss of the allergenicity through modification of protein conformation, thus preventing the binding of the IgE.

The foods most frequently involved in allergies are milk, eggs, nuts, fish, celery, sesame and soy, as recently stated by 2007/68/EC directive (Annex IIIa), that introduced the lists all the allergenic foods that must be labelled as well as a few products derived from these foods for which allergen labelling is not required.

2.2.1.1 WHEAT IgE-MEDIATED ALLERGIES

The IgE mediated immunological reactions to wheat have a prevalence ranging from 0.4 to 1.3% in childrens[21] and, as it often happens for allergies, it is less widespread among adults ranging from 0.2% to 0.9%.[22]

Wheat allergies can be classified in three main groups:

- ✓ Food allergies properly called, that have an oral route of exposition and are due to a wide spectrum of proteins, such as non specific lipid transfer proteins (ns-LTPs), chloroform/methanol soluble proteins (CM proteins), gluten proteins, 2S albumins, ...[23]
- ✓ Wheat dependent exercise induced anaphylaxis (WDEIA) that occurs in allergic subjects after wheat consumption and a physical effort and it is mainly due to ω_5-gliadins[24]

- ✓ Baker's asthma that, unlike the previous ones, has a respiratory route of sensitization. The exposition in fact occurs through the inhalation of the flour dust. The main trigger of baker's asthma are α-amylase/trypsin inhibitors, a group of defense protein that are soluble in chloroform/methanol mixtures.[25]

2.2.2 FOOD CELL-MEDIATED ALLERGIES

Food cell-mediated allergies are adverse reaction to certain food components that occur in predisposed subject and, unlike IgE mediated food allergies, other antibodies or cells of the immune system (like T lymphocytes) are involved. The non-IgE mediated food allergies present in a more subacute or chronic nature and more commonly affect only the gastrointestinal tract. The primary disorders in this category include food protein-induced enterocolitis (FPIES), food protein-induced proctitis, enteropathy and celiac disease.[26] There are no medications currently recommended by the NIAID (National Institute of Allergy and Infectious Diseases) to prevent non-IgE mediated reactions. The only treatment is the avoidance of the triggering food.

2.2.2.1 CELIAC DISEASE

Celiac disease is an autoimmune entheropathy that occurs in celiac patients after the ingestion of wheat gluten or related proteins of barley (hordeins) or rye (secalins). Beside this environmental factor, a genetic predisposition is necessary to develop celiac disease. A particular HLA-DQ heterodimer, encoded by the DQA1*0501 and DQB1*0201 genes in *cis* or *trans* configuration, confers the primary disease susceptibility.[27] HLA-DQ2 is found in 95% and the related HLA-DQ8 in most of the remaining patients with celiac disease.[28, 29] The expression of these heterodimers is a necessary condition but not sufficient alone to trigger the disease. Other environmental factors must be present, such as high gluten exposure (especially during weaning), type and duration of wheat dough fermentation, the spectrum of intestinal microbiota and its changes over time, enteric infections, and stressors in general.[30]

After the ingestion of foods containing gluten or related proteins of barley or rye, some high molecular weight peptides are formed in the gastrointestinal tract. The high presence of proline in gluten proteins in fact leads to a poor digestibility of these proteins because the cyclic side chain of this amino acid hinder the peptidic bond at his N-term, blocking the access of proteolytic enzymes.[31] Some of these peptides contain sequences implied in celiac disease.

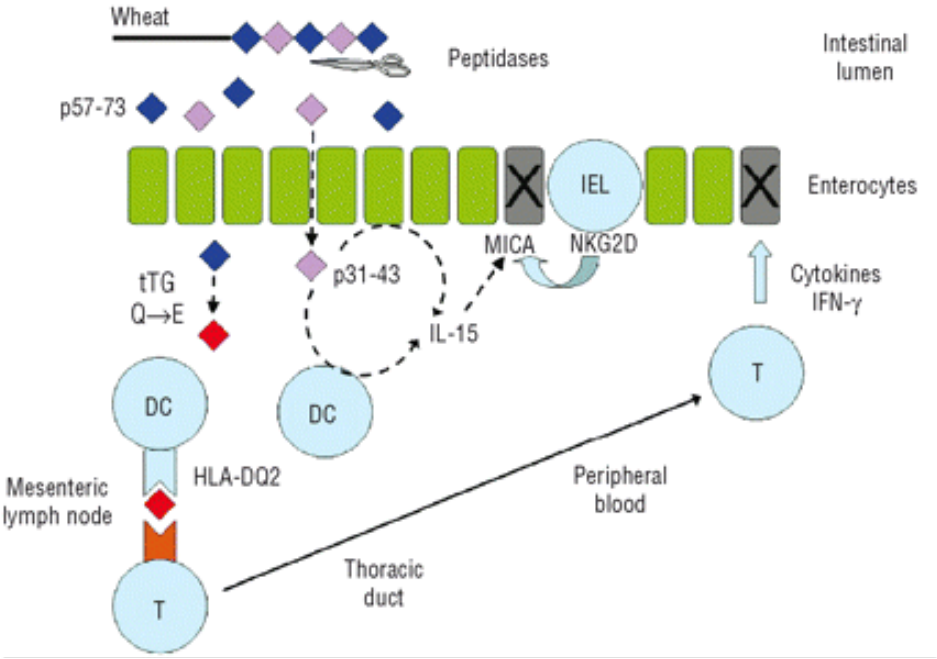

Figure 2.3 Innate and adaptive immune response mechanism in celiac disease.[32]

The immunological response to gluten in celiac disease patients involves both the adaptive and the immune system (Figure 2.3).

Adaptive immune response. Some gluten peptides can pass the intestinal epithelium taking advantages of some mucosal defects and reach the lamina propria. Here the glutamine residue (also very abundant in gluten proteins) are deamidated by the enzyme tissue transglutaminase to glutamic acid. The presence of negative charges on the peptides enhance the binding by the HLA-DQ2 (or DQ8) heterodimer on Antigen Presenting Cells (APC). This complex is recognized by the lymphocytes T, which, by secreting pro-inflammatory cytokines, are responsible for degradation of both extracellular matrix and basement membrane. Activated T cells also become able to trigger enterocyte apoptosis leading to the characteristic mucosal lesions. Finally, stimulated CD4+ T cells are also able to induce lymphocyte B differentiation into plasma cells producing specific antigliadin and anti-tissue transglutaminase antibodies (IgA and IgG type).[33]

Innate immune response. Other high molecular weight peptides generated from the incomplete gluten digestion contain sequences that elicit the dendritic cells to produce cytokines like interleukin-15 (IL-15). IL-15 stimulates intra-epithelial lymphocytes (IELs) to express NKG2D receptors and epithelial cells to express MICA molecules. Upon engagement of NKG2D receptor with MICA ligand, the IELs kill the epithelial cells causing the tissue destruction.[34]

These combined immunological response result in the typical destruction of the intestinal epithelial structure that occurs in celiac disease patients, with villous atrophy and crypts hyperplasia. Frequent in celiac disease patients are deficiencies of iron, folic acid, calcium and fat-soluble vitamins.[35] The disruption of the gut architecture leads to a wide spectrum of symptoms like diarrhoea, abdominal distension, vomiting, constipation, irritable bowel syndrome, dyspepsia or oesophageal reflux but also to systemic manifestation due to the decreased absorption

capacity of small intestine: short stature, anaemia, neurological symptoms, dermatitis herpetiformis, reduced bone density, infertility or autoimmune disorders.[36]

2.2.3 REGULATIONS FOR ALLERGENS

Given the seriousness of the problem, the European Commission has issued the Labelling Directive (Directive 2000/13/EC) concerns the labelling of foodstuffs to be delivered as such to the ultimate consumer and certain aspects relating to the presentation and advertising thereof. This Directive states that manufacturers must declare all ingredients present in pre-packaged foods sold in the EU; it was subsequently modified by the Directive 2003/89/EC and Directive 2007/68/EC) that introduced a list of the **allergenic ingredients that must be declared on the food label**. These ingredients can cause adverse health effects, and in some cases exposure to these can be fatal. These are the most common food allergens which are generally resistant to food processing and they have the capacity to trigger an allergic reaction in an allergic consumer if they are added to foods. Some of these allergens are very widely distributed all throughout Europe, while others, such as mustard and celery, are more geographically restricted. As stated in the EFSA Opinion of the Scientific Panel on Dietetic Products, Nutrition and Allergies on a request from the Commission relating to the evaluation of allergenic foods for labelling purposes, this list should be kept under review in the light of changing food practices and emergence of new clinical observations and other kind of scientific information. The list includes up to now 14 food ingredients that are reported in the following list:

1. Cereals containing gluten (i.e. wheat, rye, barley, oats, spelt, kamut or their hybridised strains) and products thereof, except wheat-based glucose syrups including dextrose, wheat-based maltodextrins, glucose syrups based on barley, cereals used for making distillates or ethyl alcohol of agricultural origin for spirit drinks and other alcoholic beverages.
2. Crustaceans and products thereof.
3. Eggs and products thereof.
4. Fish and products thereof, except fish gelatine used as carrier for vitamin or carotenoid preparations, fish gelatine or Isinglass used as fining agent in beer and wine.
5. Peanuts and products thereof.
6. Soybeans and products thereof, except fully refined soybean oil and fat, natural mixed tocopherols (E306), natural D-alpha tocopherol, natural D-alpha tocopherol acetate, natural D-alpha tocopherol succinate from soybean sources, vegetable oils derived phytosterols and phytosterol esters from soybean sources, plant stanol ester produced from vegetable oil sterols from soybean sources.
7. Milk and products thereof (including lactose), except whey used for making distillates or ethyl alcohol of agricultural origin for spirit drinks and other alcoholic beverages, lactitol.
8. Nuts, i.e. almonds (*Amygdalus communis* L.), hazelnuts (*Corylus avellana*), walnuts (*Juglans regia*), cashews (*Anacardium occidentale*), pecan nuts (*Carya illinoiesis* (Wangenh.) K. Koch), Brazil nuts (*Bertholletia excelsa*), pistachio nuts (*Pistacia vera*), macadamia nuts and Queensland nuts (*Macadamia ternifolia*), and products thereof, except nuts used for making distillates or ethyl alcohol of agricultural origin for spirit drinks and other alcoholic beverages.
9. Celery and products thereof.
10. Mustard and products thereof.

11. Sesame seeds and products thereof.
12. Sulphur dioxide and sulphites at concentrations of more than 10 mg/kg or 10 mg/litre expressed as SO_2.
13. Lupin and products thereof.
14. Molluscs and products thereof.

2.3 PROTEIN GASTROINTESTINAL DIGESTION

Understanding food digestion is very important for allergens management, since the physiological processing of food in our gastrointestinal tract can modify its allergenic potential.[37] The allergenic proteins contained in the food undergo to proteolytic cleavage by digestive enzymes, generating oligopeptides.[38] The resistance of the protein to digestion is a key factor in determining its allergenicity, because the part of allergen left intact or anyway in an immunologically active form, can be taken up by the gut and sensitize the mucosal immune system.[39]

The first step of food digestion is chewing, aimed to reduce the dimensions of food particles and thus achieve a better enzymatic action in the following digestion steps. During chewing begins also the action of ptyalin (or salivary amylase) that starts the digestion of cooked starch at an almost neutral pH. The so formed bolus then reaches the stomach, where the environment becomes strongly acidic due to the secretion of hydrochloric acid (HCl). Hydrochloric acid is necessary to activate the secreted pro-enzyme pepsinogen in pepsin (the active form). Protein digestion takes place mostly in the stomach: the acidic environment favours denaturation (with loss of the tertiary structure) and then exposure of the peptidic bonds to the action of pepsin. Pepsin is a proteolytic enzyme that cleaves peptides with an aromatic amino acid on either side of the peptide bond. Sulfur-containing amino acids increase susceptibility to hydrolysis when they are close to the peptide bond. Pepsin preferentially cleaves at the carboxyl side of phenylalanine and leucine and at the carboxyl side of glutamic acid residues. In mouth and stomach also lipase is present (lingual lipase and gastric lipase) that starts the lipids digestion. After the gastric phase the chime passes in the small intestine were the protein digestion continue with the action of pancreatic proteases: trypsin, chymotrypsin, elastase and carboxypeptidase.[40] Trypsin cleaves peptide chains mainly at the carboxyl side of the amino acids lysine or arginine. Chymotrypsin is a serine protease that hydrolyzes peptide bonds with aromatic or large hydrophobic side chains (Tyr, Trp, Phe, Met, Leu) on the carboxyl end of the bond. Elastase usually cleaves at the C-term of alanine residues. Trypsin, chymotrypsin and elastase are thus endoproteases, cleaving the peptidic bond of non terminal amino acids. Instead, carboxypeptidease is an exoprotease, cleaving the aminoacid at the C-term of peptides. The oligopeptides generated from protein digestion anyway cannot be absorbed as such, but they need to be further proteolyzed by brush border membrane enzymes such as amino-peptidase (an exoprotease cleaving the peptidic bond at the N-term of the peptide), aminooligopeptidase, dipeptidylaminopeptidase and angiotensin converting enzyme.[41] The amino acids, dipeptides and tripeptides generated can be thus absorbed by the enterocytes. In the small intestine the digestion of starch by pancreatic amylase and the lipids breakdown by pancreatic lipase also occurs, with the strong contribution of the bile salts, and their following absorption.

The ideal condition would be to test the allergen resistance in an *in vivo* model with a real food matrix, because at physiological level there are many parameters (salt concentrations, viscosity, enzyme:protein ratio, antinutritional factors, ...) that cannot be taken into account in any other manner. Anyway, *in vivo* model are very difficult to carry out both for the high cost and for ethical reasons. So, a lot of in vitro digestion methods were

developed for a wide variety of allergens; they can be classified in two basic groups: static and dynamic. Both consist usually of a salivar, a gastric and a duodenal phase; the difference between them consist in the removal of the digestion products and in the physical processes (mixing, shearing, ...) that are present only in dynamic models.[42]

2.4 LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY APPLIED TO PEPTIDE ANALYSIS

2.4.1 LIQUID CHROMATOGRAPHY

Peptides are organic molecules formed by two or more (up to 50 or even more) amino acids bound together through an amidic bond. The presence of charged groups (N- and C-term and the side chain of basic and acid amino acids) makes peptides very polar compounds, suited to be analysed by liquid chromatography. In this technique (Figure 2.4), analytes are separated on the basis of their affinity for the stationary phase and for the mobile phase flowing through it. Compounds that interact more with the stationary phase will be more retained, while compounds with characteristics more similar to mobile phase will be eluted first. Numerous types of chromatography exists, differing both for the type of stationary phase and of mobile phase used.

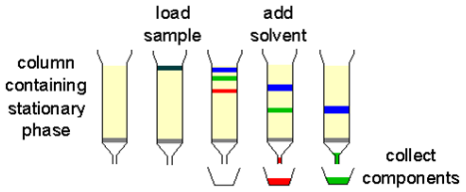

Figure 2.4 Basic principles of chromatographic separation.

The technique of choice for peptide analysis is thus liquid chromatography, especially reversed phase. Other chromatography types are used for particular purposes, such as normal phase, size exclusion, ion exchange, hydrophilic interaction and affinity chromatography, but they are less suited for interfacing with mass spectrometry. The normal phase (NP) chromatography is rarely used in peptide analysis, because usually the stationary phase is silica/alumina and the mobile phase is an apolar organic solvent such as chloroform or hexane, that makes the dissolution of polar compounds like peptides quite impossible. Anyway some methods were developed to separate very polar peptides (that cannot be resolved in reverse phase because they elute with the solvent front), using an amide derivatized stationary phase and a gradient from acetonitrile to acetonitrile/water (55:45) for the elution.[43] Like normal phase chromatography, also hydrophilic interaction chromatography (HILIC) is useful for those short and polar peptides that are not well retained in reverse phase chromatography. The stationary phase is usually comprised of silica or alumina either modified by adsorption with a polar liquid or converted by chemical reaction to form a polar bonded phase. HILIC uses organic mobile phases that are water miscible.[44] In both NPLC and HILIC, compounds elute in order of decreasing hydrophobicity or increasing polarity. Size exclusion chromatography is more suitable for separation of proteins than peptides, even if now new stationary phase are available for relatively low molecular weight polypeptides, so that this technique can be used for the estimation of the molecular weight and for purification purposes.[45] There are two basic types of size exclusion chromatography, but the only employed for polypeptide analysis is gel filtration chromatography (GFC), which uses a hydrophilic packing material and an aqueous mobile phase to separate, fractionate, or measure the molecular weight distribution of molecules soluble in water, such as polysaccharides and proteins.

Polipeptides are separated only on the basis of their dimension and molecular weight: big polypeptides will elute first because they cannot enter the pores of the stationary phase, followed by smaller polypeptides that can enter all the pores then covering a longer distance. Ion exchange chromatography is often used for peptide purification in a complementary way to reverse phase chromatography: the stationary phase can be cationic or anionic and peptides are separated on the basis of their charge. The pH of the buffering mobile phase is chosen on the basis of the pI of ionic groups of the analytes and of the stationary phase.[46] Affinity chromatography takes advantage of biochemical interaction between analyte and stationary phase (often antigen-antibody interaction, but also enzyme-substrate or receptor-ligand). Given the specificity of this interaction, it is used to purify peptides with a high level of purity from complex mixtures.[47] In reverse phase chromatography the stationary phase is constituted by silica particles derivatized with alkyl chains of variable length: the most used have 18, 8 or 4 carbon units (in a decreasing order of lipophilicity). Given the complex composition of food derived peptide mixtures, usually a gradient elution is performed, starting with a polar aqueous eluent and finishing with an organic solvent (for peptide analysis often acetonitrile, sometimes methanol). So the order of elution depends both on the polarity feature of peptides and their size: the firsts to elute will be the smallest and polar peptides, followed by the biggest and apolar ones. To improve chromatographic resolution an acid modifier is usually added to both eluent: in this way peptides are present only as positively charged species and form an ionic pair with the conjugate base of the acid. The ionic pair (formally neutral) is better retained by the lipophilic stationary phase, thus leading to sharp chromatographic peaks. If mass spectrometry is chosen as detection method, the acid modifier added should be volatile and should not form a too strong ionic pair in order to avoid ionic suppression. For these reasons, organic acids are often used, such as formic, acetic and trifluoroacetic acid. In the case of mass spectrometry detection, the use of the organic solvent (acetonitrile or methanol) can be useful for ionization, promoting the desolvation of the drops in electrospray ionization.[48]

2.4.2 DETECTION METHODS

The detection of peptides can be carried out in different ways, but the most frequently used are ultraviolet detection (UV), fluorescence detection (FLD) and mass spectrometry (MS). Ultraviolet (UV) detection of peptides can be carried out at 214 nm exploiting the absorption of the peptidic bond (but having more interfering compounds) or at higher wavelength (254 nm or 280 nm) exploiting the absorption of the side chain of aromatic amino acids (phenylalanine, tyrosine and tryptophan) and being thus more selective. Anyway peptides can be derivatized with chromophoric moieties in order to enhance their UV absorption. This detection method requires a very good peptide separation in the chromatographic column, otherwise the detection and eventual quantification of the peptides can be affected by co-elution of compounds with a similar retention time. Such a separation might be impossible to be reasonably achieved in very complex peptide mixtures.

Fluorescence detection (FLD) is more sensitive and selective than UV and can be performed on native peptides that contain tryptophan and, in a lesser extent, tyrosine. Typically, tryptophan has a wavelength of maximum absorption of 280 nm and an emission peak that is solvatochromic, ranging from ca. 300 to 350 nm depending on the polarity of the local environment. Also in the case of FLD detection the peptidic mixture can be derivatized with fluorescent reagents that enhance sensitivity.[49] Notwithstanding the selectivity of the method, even in this case a good resolution is required for a correct quantification of the peptides.

In the field of peptide identification and quantification in complex mixtures, **mass spectrometry (MS)** detection surely plays a key role with a wide range of hyphenated liquid chromatography-mass spectrometry techniques. Mass spectrometers are used either to measure simply the molecular mass of a polypeptide or to determine additional structural features including the amino acid sequence or the site of attachment and type of posttranslational modifications. In the former case, single-stage mass spectrometers are used. In the latter case, after the initial mass determination, specific ions are selected and subjected to fragmentation through collision. In such experiments, referred to as tandem mass spectrometry (MS/MS), detailed structural features of the peptides can be inferred from the analysis of the masses of the resulting fragments.[50]

The most frequently used ionization modes, in the field of peptide analysis, are electrospray ionization and matrix-assisted laser-desorption ionization. In electrospray ionization (ESI) the ionic species (usually positively charged in peptides analysis) are already present in the mobile phase, due to acidification with a volatile acid such as formic, acetic or trifluoroacetic acid. The electrospray source operates at ambient pressure and the mobile phase exiting from the column is nebulised through a nozzle in a strong electric field. After desolvation of the droplets, the charged analytes enter the mass analyzer. With this technique the formation of multicharged ions is possible. On the contrary, in matrix-assisted laser-desorption ionization (MALDI), mostly single charged ions are formed. The sample is mixed with a suitable matrix and placed on a supporting plate, where a laser is shone onto it. The energy from the laser helps to generate ions, which, propelled by the electrostatic field, are attracted into the mass analyzer. MALDI cannot be directly and easily interfaced with liquid chromatography, as it is usually used as a standalone technique.

In mass spectrometric analysis of peptides, many different analyzers are used: quadrupoles, ion traps, time of flight, cyclotron instruments and hybrid mass analyzers obtained from a combination of the previous one. The quadrupole mass analyzers consist of four parallel metal rods and to each couple of opposite rods a radio frequency voltage is applied. Ions run longitudinally through the space delimited by the four rods following a sinusoidal path. Only ions of a certain mass-to-charge ratio will reach the detector for a given ratio of voltages: other ions have unstable trajectories and will collide with the rods. This permits to scan for a range of m/z-values by continuously varying the applied voltage.[51] Quadrupole mass filters are capable of producing excellent linear dynamic ranges and abundance sensitivities. Triple quadrupole, able to perform MS/MS experiments, are used for quantification in selective reaction monitoring (SRM) mode, post-translational modification detection in precursor ion and neutral loss scanning modes. The ion trap mass analyzers consists a system of three electrodes (a ring electrode and two end-cap electrodes of hyperbolic cross-section) within which ions are trapped. Like quadrupole analysers, ions are subjected to forces applied by a radio frequency field but the forces occur in all three, instead of just two, dimensions. They have high sensitivity and resolution, thus they are used for protein and post-translational modification identification.[52] In time of flight (TOF) analyzers ions are accelerated by an electric field of known strength. The speed of the ions depends on their mass-to-charge ratio and, knowing the time taken for the ions to reach the detector, the mass-to-charge ratio can be determined. TOF can give full isotopic resolution for molecules up to 15 kDa and it is widely used for protein identification from in-gel digestion of gel separated protein band by peptide mass fingerprinting. In cyclotrone instruments (FT-ICR) the ions are trapped in a magnetic field with electric trapping plates, where they are excited to a larger cyclotron

radius by an oscillating electric field orthogonal to the magnetic field. After the excitation field is removed, the ions are rotating at their cyclotron frequency in phase. These ions induce a charge (detected as an image current) on a pair of electrodes as the packets of ions pass close to them and the signal is converted with the Fourier transform into a mass spectrum.[53] Given its high mass accuracy it is a good instrument for top-down proteomics and post-translational modification characterization. The Orbitrap mass analyzer leverages both principles of TOF and FT-ICR: it consists of three electrodes, one spindle like central electrode and the other two forming an outer shell. Ions are injected into the volume between the central and the outer electrodes (in a radial electric field), to which a voltage is applied. The ions remain on a nearly circular spiral inside the trap, and the current generated from the axial oscillation is converted in a mass spectra. The most striking aspect of the Orbitrap's performance is the very high resolution obtained.[54] Hybrid instruments have been designed to combine the capabilities of different mass analyzers and include the Q–q–Q, Q–q–LIT, Q–TOF, TOF–TOF, and LTQ–FTICR.

Depending on the type of mass analyzer used and on the type of analysis, different acquisition modes can be used. In MS experiment three different types of acquisition can be used:

1. Full Scan modality, in which a range of mass to charge ratio is monitored (typically for peptide analysis the range is within 70-2000 m/z);
2. Selected Ion Recognition, were only one (or more) mass to charge ratio is monitored. In this way the sensitivity increases because the scan duration will be lower, thus a higher number of scans for the selected m/z can be made.

In MS/MS experiments, several acquisition mode can be used (Figure 2.5):

1. In Product Ion Scan (also called Daughter Scan), one (or more) mass to charge ratio is selected and fragmented, and the generated ions will be detected; it is a useful technique for compound identification, because every molecule has its characteristic fragmentation spectra.
2. In Precursor Ion Scanning, all the mass to charge ratios that, after fragmentation, generate a fixed m/z ion, are detected. This mode is usually employed when a class of compound with the same moiety (phosphate, sugar, ...) must be detected, because all the compounds of that class will generate the specific fragment.
3. In Neutral Loss Scanning only the compounds that show a specific m/z difference after fragmentation are detected. Since the fixed m/z difference is linked to a specific functional group lost upon fragmentation, this mode is used to detect specific classes of molecules.
4. In Multiple Reactions Monitoring only the fixed m/z ratio that, after fragmentation, generate a fixed m/z ion is detected. It is the more specific detection method, thus it is employed for the detection of a specific analytes with known fragmentation properties in complex samples.

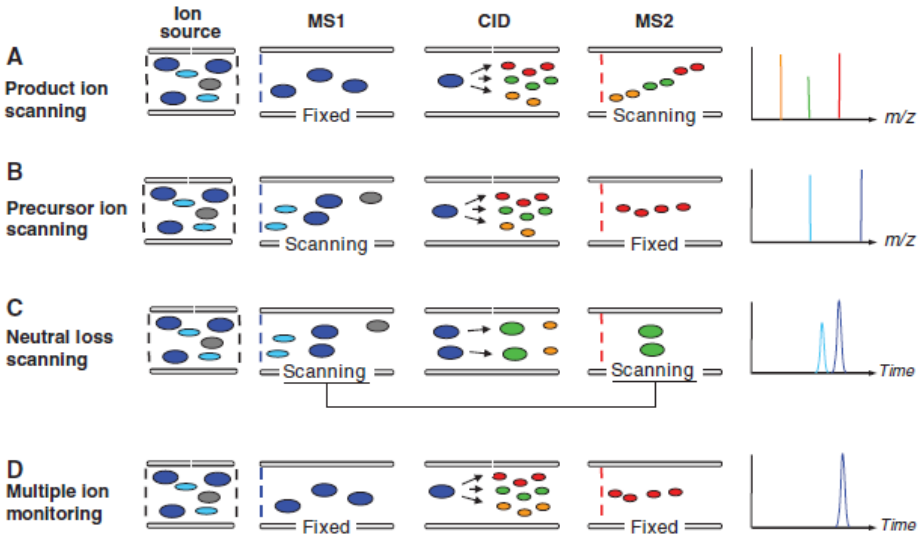

Figure 2.5 Possible acquisition modes in mass spectrometry.[55]

---

[1] Clayton WD, Renvoize SA (1986) In *Genera graminum grasses of the world*. pp. 146–158. HMSO, London.

[2] Matsuoka Y (2011) Evolution of polyploid triticum wheats under cultivation: the role of domestication, natural hybridization and allopolyploid speciation in their diversification. *Plant Cell Physiol*. 52(5): 750–764

[3] Saltini A (2010) In *I semi della civiltà. Frumento, riso e mais nella storia delle società umane*. Nuova Terra Antica, Bologna.

[4] Leff B, Ramankutty N, Foley JA (2004) Geographic distribution of major crops across the world. *Global Biogeochemical Cycles*, 18, GB1009, doi:10.1029/2003GB002108.

[5] Bushuk W, Rasper VF (1994) in *Wheat: Production, Properties and Quality*, pp. 12-13. Springer, ISBN 978-0-7514-0181-3.

[6] International Grain Council, 2002.

[7] Abebe W, Bultos G, Lemessa F (2011) Grain and starch properties of six durum wheat (*Triticum turgidum* L. var. *durum* Desf) varieties grown at Debre Zeit, Ethiopia. *Ethiop J Appl Sci Technol* 2(1): 67 - 74

[8] Fares C, Menga V (2011) Effects of toasting on the carbohydrate profile and antioxidant properties of chickpea (*Cicer arietinum* L.) flour added to durum wheat pasta. *Food Chem* 131: 1140–1148

[9] Pomeranz Y, Chung O, Robinson R J (1966) Lipids in wheat from various classes and varieties. J Am Oil Chem Soc 43(8): 511-514.

[10] National Nutrient Database for Standard Reference, USDA. http://ndb.nal.usda.gov/ndb/foods/show/6306

[11] Tatham AS, Miflin BJ, Shewry PR (1985) The Beta-Turn Conformation in Wheat Gluten Proteins: Relationship to Gluten Elasticity. *Cereal Chem* 62(5): 402-412.

[12] Woychik JH, Boundy JA, Dimler RJ (1961) Starch gel-electrophoresis of wheat gluten proteins with concentrated urea. *Arch Biochem Biophys*. 94: 477-482.

[13] Bietz JA, Huebner FR, Sanderson JE, Wall JS (1977) Wheat gliadin homology revealed throught N-terminal amino acid sequence analysis. *Cereal Chem* 54: 1070 – 1083.

[14] Montserrat F, Miles F (2004) Food allergies: clinical and psychosocial perspectives, in *Plant Food Allergens*, Mills NE, Shewry PR (eds)., Blackwell Science, Oxford.

[15] Sicherer SH, Sampson HA (2006) Food allergy. J Allergy Clin Immunol 117(2): S470-S475.

[16] European Accademy of Allergy and Clinical Immunology, http://www.eaaci.org/resources/what-is-food-allergy.html

[17] http://asthmacenter.textalk.com/page32342.html

[18] Coico R, Sunshine G (2009) *Immunology: A Short Course*. John Wiley & Sons, Manhattan, pp 222-226.

[19] Sampon HA (2003) Update on food allergy, *J Allergy Clin Immunol* 113(5) 805-819.

[20] Cifuentes A (2013) *Foodomics: Advanced Mass Spectrometry in Modern Food Science and Nutrition*. John Wiley & Sons, Manhattan, pp 73-74.

[21] Emmett SE, Angus FJ, Fry JS, Lee PN (1999) Perceived prevalence of peanut allergy in Great Britain and its association with other atopic conditions and with peanut allergy in other household members. *Allergy* 54: 380–385.

[22] Venter C, Pereira B, Grundy J, Clayton CB, Arshad SH, Dean T (2006) Prevalence of sensitization reported and objectively assessed food hypersensitivity amongst six-year-old children: A population-based study. *Pediatr Allergy Immu* 17: 356–363

[23] Pastorello EA, Farioli L, Conti A, Pravettoni V, Bonomi S, Iametti S, Fortunato D, Scibilia J, Bindslev-Jensen C, Ballmer-Weber B, Robino AM, Ortolani C (2007) Wheat IgE mediated food allergy in European patients: a-Amylase inhibitors, lipid transfer proteins and low molecular weight glutenins. *Int Arch Allergy Imm* 144: 10–22.

[24] Morita E, Matsuo H, Mihara S, Morimoto K, Savage AW, Tatham AS (2003) Fast x-5 gliadin, is a major allergen in wheat dependent exercise induced anaphylaxis. *J Dermatol Sci* 33: 99–104.

[25] Tatham AS, Shewry PR (2008) Allergens in wheat and related cereal. *Clin Exp Allergy* 38: 1712–1726.

[26] Guandalini S, Newland C (2011) Differentiating food allergies from food intolerances. *Current Gastroenterology Reports* 13(5): 426-434

[27] Sollid LM, Thorsby E (1993) HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis. Gastroenterology 105(3):910-922

[28] Howell MD, Austin RK, Kelleher D, Nepom GT, Kagnoff MF (1986) An HLA-D region restriction length polymorphism associated with celiac disease. J Exp Med164:333–339

[29] Sollid L, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E (1989) Evidence for a primary association of celiac disease to a particular HLA-DQ-a/b heterodimer. J Exp Med169:345–350

[30] Baton SH, Murray JA (2008) Celiac disease and autoimmunity in the gut and elsewhere. *Gastroenterol Clin North Am* 37(2): 411–vii

[31] Dewar D, Pereira SP, Ciclitira PJ (2004) The pathogenesis of coeliac disease. *Int J Biochem Cell B* 36: 17-24

[32] Van Heel DA, West J (2006) Recent advances in celiac disease. *Gut 55: 1037–1046*

[33] Ciccocioppo R, Di Sabatino A, Corazza GR (2005) The immune recognition of gluten in coeliac disease. *Clin Exp Immunol* 140: 408–416

[34] Stepniak D, Koning F (2006) Celiac Disease—Sandwiched between innate and adaptive immunity. *Hum Immunol* 67: 460–468

[35] Feighery CF (1999) Coeliac disease. *Brit Med J* 319: 236–239.

[36] Green PHR, Jabri B (2003) Coeliac disease. *Lancet* 362: 383–391

[37] Lehrer SB, Ayuso R, Reese G (2006) Current understanding of food allergens. *Ann Ny Acad Sci* 964: 69-85

[38] Breitender H, Mills, ENC (2004) Molecular properties of food allergens. *J Allergy Clin Immunol* 115(1): 14-23

[39] Moreno FJ (2007) Gastrointestinal digestion of food allergens: effect on their allergenicity. *Biomed Pharmacother* 61: 50-60

[40] McGuire M, Beerman KA (2011) *Nutritional sciences: from fundamentals to food.* Cengage Learning, Belmont, pp 191-192

[41] Bravo-Osuna I, Vauthier C, Ponchel G (2007) Core-shell polymer nanoparticle formulations for the oral administration of peptides and proteins. In Mashkevich BO, *Drug delivery research advances*, Nova Science Publisher, New York, pp 43-44.

[42] Thomas K, Herouet-Guicheney C, Ladics G, Bannon G, Cockburn A, Crevel R, Fitzpatrick J, Mills C, Privalle L, Vieths S. Evaluating the effect of food processing on the potential human allergenicity of novel proteins: International workshop report. *Food Chem Toxicol* 45: 1116-1122.

[43] Yoshida T (1997) Peptide separation in normal phase liquid chromatography. *Anal Chem* 69: 3038-3043

[44] Gilar M, Jaworsky A (2011) Retention behavior of peptides in hydrophilic-interaction chromatography. *J Chrom A* 1218: 8890-8896

[45] Irvine GB (2003) High-performance size-exclusion chromatography of peptides. *J Biochem Biophys Methods* 56: 233-242

[46] Andrews PC (1988) Ion-exchange HPLC for peptide purification. Pept Res 1(2): 93-99

[47] Frey DD, Kang X (2005) New concepts in the chromatography of peptides and proteins. *Curr Opin Biotech* 16: 552-560.

[48] Fekete S, Veuthey JL, Guillarme D (2012) New trends in reversed-phase liquid chromatographic separations oftherapeutic peptides and proteins: theory and applications. *J Pharm Biomed Anal* 69: 9-27

[49] Issaq HJ, Chan KC, Blonder J, Ye X, Veenstra TD (2009) Separation, detection and quantitation of peptides by liquid chromatography and capillary electrochromatography. *J Chrom A* 1216: 1825-1837

[50] Domon B, Aebersodl R (2006) Mass Spectrometry and Protein Analysis. Science 312: 212-217

[51] de Hoffmann E, Stroobant V (2003). *Mass Spectrometry: Principles and Applications.* John Wiley & Sons, Toronto, p. 65.

[52] Wong PSH, Cooks RG (1997) Ion Trap Mass Spectrometry. *Current separation* 16:85-92.

[53] Marshall AG, Hendrickson CL, Jackson GS (1998) Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom Rev* 17: 1-35.

[54] Zubarev RA, Makarov A. Orbitrap mass spectrometry. *Anal Chem* 85: 5288-5296.

[55] Domon B, Aebersold R (2006) Mass spectrometry and protein analysis. *Science* 312(5771): 212-217

3 DIFFERENT SIMULATED GASTROINTESTINAL DIGESTION METHODS APPLIED TO WHEAT AND PASTA: IMPLICATIONS FOR THE CELIAC DISEASE

3.1 INTRODUCTION

The protein fraction of wheat includes four main classes: albumins and globulins (functional proteins, soluble in water and in dilute salt solutions, respectively), gliadins and glutenins (storage proteins, soluble in 70% aqueous ethanol and in aqueous 50% 1-propanol with denaturing and reducing agents, respectively) as described by Osborne (1907).[1] These two latter protein classes are called together "gluten" and constitute about 85% of wheat total protein content. They do not have a very high biological value because they are lacking essential amino acids (lysine and tryptophan in particular) and are rich in proline and glutamine.[2] The high content of glutamine is necessary for the plant as a nitrogen source, but the massive presence of proline makes gluten proteins quite resistant to digestion by mammalian proteases.[3] In fact, this amino acid has a cyclic side chain, whose steric hindrance limits the access of proteolytic enzymes to the peptidic bonds. This leads to the formation of high molecular weight peptides in the gastrointestinal tract,[4, 5, 6] some of which are known to be involved in triggering celiac disease.[7, 8] Celiac disease is an autoimmune disorder caused by the ingestion of wheat gluten or related proteins of barley and rye by some genetically predisposed individuals.[9] The immunological reactions involve both the innate and the adaptive immune system: in fact some sequences (PSQQ, QQQP, QQPY and QPYP) stimulate the production of IL-15, other sequences instead are recognized by HLA-DQ2 or DQ8 after deamidation by tissue transglutaminase, and activate a T cell mediated response.[10] The resistance to digestion of some peptides containing sequences previously demonstred to be involved in innate immune response (LGQQQPFPPQQPYPQPQPFPSQQPY) was already assessed[11] as well as that of some immunogenic peptides (QLQPFPQPQLPY, PQPQLPYPQPQLPY by Hausch et al in 2002,[12] FLQPQQPFPQQPQQPYPQQPQQPFPQ by Shan et al in 2005,[13] LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF by Shan et al in 2002[14]). Since the epitopes from the α-gliadin seem to have the highest clinical relevance, many experiments reported in the literature focused on the resistance to simulated gastrointestinal digestion of this fraction: some of them aimed at identifying peptides derived from a recombinant gliadin digested with the main gastrointestinal and pancreatic proteases (pepsin, chymotrypsin, elastase and carboxypeptidase) [11, 13, 14] but also complex matrices, such as flour[15] and whole bread[16], underwent simulated gastrointestinal digestion with similar procedures, and the digestion products were both from the chromatographic profile point of views and/or performing ELISA and T-cells proliferation assays.

In all the above reported digestion models on gluten, the digestion was carried out only considering simple models taking into account protease types, temperature, pH, time and enzyme:substrate ratio. Anyway, in literature are present several papers reporting the development of artificial fluids (saliva, gastric juice, duodenal juice and bile) whose composition strictly simulates the physiological one, in order to better model the environment experimented by a given food during the digestion.[17, 18, 19] Given the complexity of the matrices containing gluten, different experiments performed with different matrices (pure gluten proteins[14] or more complex food matrices[16]) with different proteolytic systems (very simple mixtures of proteolytic enzymes)[11] or more complex mixtures simulating gastric and duodenal juices[20] can lead to conflicting results in terms of gluten

peptides actually produced during the digestion, a molecular characteristic very important in celiac disease-related studies.

In this work a detailed comparison between a simplified digestion model and a complex digestion model with artificial juices were compared. First the simplified model was applied both to a 70% ethanol wheat extracts and to the whole wheat in order to outline the differences induced by the matrix on the peptide mixtures in a simple model. Then the simplified method and the complex method were both applied to wheat and to various gluten-containing samples (from the pasta production chain going from kernels to cooked pasta). The peptides generated were identified in all cases and compared from a qualitative and quantitative point of view, in particular those containing sequences involved in celiac disease, taking into account the different starting matrix (extract or intact food) and the different digestion conditions.

3.2 MATERIALS AND METHODS

3.2.1 Samples

Samples were provided by Council for Research and Experimentation in Agriculture (CRA, Rome). Durum wheat samples (*Triticum turgidum* spp *durum*) were taken at different steps of the pasta production chain (ground kernels, semolina, dough, extruded pasta, dried pasta and cooked pasta) for three different pure varieties (Svevo, Saragolla and Meridiano, provided by PSB, Bologna). Samples were lyophilized and ground with screen 0.5 mm. Traditional pasta (spaghetti shape) were produced by a pilot-scale press and dried in low temperature conditions (maximum peak = 65°C). Pasta cooking was performed with boiling tap water (ratio pasta-water 1:10) without kitchen salt. Cooked pasta were immediately frozen at T = -20°C and then lyophilized.

3.2.2 Reagents and solvents

Deionised water was obtained from a Millipore Alpha Q-Waters purification system (Billerica, MA, USA). Pepsin from porcine gastric mucosa, trypsin from porcine pancreas, α-Chymotrypsin from bovine pancreas, α-amylase from barley malt (type VIIIa), uric acid, mucin from porcine stomach (type III), glucose, glucuronic acid, glucoseamine hydrochloride, albumin bovine, pancreatin from porcine pancreas, lipase from porcin pancreas (type II), bile from bovine and ovine, sodium dihydrogen phosphate, potassium chloride, urea, acetonitrile, ethanol, Fmoc-glutamine(Trt)-OH, piperidine, diisopropylethylamine, dichloromethane, triisopropylsilane, tyrosine, phenylalanine, Tris(hydroxymethyl)-aminomethane, urea, dithiothreitol and diethyl ether were purchased from Sigma-Aldrich (St. Louis, MO, USA). Chloridric acid (37% V/V), sodium hydroxide, sodium hydrogen carbonate, calcium chloride, ammonium chloride and dimethylformamide were purchased from Carlo Erba (Milan, Italy). Fmoc-leucine-OH, Fmoc-proline-OH, HBTU and Fmoc-tyrosine(tBu)-Wang resin were purchased from Novabiochem (Darmstadt, Germany). Formic acid and trifluoroacetic acid were purchased from Acros Organics (Geel, Belgium). Sodium chloride was purchased from analaR Normapur (Milan, Italy). Potassium thiocyanate was purchased from Riedel De Haen (Seelze, Germany). Magnesium chloride was purchased from Merck (Darmstadt, Germany).

3.2.3 Synthesis of the internal standard

The peptide LQLQPFPQPQLPY (isotopically labeled on the phenylalanine residue) was synthesized on solid phase according Fmoc/t-butyl strategy using a Syro I Fully Automated Peptide Synthesizer (Biotage, Uppsala, Sweden). The peptide was cleaved from the Wang-resin using a TFA:TIS:H_2O (95:2.5:2.5) solution and purified using a semipreparative RP-HPLC-UV (λ=280 nm). The purified product was quantified via spectrophotometric method at 280 nm using an external calibration curve made with an equimolar solution of tyrosine and phenylalanine.

3.2.4 Simplified digestion

1 gram of each sample was extracted with 10 ml of a 70% ethanol aqueous solution for 2.5 hours. 1 ml of extract was spiked with 10 µl of internal standard solution (0.477 mM) and dried under nitrogen flux. The sample was reconstituted with 500 µl of HCl 10 mM and added with 40 µl of a pepsin solution (1 mg/ml). After 3 hour of incubation at 37°C the pH was set to 7.2 with 300 µl of phosphate buffer (100 mM) and added with 40 µl of a trypsin and chymotrypsin solution (1 mg/ml). After 4 hour of incubation at 37°C the sample was dried under nitrogen flux. For the digestion of the whole wheat, 10 mg of ground wheat were added of 5 ml of HCl 10 mM pH 2 and 200 µl of pepsin (1 mg/ml). After 3 hour of incubation at 37°C the pH was set to 7.2 with 3 ml of phosphate buffer (100 mM) and added with 200 µl of a trypsin and chymotrypsin solution (1 mg/ml). After 4 hour of incubation at 37°C the sample was centrifuged and supernatant dried under nitrogen flux.

3.2.5 Digestion with artificial digestive juices

The samples were digested as described by Versantvoort et al (2005),[17] adapting the volume of digestive juices to the smallest amount of sample. Briefly, 450 mg of sample were incubated 5 minutes with 600 µl of saliva; after this phase, 2.4 ml of gastric juice were added and the sample was incubated for 2 hours. To set the pH for the intestinal phase, 400 µl $NaHCO_3$ 1 M were added; after, 2.4 ml of duodenal juice and 1.2 ml of bile were added and the sample was incubated for 2 hours. All the digestion steps were carried out at 37°C. At the end of the digestion, 58.3 µl of HCl 37% were added and the sample was centrifuged at 8965 g at 4°C for 45 minutes, in order to precipitate insoluble compounds and undigested proteins. Prior to LC-MS analysis, all samples were filtrated with a cut off of 0.45 µm. For the quantification, 295 µl of digested sample were spiked with 5 µl of standard peptide solution (0.477 mM).

3.2.6 HPLC/ESI-MS/MS analyses

The digested samples were separated by a RP column (JUPITER 5 µm C18 300 Å 250*2 mm) in an HPLC/ESI-MS/MS (HPLC Waters Alliance 2695 with a triple quadrupole mass spectrometer Waters 4 Micro), using a gradient elution. Eluent A was water with 0.1% formic acid and 0.2% acetonitrile, eluent B was acetonitrile with 0.1% formic acid; gradient: 0-12 min 100% A, 12-77 min from 100% A to 50% A, 77-81 min 50% A, 81-82 min from 50% A to 0% A, 82-90 min 0% A, 90-91 min from 0% A to 100% A, 91-110 min 100% A. The samples were first analyzed in Full Scan mode, to identify the characteristic ions and the retention time of the unknown compounds, and then in Daughters Scan modality using a variable collision energy on the basis of the mass and charge of the ion to be fragmented. HPLC/ESI-MS/MS parameters were: flow 0.2 ml/min; analysis time 110 min; column temperature 35°C; injection volume 40 µl; acquisition time 7-90 min; ionization type positive ions; scan range 100-2000 m/z; capillary voltage 3.2 kV; cone voltage 35 V; source temperature 100°C; desolvation

temperature 150°C; cone gas flow 100 l/h; desolvation gas flow 650 l/h. The peptide sequences were assigned on the basis of the mass spectra obtained.

3.2.7 LTQ-OrbiTrap analyses

The samples were analysed by HPLC-LTQ-ORBITRAP using a C18 column and a gradient elution; eluent A was water with 0.1% acetonitrile and 0.1% formic acid and eluent B was acetonitrile with 0.1% formic acid (gradient: 0-4 min from 100% A to 95% A, 4-60 min from 95% A to 50% A, 60-62 min from 50% A to 10% A, 62-72 min 10% A, 72-74 min from 10% A to 95% A, 74-90 min 95% A). The analysis parameters were: flow 5 µl/min; analysis time 90 min; column temperature 30°C; sample temperature: 10°C; injection volume 5 µl; acquisition time 0-75 min; ionization type positive ions; scan range 200-1800 m/z; source voltage 3,5 kV; capillary voltage 35 V; source temperature 275°C.

3.2.8 UPLC/ESI-MS analysis

In order to quantify the peptides in the digested mixtures, the dried digested samples were redissolved in 300 µl of HCOOH 0.1% and separated by a RP column (ACQUITY UPLC BEH 300 C18 1.7 µm 2.1*150 mm) in an UPLC/ESI-MS system (UPLC Acquity Waters with a single quadrupole mass spectrometer Waters AcquityUltraperformance) using a gradient elution. Eluent A was water with 0.1% formic acid and 0.2% acetonitrile, eluent B was acetonitrile with 0.1% formic acid; gradient: 0-7 min 100% A, 7-50 min from 100% A to 50% A, 50-52,6 min 50% A, 52,6-53 min from 50% A to 0% A, 53-58,2 min 0% A, 58,2-59 min from 0% A to 100% A, 59-72 min 100% A. The digested sample extracts were analysed with UPLC/ESI-MS in the Full Scan mode (flow 0.2 ml/min; analysis time 72 min; column temperature 35°C; sample temperature 18°C; injection volume 5 µl; acquisition time 7-58,2 min; ionization type positive ions; scan range 100-2000 m/z; capillary voltage 3.2 kV; cone voltage 30 V; source temperature 100°C; desolvation temperature 200°C; cone gas flow 100 l/h; desolvation gas flow 650 l/h), the characteristic ions of every peptide were extracted, obtaining eXtract Ion Chromatograms (XICs), in which the identified peptides and internal standard LQLQPF(d_5)PQPQLPY were integrated with the MassLynx software. The quantification value was obtained as the ratio peptide area/internal standard area multiplied by the moles of internal standard.

3.3 RESULTS AND DISCUSSION

3.3.1 Peptide composition of mixtures obtained by applying a simplified digestion model to prolamin extracts and to whole wheat

To elucidate the role of the digestion model in the outcome of the simulated digestion, first both the prolamin extract and the whole wheat were digested with the simplified method (pepsin at pH 2 for the gastric phase and trypsin/chymotrypsin at pH 7.2 for the intestinal phase). The digests were then analyzed by UPLC/ESI-MS (Figure 3.1A and C).

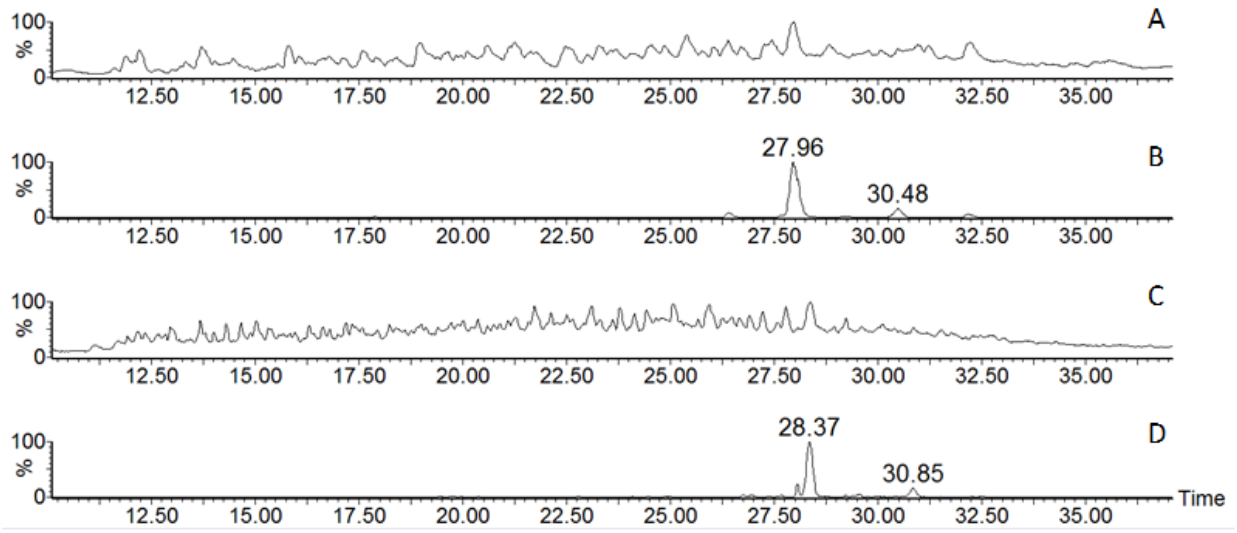

Figure 3.1 Total Ion Chromatograms obtained from UPLC/ESI-MS analysis of digested prolamin extract (A) and whole wheat (C), obtained trough peptic-tryptic/chymotryptic treatment of the sample. In Figure B and D Exctract Ion Chromatogram of the toxic peptide VRVPVPQLQPQNPSQQQPQEQVPL and of the immunogenic peptide QLQPFPQPQLPY obtained from UPLC/ESI-MS analysis of digested prolamin extract (B, tR 27.96 and 30.48 min respectively) and of the digested whole wheat (D, tR 28.37 and 30.85 min respectively).

The compounds giving the most intense chromatographic signals were identify by HPLC/ESI-MS/MS and LTQ-OrbiTrap experiments, determining the amino acidic sequence of the peptides by the mass spectra generated from the collision induced dissociation. No evident differences in the peptide composition in both mixtures were observed: an example of peptides identified present after the in vitro digestion of both ethanolic extract and of whole wheat is shown in Figure 3.1B and D. The peptides identified in both samples digested with the simplified method are reported in Table 3.1, first column.

Table 3.1 Peptides identified through MS/MS experiments in the digested samples obtained with the two different models.

SIMPLIFIED DIGESTION MODEL ON PROLAMIN EXTRACTS SIMPLIFIED DIGESTION MODEL ON WHOLE WHEAT			PHISIOLOGICAL DIGESTION MODEL ON WHOLE WHEAT		
Rt (min)	PEPTIDE	MW (Da)	Rt (min)	PEPTIDE	MW (Da)
11.0	SL	218.1	11.5	VM	248.1
12.3	QL	259.2	11.7	LQPH	493.3
12.8	NS/SN	219.1	12.4	QLA	329.2
13.0	EEIR	545.3	12.7	RSQPQ	614.3
14.0	I/LR	287.2	13.2	VL	230.2
15.6	QAI	330.2	13.6	QAFPQPQ	813.4
15.7	EIAR	487.3	14.5	SQQQQPV	813.4
16.1	VI/L	230.2	15.4	QIPQ	484.3
16.2	IILHQQQK	1006.6	16.1	YPTSPQ	691.3
17.3	VSSL	404.2	16.4	QALPQ	554.9
17.5	SKLPEWMTS	1077.5	16.5	QQQPL	612.3
17.8	VQQVK	600.4	16.6	LQPQNPSQQQPQ	1392.6
18.4	VQQQQF	776.4	16.7	TLPT	430.3
19.1	L/ISAV	388.2	17.5	VPVPQ	538.3
19.4	SIVAG	445.3	17.6	QTLPA	528.3
19.7	APF	333.2	17.7	QTFPHQPQ	981.5
19.8	I/LI/L	244.2	18.0	RPQQPYPQPQPQ	1462.7
19.9	EEIRNL	772.4	18.4	IL	244.2
20.1	LQQKCSPL	915.5	19.8	QQPPFS	702.3
20.6	SQVLQQSTY	1052.5	19.8	QQQPPFS	830.4
20.7	ISMILPR	828.5	20.3	PQQPPFS	799.4
21.8	IMRPL	628.4	20.5	TQQPQQPFPQ	1197.6
21.9	QQIL	500.3	20.7	NVPL	441.3
22.3	QLPSL	556.3	20.8	LQL	372.2
22.3	ILRPL	610.4	21.0	VRVPVPQ	794.5
22.7	WQIPEQSR	1042.5	21.3	SQQPQQPFPQPQ	1408.7
22.8	LQL	372.2	21.4	QQIL	500.3
23.2	TTTRVPF	820.4	21.9	VPF	361.2
23.3	DVVL	444.3	22.0	PQQPPFSQQQQPV	1507.7
23.7	LQPHQIAQL	1046.6	23.1	NLPLPQ	680.4
24.1	IPCM	462.2	23.1	QLPQIPEQS	1038.5
24.3	NLAL	429.3	23.6	WQIPEQS	886.4
24.6	GIF	335.2	23.9	LQPQNPSQQQPQEQVPL	1959
24.7	EVIRSLV	814.5	24.0	GIF	335.2
24.8	LVQGQGIIQPQQPAQ	1603.9	24.2	PSPF	446.2
25.3	IIL	357.3	24.4	QAFPQQPQQPFPQ	1539.8
25.3	EEIRNLAL	956.5	24.7	IIL	357.3
25.5	IIRAPF	715.4	24.9	TQQPQQPFPQQPQQPFPQ	2148
25.61	SHIPGLEKPSQQQPLPL	1868	25.4	QQQPLPL	822.5
25.84	QLVQGQGIIQPQ	1307.7	25.5	QQPPFSQQQPPPFS	1611.8
25.85	SVVHSIIM	884.5	25.6	QVPPL	552.3

26.13	SLVL	430.3
26.46	LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF	3909
26.38	IALPVPSQPVDPR	1388.2
26.63	VRVPVPQLEPQNPSQQQPQEQVPL	2734.4
26.82	QNPSQQQPQEQVPLVQQQ	2103
26.88	LSLL	444.3
27.34	QLPQFEEIR	1158.6
27.77	VPVPQLQPQNPSQQQPQEQVPL	2478.3
28.20	VRVPVPQLQPQNPSQQQPQEQVPLVQQQQF	3491.8
28.77	VRVPVPQLQPQNPSQQQPQEQVPL	2734.4
28.94	QPFLQPQLPY	1229.6
29.03	HIFLPLSQQQQVGQGSL	1879
29.26	FIALPVPSQPVDPR	1534.9
29.84	VRFPVPQLQPQNPSQQQPQEQVPL	2781.5
29.94	QPQQPFPQQPGQIIPQQPQQPFP	2652.3
30.57	HILLPL	704.5
30.74	QLQPFPQPQLPY	1454.8
30.95	IFLPLSQQQQVGQGSL	1741.9
31.48	IVLQQQPPFL	1181.7
32.56	LQLQPFPQPQLPY	1567.8
32.67	QLQPFPQPQLPYPQPQPF	2149.1
33.32	QLQPFPQPQLPYPQPHLPYPQPQPF	2981.5
34.13	LVLPQQQIPFVHPSIL	1828.1
34.05	LQLQPFPQPQLPYPQPQPF	2262.2
35.25	RQPVLPQQPPFSQQQQLVLLQQ	2599.4
26.3	LSLL	444.3
26.7	PQQPYPQPQPFPSQQPYLQLQPFPQPQP	3299.6
26.8	PQTQQPQQPFPQFQQPQQPFPQPQQP	3098.5
26.8	IALPVPSQPVDPR	1387.8
27.3	LGQQQPFPPQQPYPQPQPFPS	2405.2
29.0	PQLPYPQPQPFPPQQSYPQPQPQYPQPQQPI	3637.8
29.3	FPQQPQLPFPQQPQQPFPQPQQPQ	2856.5
29.3	PFPQPQQPQQPFPQSQQPQQPFPQP	2925.4
29.8	QPQLPFPQQPQQPFPQPQQPQQPSPQSQQPQQPFPQ	4186.1
30.2	QLQPFPQPQLP	1291.7
30.7	QQPQQPFPQPQQTFPQQPQLPFPQQPQQPFP	3680.8

Almost all the peptides identified belonged to gliadins, mainly α-gliadin: this was somehow expected for the prolamine extracts, which are mainly composed by gliadins, but this turned out to be the outcome also for the whole wheat. A possible explanation of this latter result is that gluten proteins are much more abundant in wheat as compared to citosolic proteins and, among gluten, gliadins (monomeric) are more soluble in the simple digestion medium used than glutenins (polymeric), so they were more easily proteolyzed by the enzymes used.

Thus, as far as peptide composition was concerned, it was possible to conclude that, when using a simplified digestion model, the influence of the matrix is negligible, but possibly a bias towards the prevalent identification of gliadin-derived peptides is introduced.

3.3.2 Peptide composition of mixtures obtained by applying a simplified digestion model and a physiological digestion model to whole wheat

Then, the simplified digestion model applied to the whole wheat was compared to a physiological one, that requires the use of artificial fluids simulating the complex composition of digestive juices (details in the experimental section). The digested samples were analyzed with UPLC/ESI-MS (Figure 3.2 A and C).

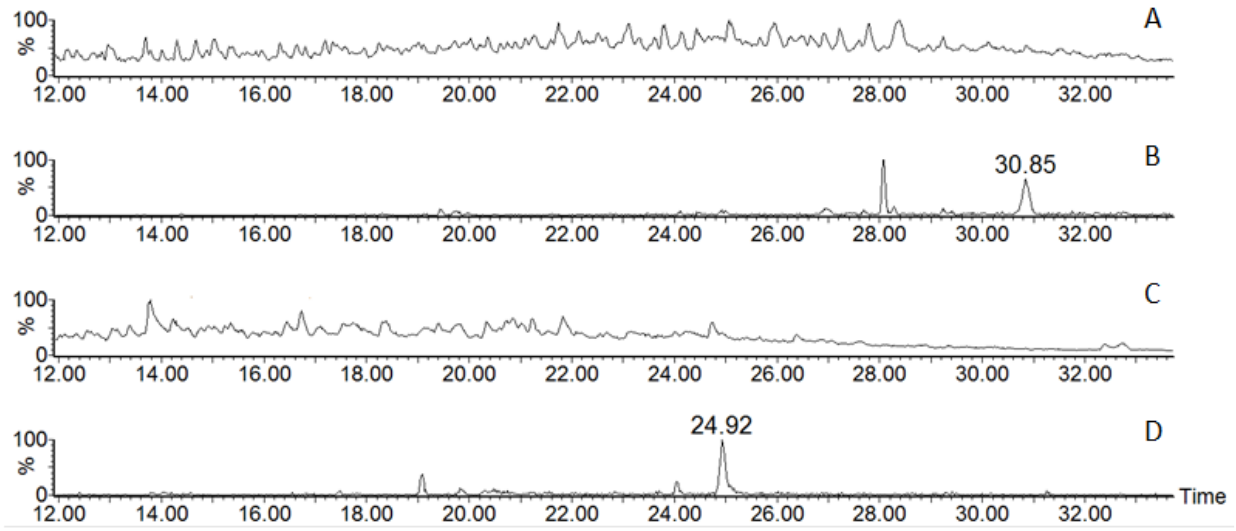

Figure 3.2 Total Ion Chromatograms obtained from UPLC/ESI-MS analysis wheat underwent simplified digestion (A) and the most physiological method (C). In Figure B and D are reported the Extract Ion Chromatogram of the immunogenic peptide QLQPFPQPQLPY (only present in the simplified digestion model, B, t_R 30.85 min) and of the immunogenic peptide TQQPQQPFPQQPQQPFPQ (only present in the more physiological digestion method, D, t_R 24.92 min).

In Figure 3.2 B and D are reported the Extract Ion Chromatogram of two immunogenic peptides (QLQPFPQPQLPY and TQQPQQPFPQQPQQPFPQ) that are present only in function of the digestion method applied.

The molecular weight distribution of the peptides generated by the two types of digestions have similar patterns: the most represented peptides have a length of less than 15 aminoacids, but also high molecular weight peptides (from 15 up to 40 aminoacids) are generated. Instead, a difference was observed in the proteins originating the peptides identified: while in the simplified digestion most of the peptides belonged to α-gliadins, in the complete digestion they were equally distributed among α-gliadins, γ-gliadins and LMW-glutenins. This can be attributed to the different extractability of these protein classes in the digestion medium of the simplified digestion (only 10 mM HCl for the gastric phase and phosphate buffer for the intestinal one) or in presence of surfactant agents such as bile salts in the complex digestion mixtures. The composition of digestive juices in the latter case includes also non proteolytic enzymes such as α-amylases and lipases that, even if not directly implied in protein digestion, contribute to matrix degradation improving protein bioaccessibility and degradation. Moreover, it can be noted that in the simplified digestion the use of three endopeptidases lead mainly to the formation of peptides generated from the recognition of specific cleavage sites (F, Y, W, L, H, M, K and R), whereas the presence of carboxypeptidases in pancreatin (used for the more physiological method of digestion) lead to the formation of shorter peptides with C-term aminoacid not corresponding at any known cleavage site (Table 3.2, second column).

Table 3.2 Pathogenic peptides identified in the digested samples obtained with the two different models, with the protein of origin. Epitopes and toxic sequences are underlined.

Simplified digestion		Phisiological digestion	
Adaptive immune response	**Protein**	**Adaptive immune response**	**Protein**
QLQPFPQPQLPY	α	TQQPQQPFPQ	γ
QLQPFPQPQLPYPQPQPF	α	SQQPQQPFPQPQ	γ
LQLQPFPQPQLPY	α	QAFPQQPQQPFPQ	γ
LQLQPFPQPQLPYPQPQPF	α	TQQPQQPFPQQPQQPFPQ	γ
QLQPFPQPQLPYPQPQLPYPQPQPF	α	PQTQQPQQPFPQFQQPQQPFPQPQQP	γ
QLQPFPQPQLPYPQPHLPYPQPQPF	α	FPQQPQLPFPQQPQQPFPQPQQPQ	γ
LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF	α	PFPQPQQPQQPFPQSQQPQQPFPQP	γ
LPFPQQPQQPFPQPQ	γ	QPQLPFPQQPQQPFPQPQQPQQPSPQSQQPQQPFPQ	γ
		QQPQQPFPQPQQTFPQQPQLPFPQQPQQPFP	γ
Innate immune response	**Protein**	**Innate immune response**	**Protein**
VRVPVPQLQPQNPSQQQPQEQVPLVQQQQF	α	LQPQNPSQQQPQ	α
QNPSQQQPQEQVPLVQQQ	α	RPQQPYPQPQPQ	α
VPVPQLQPQNPSQQQPQEQVPL	α	LQPQNPSQQQPQEQVPL	α
VRVPVPQLEPQNPSQQQPQEQVPL	α	LGQQQPFPPQQPYPQPQPFPS	α
VRVPVPQLQPQNPSQQQPQEQVPL	α	SQQQQPV	γ
VRFPVPQLQPQNPSQQQPQEQVPL	α	QQQPL	LMW
PSSQVQWPQQQPVPQ	γ	QQQPPFS	LMW
NMQVDPSGQVQWPQQQPF	γ	PQQPPFSQQQQPV	LMW
SHIPGLEKPSQQQPLPL	LMW	QQPPFSQQQPPPFS	LMW
		QQQPLPL	LMW

If we consider the peptide sequence known to be involved in celiac disease (Table 3.2), by considering the known peptides containing epitopes recognized by T cells after deamidation by tissue transglutaminase, it is easy to see that they belonged mainly to α-gliadin in the case of simplified digestion, while belonged exclusively to γ-gliadins in the most physiological digestion method. The immunogenic peptides present in the simple digest in fact are further proteolyzed in the complex digestion, generating the shorter peptide QLQPFPQPQLP. So, the absence of α-gliadin derived peptides containing epitopes in the more physiologically digestion system seems to be due not to the effect of the matrix), but to the other digestion parameters, and particularly to the presence of more enzymes: pancreatin in fact not only contains trypsinogen and chymotrypsinogen (used as trypsin and chymotrypsin in the simplified digestion), but also elastase and carboxypeptidase; in particular, this latter enzymes is an exopeptidase cleaving the amino acid at the C-term of the peptide, as happened in the case of the above mentioned immunogenic peptides where C-term tyrosine was removed. The strong presence of γ-gliadin derived peptides in the complex digestion model, instead, seems to be due both to a better extractability of these proteins in the presence of the digestive juices (containing bile salts and with strong presence of organic and inorganic compounds) and to a better bioaccessibility obtained with a more complex enzyme composition.

For what concerns instead peptides able to trigger the innate immunogenic reaction, while in the simplified digestion were present only the sequences PSQQ and QQQP, in the complex digestion of whole wheat also the toxic sequences QQPY and QPYP were detected. These peptides, belonging to the N-term of α-gliadins, were detected in both cases, even if in the more complex system they were shorter due to the stronger proteolysis. Interestingly, in the complete digestion, also the peptide LGQQQPFPPQQPYPQPQPFPS was identified, which is known in literature to induce anti-endomysial antibody production.[21] This peptide is present only in complex digestion system, since its generation requires the action of enzymes such as carboxypeptidase.

3.3.3 Peptides quantification

For both types of digestion the peptides generated were quantified using an internal standard isotopically labeled (LQLQPF(d_5)PQPQLPY, details in the experimental section). The quantification was carried out for six steps of pasta production (involving only durum wheat), in order to verify if some technological treatment have an influence on protein extractability/digestibility. Three different varieties were analyzed in order to exclude variation due to genotype. Results are shown in Figure 3.3.

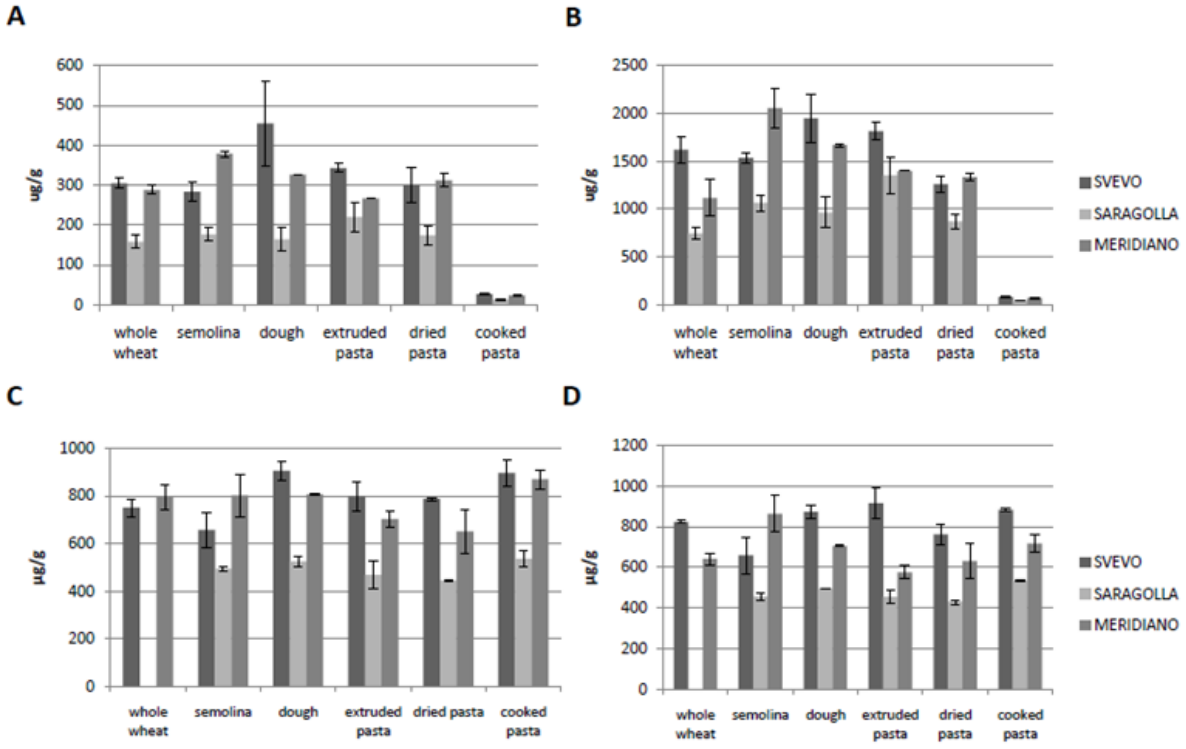

Figure 3.3 Peptides containing sequences involved in adaptive (A) and innate (B) immune response quantified after simplified digestion of the extract; peptides containing sequences involved in adaptive (C) and innate (D) immune response generated after the complex digestion.

As it can be seen, after the more physiological digestion method the amount of peptides containing sequences involved in celiac disease detected was approximately a half as compared to the simplified method, while detected immunogenic peptides were approximately twice, due to the different type of peptides generated. This should be taken into account when biological tests are done in order to assess immunological responses. It can

be noted that the amount of pathogenic peptides remains largely constant along the pasta chain of production, so in case of studies aimed at determining the amount of peptides involved in celiac disease in past, there is no need to obtain the end product but is sufficient to test the basic wheat variety. The difference between the two digestion methods becomes more evident after pasta cooking: in fact heat causes polymerization of gliadins through intermolecular disulphide bridge formation and, to a lesser extent for dehydroalanine formation.[22] This lead to a loss of extractability of gliadin, which in turn lead to a high underestimation of peptides generated after digestion of the ethanolic extract. The complete digestion model clearly outlined that the peptides relevant for celiac disease were still generated in large amount also after pasta cooking. However it is interesting to note that independently from the method adopted, the differences between varieties maintained the same trend in all the steps of processing.

3.4 CONCLUDING REMARKS

Two different digestion models (a simplified digestion procedure and a more physiological one) were tested on wheat-based matrices and the peptides generated were identified through LC-MS techniques, in order to verify the comparability of the two methods. Strong differences were found, both in the aminoacidic sequence of the peptides obtained and in the protein of origin of such peptides, when applying the simplified and the complete method to complex cereal-based matrices. The production of peptides containing sequences implicated in celiac disease was very different in the two cases, both in a qualitative than in a quantitative way. The use of a digestion method involving fluids strictly simulating the composition of digestive juices is thus to be suggested when studying the immunological implications of cereal-based products at the molecular level. Notwithstanding the differences present between the two methods when applied to cereal digestion, the production of toxic and immunogenic peptides had a good correlation between the two methods (Pearson coefficients 0.517 and 0.817, $p<0.01$), meaning that independently from the method adopted, significant differences between varieties remain the same, so, if only a varietal screening has to be performed, the simplified method allows to obtain a hierarchy of the varieties and it is less laborious and expensive.

Finally, the two methods were applied to samples taken at the six basic step of pasta production, showing that the content of pathogenic peptides, determined with both methods, remains constant along the production chain. Anyway, the simplified method is not applicable after pasta cooking, likely due to gliadin polymerization which hampers its extractability and digestion.

3.5 ACKNOWLEDGEMENTS

The AGER Foundation is gratefully acknowledged for its financial support to the project “From seed to pasta.”

---

[1] Osborne TB (1907) *The Proteins of the Wheat Kernel*. Carnegie Institution of Washington

[2] Coultate TP (2002) *Food. The chemistry of its components*. Cambridge: The royal society of chemistry

[3] Thomson R, Hodgman TC, Yang ZR, Doyle AK (2003) Characterizing proteolytic cleavage site activity using bio-basis function neural networks. *Bioinformatics* 19, 1741–1747

[4] Marti T, Molberg O, Li Q, Gray GM, Khosla C, Sollid LM (2005) Prolylendopeptidase mediated destruction of t cell epitopes in whole gluten—chemical and immunological characterization. *J Pharmacol Exp Ther* 312, 19–26

[5] Piper JL, Gray GM, Khosla C (2004) Effect of prolylendopeptidase on digestive-resistant gliadin peptides in vivo. *J Pharmacol Exp Ther* 311, 213–219

[6] Shan L, Marti T, Sollid LM, Gray GM, Khosla C (2004) Comparative biochemical analysis of three bacterial prolylendopeptidases: implications for coeliac sprue. *Biochem J* 383, 311–318

[7] Koning F (2008) Celiac disease: across the threshold of tolerance. In: Fasano, A., Troncone, R., & Branski, D. (Eds), *Frontiers in Celiac Disease* (pp 82-88). Basel: Karger

[8] Stepniak D, Wiesner M, de Ru AH, Moustakas AK, Drijfhout JW, Papadopoulos GK, van Veelen PA, Koning F (2008) Large-scale characterization of natural ligands explains the unique gluten binding properties of HLA-DQ2. J Immunol, 180, 3268–3278

[9] Sollid LM (2002) Celiac disease: dissecting a complex inflammatory disorder. *Nat Rev Immunol* 2, 647–655

[10] Ciccocioppo R, Di Sabatino A, Corazza GR (2005) The immune recognition of gluten in celiac disease. *Clin Exp Immunol* 140, 408–416

[11] Mamone G, Ferranti P, Rossi M, Roepstorff P, Fierro O, Malorni A, Addeo F (2007) Identification of a peptide from α-gliadin resistant to digestive enzymes: implications for celiac disease. *J Chromatogr B* 855, 236-241

[12] Hausch F, Shan L, Santiago NA, Gray GM, Khosla C (2002) Intestinal digestive resistance of immunodominant gliadin peptides. *Am J Physiol-Gastr* L 283, G996-G1003

[13] Shan L, Qiao SW, Arentz-Hansen H, Molberg Ø, Gray GM, Sollid LM, Khosla C (2005) Identification and analysis of multivalent proteolytically resistant peptides from gluten: implications for celiac sprue. *J Proteome Res* 5, 1732-1741

[14] 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, 2275-2279.

[15] Spaenij-Dekking L, Kooy-Winkelaar Y, van Veelen P, Drijfhout JW, Jonker H, Van Soest L, Smulders MJM, Bosch D, Gilissen LJWJ, Koning F (2005) Natural variation in toxicity of wheat: potential for selection of nontoxic varieties for celiac disease patients. *Gastroenterology* 129, 797-806

[16] Gass J, Bethune MT, Siegel M, Spencer A, Khosla C (2007) Combination of enzyme therapy for gastric digestion of dietery gluten in patients with celiac sprue. *Gastroenterology* 133, 472-480

[17] Versantvoort CHM, Oomen AG, Van de Kamp E, Rompelberg CJM, Sips AJAM (2005) Applicability of an in vitro digestion model in assessing the bioaccessibility of mycotoxins from food. *Food Chem Toxicol* 43, 31-40

[18] Oomen AG, Rompelberg CJM, Bruil MA, Dobbe CJC, Pereboom DPKH, Sips AJAM (2003) Development of an in vitro digestion model for estimating the bioaccessibility of soil contaminants. *Arch Environ Con Tox* 44, 281-287

[19] Brandon EFA, Oomen AG, Rompelberg CJM, Versanvoort CHM, van Engelen JGM, Sips AJAM (2006) Consumer product in vitro digestion model: bioaccessibility of contaminants and its application in risk assessment. *Regul Toxicol Pharm* 44, 161-171

[20] Hur SJ, Lim BO, Decker EA, McClements DJ (2011) In vitro human digestion models for food applications. *Food Chem* 125, 1–12

[21] Picarelli A, Di Tola M, Sabbatella L, Anania MC, Di Cello T, Greco R, Silano M, De Vincenzi M (1999) 31-43 amino acid sequence of the α-gliadin induces anti-endomysial antibody production during in vitro challenge. *Scand J Gastroentero* 11, 1099-1102

[22] Lagrain B, Rombouts I, Brijs K, Delcour JA (2011) Kinetics of heat-induced polymerization of gliadin. *J Agr Food Chem* 59, 2034-2039

4 COMPOSITION OF PEPTIDE MIXTURES DERIVED FROM SIMULATED GASTROINTESTINAL DIGESTION OF PROLAMINS FROM DIFFERENT WHEAT VARIETIES

Based on: Prandi B., Bencivenni M., Faccini A., Tedeschi T., Dossena A., Marchelli R., Galaverna G., Sforza S. (2012) Composition of peptide mixtures derived from simulated gastrointestinal digestion of prolamins from different wheat varieties. Journal of Cereal Science 56(2): 223-231.

4.1 INTRODUCTION

Gluten is a complex of reserve proteins found in wheat, barley and rye, composed of gliadins (the alcohol soluble fraction) and glutenins (soluble in dilute acids with denaturing agents).[1] Gluten-derived peptides are formed in the gastrointestinal tract from the incomplete digestion of gluten proteins. As a matter of fact, gluten is not completely digested from gastric, pancreatic and intestinal proteases due to its high content of proline,[2] since the cyclic structure of this amino acid interferes with the enzyme accessibility to the peptidic bond. Thus, from the gastrointestinal digestion of gluten, proteolitically resistant peptides are formed, some of them implicated in the pathogenesis of celiac disease.[3]

Celiac disease is an autoimmune disease that develops in some genetically predisposed subjects after the consumption of foods containing gluten,[4] in fact celiac disease is associated with the expression of HLA-DQ2 or DQ8 molecules.[5] Pathogenesis is due to both innate and adaptive immune response.[6] The innate immune response is activated by some sequence of gluten-derived peptides (PSQQ, QQQP, QPYP and QQPY),[7, 8] that induce the production of interleukin-15.[9] The adaptive immune response, instead, begins with the recognition of some T-cells epitopes in gluten-derived peptides,[10] which are bound by antigen presenting cells that express the human leukocyte antigens DQ2 or DQ8.[11] Moreover, in the lamina propria, the enzyme tissue transglutaminase catalyzes the deamination of specific glutamine residues in the immunogenic peptides,[12] causing a stronger immunological response.[13] These complexes are recognized by T-cells that induce the intestinal epithelium damages.[14] Both toxic and immunogenic peptides can reach the lamina propria taking advantage of mucosal defects allowing their passage between or through the epithelial cells.[15]

Many studies on celiac disease hinted at the prolamin fraction[16] as the major external trigger in celiac disease, which is mainly composed of gliadins. Gliadins are classified according to their electrophoretic mobility in α/ β-, γ- and ω-type.[17] A recent study however defined a hierarchy of the immunogenic peptides, showing a strong contribution, besides α-gliadins, of ω-gliadins.[18] Many peptides have been reported in the literature as able to elicit a cell-mediated response in celiac patients[19, 20] or to be toxic for the intestinal mucosa of celiac subjects.[21] These peptides have been identified in gastric/tryptic digest of a single purified gliadin or of a pure recombinant gliadin,[22] and adverse effects of the gliadin peptides have been demonstrated on T-cells from intestinal biopsies and/or peripheral mononuclear cells from celiac disease patients using the synthetic peptides. However, so far, an extensive study of the peptides formed during gastrointestinal digestion of the prolamin fraction extracted from real wheat samples has never been performed.

This issue is not trivial for several reasons: first, celiac disease is strictly dependent on the amount of gluten ingested. The threshold level of gluten exposure that triggers wheat allergy is unknown, but for celiac disease 50

mg daily is associated with intestinal inflammation, so a food is considered safe for celiac subjects when its gluten content is below 20 ppm (CE regulation 41/2009),[23] therefore the most dangerous peptides are presumably those which are present in a gluten digests in the highest amounts. Second, wheat varieties highly differ for amount and type of gliadins, and therefore it is very important to define the amount and the type of pathogenic peptides released by gastrointestinal digestion of different wheat varieties. Accordingly, recently a study suggested that celiac disease autoimmunity may develop at any age, even in the elderly; the sharp increase of celiac disease prevalence is related to an increasing number of subjects that, in their adulthood, lose the immunological tolerance to gluten.[24] The amount and the quality of ingested gluten, type and duration of wheat dough fermentation, the spectrum of intestinal microbiota and its changes over time, enteric infections, and stressors in general are all possible switches of the tolerance/immune response balance.[25]

The aim of the present study was to characterize (with LC/ESI-MS techniques) the most abundant peptides derived from the simulated gastrointestinal digestion of prolamin extracts from wheat samples, belonging to different *T. durum* and *aestivum* varieties. The peptides, including in particular those containing toxic sequences or T-cells epitopes already known in literature, have been quantified with an isotopically labelled immunogenic peptide, allowing the comparison of their content in the different wheat varieties. Toxic and immunogenic peptides mostly formed during the simulated gastrointestinal digestion indicate the main peptides which could come in contact with our intestinal cells and, at the same time, allow to identify *T. durum* varieties potentially having a lower pathogenic impact in celiac disease.

4.2 MATERIALS AND METHODS

4.2.1 Reagents and solvents

Deionised water was obtained from a Millipore Alpha Q-Waters (Billerica, MA, USA) purification system. Pepsin from porcine gastric mucosa, trypsin from porcine pancreas, α-Chymotrypsin from bovine pancreas, gliadin from wheat, sodium dihydrogen phosphate, acetonitrile, ethanol, Fmoc-glutamine(Trt)-OH, piperidine, diisopropylethylamine, dichloromethane, triisopropylsilane and diethyl ether were purchased from Sigma-Aldrich (Stockholm, Sweden). Chloridric acid (37% V/V) and dimethylformamide were purchased from Carlo Erba (Milan, Italy). Fmoc-leucine-OH, Fmoc-proline-OH, HBTU and Fmoc-tyrosine(tBu)-Wang resin were purchased from Novabiochem (Darmstadt, Germany). Formic acid was purchased from Acros Organics (Geel, Belgium).

4.2.2 Samples

Wheat samples (24, Table 4.1) were obtained from Società Produttori Sementi SpA (Bologna, Italy).

Table 4.1 List of the analysed samples and of their total protein and prolamine content.

Code	Species	Line	Area of production	Total protein content (%, Kjeldhal)	Gliadin content (%, fluorimetric method)
1	*Triticum turgidum* spp *durum*	D240	Argelato (BO)	10.34±1.30	3.13±0.04
6	*Triticum turgidum* spp *durum*	D240	Argelato (BO)	9.66±0.11	3.38±0.02
9	*Triticum turgidum* spp *durum*	D240	Poggio Renatico (FE)	11.73±1.02	3.79±0.01
12	*Triticum turgidum* spp *durum*	D240	Poggio Renatico (FE)	11.54±1.19	3.90±0.17
13	*Triticum turgidum* spp *durum*	D240	Lucera (FG)	13.57±0.45	3.54±0.19
18	*Triticum turgidum* spp *durum*	D240	Lucera (FG)	14.85±0.23	4.07±0.16
3	*Triticum turgidum* spp *durum*	Levante	Argelato (BO)	11.61±0.28	3.57±0.08
5	*Triticum turgidum* spp *durum*	Levante	Argelato (BO)	10.22±0.56	3.54±0.19
8	*Triticum turgidum* spp *durum*	Levante	Poggio Renatico (FE)	13.45±0.51	3.91±0.03
11	*Triticum turgidum* spp *durum*	Levante	Poggio Renatico (FE)	12.46±0.45	3.93±0.06
15	*Triticum turgidum* spp *durum*	Levante	Lucera (FG)	15.69±0.17	4.43±0.10
16	*Triticum turgidum* spp *durum*	Levante	Lucera (FG)	17.65±0.45	4.60±0.03
2	*Triticum turgidum* spp *durum*	Svevo	Argelato (BO)	13.65±1.69	4.00±0.16
4	*Triticum turgidum* spp *durum*	Svevo	Argelato (BO)	12.38±0.23	3.93±0.32
7	*Triticum turgidum* spp *durum*	Svevo	Poggio Renatico (FE)	14.21±0.45	4.05±0.09
10	*Triticum turgidum* spp *durum*	Svevo	Poggio Renatico (FE)	14.57±0.40	4.63±0.08
14	*Triticum turgidum* spp *durum*	Svevo	Lucera (FG)	15.21±0.40	4.62±0.44
17	*Triticum turgidum* spp *durum*	Svevo	Lucera (FG)	18.88±0.51	4.64±0.03
19	*Triticum aestivum*	C172	Argelato (BO)	14.37±0.11	4.06±0.24
20	*Triticum aestivum*	C173	Argelato (BO)	13.65±0.34	4.20±0.05
21	*Triticum aestivum*	C174	Argelato (BO)	13.05±0.28	3.66±0.08
22	*Triticum aestivum*	C181	Argelato (BO)	14.25±0.73	3.98±0.22
23	*Triticum aestivum*	Centauro	Argelato (BO)	13.77±0.51	3.80±0.06
24	*Triticum turgidum* spp *turanicum*	Kamut	Argelato (BO)	16.09±1.75	4.43±0.10

4.2.3 Prolamin extraction

Food kernels were ground, weighed and extracted with a solution of 70% ethanol (ratio matrix:solvent 1:20 w/V). The mixture was left stirring for 2 hours and 30 minutes and centrifuged at 2486*g* for 15 min at 10°C.

4.2.4 Protein quantification

In order to assess the protein content of the prolamin extract, a Q-bit fluorometer (Invitrogen, Grand Island, NY, USA) was used. The working solution was prepared by mixing 1/200 of the Quant-it reagent with 199/200 of Quant-it protein buffer; then, 2 µl of each prolamin extract was mixed with 198 µl of working solution. The sample was left for 15 min in the dark before reading the protein content with the fluorometer. The total protein content was determined according to the Kjeldhal method.

4.2.5 SDS_PAGE of the extracts

50 µl of prolamin extracts were dried under nitrogen flux and reconstituted with 25 µl of sample reducing buffer (pH 6,8; 1% SDS; 12,5% glycerol; 0,005% bromophenol blue; 2,5% 2-mercaptoethanol); after denaturation at 95°C for 5 min, the samples were loaded into a Criterion bis-tris precast gels 12%. The running buffer was 25

mM Tris, 192 mM glycine and 0,1% SDS and the potential applied was 150V. The gel was finally stained with Coomassie Brilliant Blue.

4.2.6 Synthesis of the internal standard

The peptide LQLQPFPQPQLPY (isotopically labeled on the phenylalanine residue) was synthesized on solid phase according Fmoc/t-butyl strategy using a Syro I Fully Automated Peptide Synthesizer (Biotage, Uppsala, Sweden). The peptide was cleaved from the resin using a TFA:TIS:H2O (95:2.5:2.5) solution and purified using a semipreparative RP-HPLC-UV (λ=280 nm). The purified product was quantified via spectrophotometric method at 280 nm using an external calibration curve made with an equimolar solution of tyrosine and phenylalanine.

4.2.7 Digestion of prolamin extracts

For the enzymatic digestion of the prolamin extracts, essentially the method of Shan et al.[26] (2002) was followed with slight modifications. 1 ml of prolamin extract was dried under nitrogen flux; 500 µl of HCl 0.01 N were then added to the sample and mixed; the final pH of the solution was set at pH=2. 20 µl of a pepsin solution (1 mg/ml; enzyme:substrate ratio 1:100) were then added to the sample, the resulting solution was mixed and put in a thermostatic bath, with a magnetic stirrer, at 37°C for 3 hours (gastric phase).[27] At the end of the gastric phase, the sample was neutralized until pH=7 with 300 µl of phosphate buffer (NaH2PO4 100 mM). Then, the sample was mixed with 20 µl of a chymotrypsin solution (1 mg/ml; enzyme:substrate ratio 1:100) and 20 µl of a trypsin solution (1 mg/ml; enzyme:substrate ratio 1:100 (Shan et al., 2002)). The resulting solution was put in a thermostatic bath with a magnetic stirrer, at 37°C for 4 hours (intestinal phase). At the end of the digestion procedure, the sample was dried under nitrogen flux. In the samples required for the quantification of the generated peptides, the prolamine extract was added of 5 µl of a 0.477 mM solution of LQLQPF(d_5)PQPQLPY, used as internal standard.

4.2.8 HPLC/ESI-MS/MS analyses

The dried digested samples were redissolved in 300 µl of HCOOH 0,1% and separated by a RP column (JUPITER 5 µm C18 300 Å 250*2 mm) in an HPLC/ESI-MS/MS (HPLC Waters Alliance 2695 with a triple quadrupole mass spectrometer Waters 4 Micro), using a gradient elution. Eluent A was water with 0.1% formic acid and 0.2% acetonitrile, eluent B was acetonitrile with 0.1% formic acid; gradient: 0-12 min 100% A, 12-77 min from 100% A to 50% A, 77-81 min 50% A, 81-82 min from 50% A to 0% A, 82-90 min 0% A, 90-91 min from 0% A to 100% A, 91-110 min 100% A. The samples were first analyzed in Full Scan mode, to identify the characteristic ions and the retention time of the unknown compounds, and then in Daughters Scan modality using a variable collision energy on the basis of the mass and charge of the ion to be fragmented. HPLC/ESI-MS/MS parameters were: flow 0.2 ml/min; analysis time 110 min; column temperature 35°C; sample temperature 20°C; injection volume 40 µl; acquisition time 7-90 min; ionization type positive ions; scan range 100-2000 m/z; capillary voltage 3.2 kV; cone voltage 35 V; source temperature 100°C; desolvation temperature 150°C; cone gas flow 100 l/h; desolvation gas flow 650 l/h. The peptide sequences were assigned on the basis of the mass spectra obtained.

4.2.9 LTQ-OrbiTrap analyses

Dried digested mixtures were redissolved in 10 µl of trifluoroacetic acid 0.1% and ZipTip pipette tips (Millipore) were used for desalting the samples. Tips were first wetted with acetonitrile and then equilibrated with trifluoroacetic acid 0.1%. For peptides binding, the solution was adsorbed on the tips several times. Then, the tips were washed with an aqueous solution of 5% methanol and 0.1% trifuoroacetic acid in order to eliminate the salts and finally the peptides were eluted with an aqueous solution of methanol 50% and 0.1% formic acid. The desalted solution containing the peptides was dried (at 0°C) with an air flux, redissolved in 60 µl deionised water and quantified with the Q-Bit system. The samples were analysed by HPLC-LTQ-ORBITRAP using a C18 column and a gradient elution; eluent A was water with 0.1% acetonitrile and 0.1% formic acid and eluent B was acetonitrile with 0.1% formic acid (gradient: 0-4 min from 100% A to 95% A, 4-60 min from 95% A to 50% A, 60-62 min from 50% A to 10% A, 62-72 min 10% A, 72-74 min from 10% A to 95% A, 74-90 min 95% A). The analysis parameters were: flow 5 µl/min; analysis time 90 min; column temperature 30°C; sample temperature: 10°C; injection volume 5 µl; acquisition time 0-75 min; ionization type positive ions; scan range 200-1800 m/z; source voltage 3.5 kV; capillary voltage 35 V; capillary temperature 275°C.

4.2.10 UPLC/ESI-MS analysis

In order to quantify the peptides in the digested mixtures, the dried digested samples were redissolved in 300 µl of HCOOH 0.1% and separated by a RP column (ACQUITY UPLC BEH 300 C18 1.7 µm 2.1*150 mm) in an UPLC/ESI-MS system (UPLC Acquity Waters with a single quadrupole mass spectrometer Waters Acquity Ultraperformance) using a gradient elution. Eluent A was water with 0.1% formic acid and 0.2% acetonitrile, eluent B was acetonitrile with 0.1% formic acid; gradient: 0-7 min 100% A, 7-50 min from 100% A to 50% A, 50-52,6 min 50% A, 52,6-53 min from 50% A to 0% A, 53-58,2 min 0% A, 58,2-59 min from 0% A to 100% A, 59-72 min 100% A. The digested sample extracts were analysed with UPLC/ESI-MS in the Full Scan mode (flow 0.2 ml/min; analysis time 72 min; column temperature 35°C; sample temperature 6°C; injection volume 5 µl; acquisition time 7-58,2 min; ionization type positive ions; scan range 100-2000 m/z; capillary voltage 3.2 kV; cone voltage 30 V; source temperature 100°C; desolvation temperature 200°C; cone gas flow 100 l/h; desolvation gas flow 650 l/h), the characteristic ions of every peptide were extracted, obtaining eXtract Ion Chromatograms (XICs), in which the identified peptides and internal standard LQLQPF(d_5)PQPQLPY were integrated with the MassLynx software. The quantification value was obtained as the ratio peptide area/internal standard area multiplied by the moles of internal standard.

4.2.11 Statistical analysis

The statistical treatment of the data (analysis of variance and Pearson correlation) was carried out using SPSS statistics 17.0 software.

4.3 RESULTS AND DISCUSSION

4.3.1 Extraction and digestion of the prolamin fraction

The extraction of the prolamin fraction from the various samples was carried out with 70% ethanol for 2 hours and 30 min. The different extracts were quantified with fluorimetric method as far as their protein content was concerned. This procedure allowed a protein concentration of 1.5-2.3 mg/ml to be obtained. An SDS-PAGE

analysis of the extracts, together with a standard of gliadin, confirmed that the protein extracts were mostly composed of gliadins (Figure 4.1).

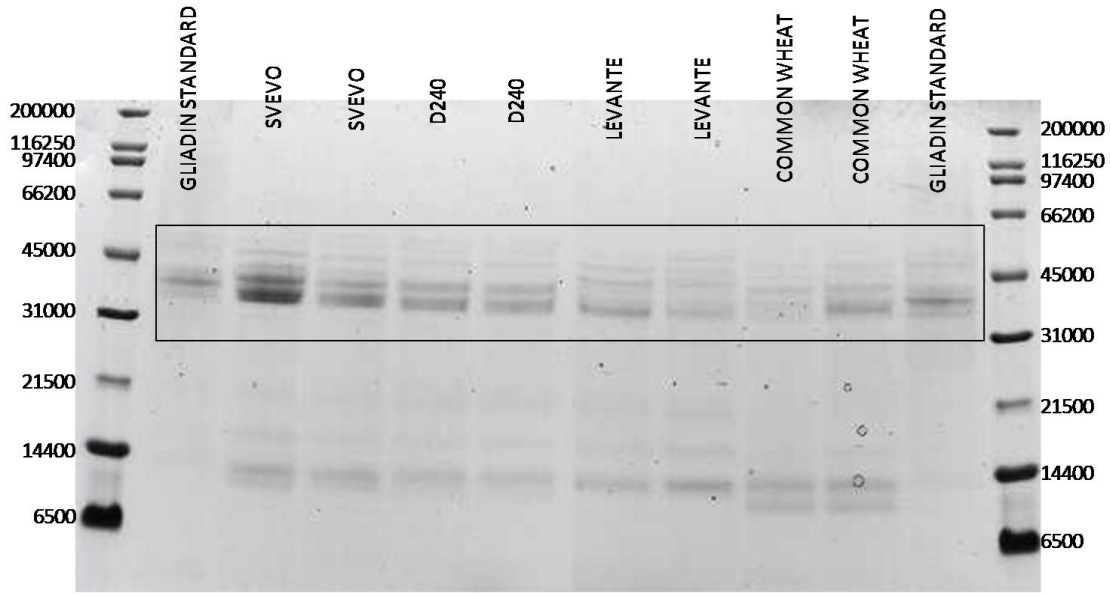

Figure 4.1 SDS-PAGE of various prolamin extracts and of the gliadin standard. Marker was loaded into the outer lanes and molecular weight is expressed in Da; gliadin bands are circled (MW = 31.0 – 66.2 KDa).

For the simulated gastrointestinal digestion, the method of Shan et al. (2002) was essentially applied, to mimic the peptides generated in the healthy proximal-mid small intestine. It is to be underlined that, since the immune response occurs after absorption (and not when the apical surface of the epithelium is exposed to gluten), the peptides generated might be further processed after absorption and before being presented to T-cells or have other biological effects. Some parameters were varied in order to optimize the digestion conditions. For simulating the gastric phase, pepsin at pH=2 was used, whereas, for simulating the intestinal phase, a mixture of trypsin and chymotrypsin at pH=7 was chosen, starting with an enzyme:substrate ratio of 1:100. The pH in the gastric phase was set by using 10 mM HCl, then the pH was raised to 7 with phosphate buffer in the intestinal phase. The possible contribution of brush border proteases was not considered in this simulation. In all cases, the outcome of the digestion was assessed by analyzing the extracts by LC/ESI-MS.

4.3.2 Peptides identification

The digested extracts were analyzed with UPLC/ESI-MS, in order to detect the compounds giving the highest chromatographic signals. The sequence of these peptides was then deduced by performing MS/MS experiments, either with triple quadrupole or OrbiTrap analyzer. Peptides with lower molecular weight (generally present as singly charged ions) were characterized in low resolution mode by HPLC/ESI-MS/MS spectra obtained with a triple quadrupole instrument, whereas high molecular weight peptides (generally present as multiply charged ions) were characterized in high resolution mode by HPLC/ESI-MS/MS spectra obtained with an LTQ-ORBITRAP instrument. The peptides identified are reported in Table 4.2.

Table 4.2 Peptides identified in the prolamin digested extracts of T. durum and T. aestivum .

avg Rt (min)	Peptide	MW (Da)	Ions (MH^+)	*T. durum*	*T. aestivum*
11.0	SL	218.1	219.1	√	√
12.3	QL	259.2	260.2	√	√
12.8	NS/SN	219.1	220.1	√	X
13.0	EEIR	545.3	546.3	X	√
14.0	I/LR	287.2	288.2	√	X
15.6	QAI	330.2	331.2	X	√
15.7	EIAR	487.3	488.3	√	√
16.1	VI/L	230.2	231.2	√	X
16.2	IILHQQQK	1006.6	504.3 $[M+2H^+]^{2+}$	√	√
			336.5 $[M+3H^+]^{3+}$		
17.3	VSSL	404.2	405.2	√	√
17.5	SKLPEWMTS	1077.5	539.8 $[M+2H^+]^{2+}$	√	√
			360.2 $[M+3H^+]^{3+}$		
17.8	VQQVK	600.4	601.4	√	√
18.4	VQQQQF	776.4	777.4	√	√
			389.2 $[M+2H^+]^{2+}$		
19.1	L/ISAV	388.2	389.2	√	√
19.4	SIVAG	445.3	446.3	√	X
19.7	APF	333.2	334.2	√	√
19.8	I/LI/L	244.2	245.2	X	√
19.9	EEIRNL	772.4	773.4	√	√
			387.4 $[M+2H^+]^{2+}$		
20.1	LQQKCSPL	915.5	458.8 $[M+2H^+]^{2+}$	√	√
20.6	SQVLQQSTY	1052.5	1053.5	√	√
			527.3 $[M+2H^+]^{2+}$		
20.7	ISMILPR	828.5	829.5	X	√
			415.3 $[M+2H^+]^{2+}$		
21.8	IMRPL	628.4	315.2	√	√
21.9	QQIL	500.3	501.3	√	√
22.3	QLPSL	556.3	557.3	√	√
22.3	ILRPL	610.4	306.3 $[M+2H^+]^{2+}$	√	√
22.7	WQIPEQSR	1042.5	1043.5	X	√
			522.3 $[M+2H^+]^{2+}$		
22.8	LQL	372.2	373.2	√	√
23.2	TTTRVPF	820.4	822.0	√	X
			411.5 $[M+2H^+]^{2+}$		
23.3	DVVL	444.3	445.3	√	√
23.7	LQPHQIAQL	1046.6	524.3 $[M+2H^+]^{2+}$	√	√
24.1	IPCM	462.2	463.2	X	√
24.3	NLAL	429.3	430.3	√	√
24.6	GIF	335.2	336.2	√	√
24.7	EVIRSLV	814.5	408.3 $[M+2H^+]^{2+}$	√	√
24.8	LVQGQGIIQPQQPAQ	1603.9	803.4 $[M+2H^+]^{2+}$	√	√
			536.0 $[M+3H^+]^{3+}$		
25.3	IIL	357.3	358.3	√	√
25.3	EEIRNLAL	956.5	957.5	√	√
			479.3 $[M+2H^+]^{2+}$		

25.5	IIRAPF	715.4	358.7 $[M+2H^+]^{2+}$	X	√
25.6	SHIPGLEKPSQQQPLPL	1868.0	935.6 $[M+2H^+]^{2+}$	√	X
			624,1 $[M+3H^+]^{3+}$		
25.8	QLVQGQGIIQPQ	1307.7	654.9 $[M+2H^+]^{2+}$	√	√
25.9	SVVHSIIM	884.5	443.6 $[M+2H^+]^{2+}$	√	√
26.1	SLVL	430.3	431.3	√	√
26.5	LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF	3909.0	1304.0 $[M+3H^+]^{3+}$	X	√
			978.2 $[M+4H^+]^{4+}$		
26.4	LVLSImSMILPR	1388.2	695.1 $[M+2H^+]^{2+}$	√	√
			463.8 $[M+3H^+]^{3+}$		
26.6	VRVPVPQLEPQNPSQQQPQEQVPL	2734.4	1368.2 $[M+2H^+]^{2+}$	√	√
			912.5 $[M+3H^+]^{3+}$		
			684.6 $[M+4H^+]^{4+}$		
26.8	QNPSQQQPQEQVPLVQQQ	2103.0	1052.5 $[M+2H^+]^{2+}$	√	X
			702.0 $[M+3H^+]^{3+}$		
26.9	LSLL	444.3	445.3	√	√
27.3	QLPQFEEIR	1158.6	580.3 $[M+2H^+]^{2+}$	X	√
27.8	VPVPQLQPQNPSQQQPQEQVPL	2478.3	827.6 $[M+3H^+]^{3+}$	√	√
28.2	VRVPVPQLQPQNPSQQQPQEQVPLVQQQQF	3491.8	1165.7 $[M+3H^+]^{3+}$	√	√
			874.5 $[M+4H^+]^{4+}$		
28.8	VRVPVPQLQPQNPSQQQPQEQVPL	2734.4	1368.2 $[M+2H^+]^{2+}$	√	√
			912.5 $[M+3H^+]^{3+}$		
			684.6 $[M+4H^+]^{4+}$		
28.9	QPFLQPQLPY	1229.6	615.8 $[M+2H^+]^{2+}$	√	√
29.0	HIFLPLSQQQQVGQGSL	1879	941.1 $[M+2H^+]^{2+}$	X	√
			627.7 $[M+3H^+]^{3+}$		
29.3	FIALPVPSQPVDPR	1534.9	768.9 $[M+2H^+]^{2+}$	√	√
			512.9 $[M+3H^+]^{3+}$		
29.8	VRFPVPQLQPQNPSQQQPQEQVPL	2781.5	1392.6 $[M+2H^+]^{2+}$	√	√
			928.7 $[M+3H^+]^{3+}$		
			696.8 $[M+4H^+]^{4+}$		
29.9	QPQQPFPQQPGQIIPQQPQQPFP	2652.3	1328 $[M+2H^+]^{2+}$	X	√
			885.7 $[M+3H^+]^{3+}$		
30.6	HILLPL	704.5	705.5	X	√
			353.2 $[M+2H^+]^{2+}$		
30.7	QLQPFPQPQLPY	1454.8	728.8 $[M+2H^+]^{2+}$	√	√
31.0	IFLPLSQQQQVGQGSL	1741.9	872.5 $[M+2H^+]^{2+}$	X	√
			582.0 $[M+3H^+]^{3+}$		
31.5	IVLQQQPPFL	1181.7	1182.7	X	√
			591.8 $[M+2H^+]^{2+}$		
32.6	LQLQPFPQPQLPY	1567.8	785.4 $[M+2H^+]^{2+}$	√	√
32.7	QLQPFPQPQLPYPQPQPF	2149.1	1075.6 $[M+2H^+]^{2+}$	X	√
			717.4 $[M+3H^+]^{3+}$		
33.3	QLQPFPQPQLPYPQPHLPYPQPQPF	2981.5	995.5 $[M+3H^+]^{3+}$	X	√
			747.2 $[M+4H^+]^{4+}$		
34.1	LVLPQQQIPFVHPSIL	1828.1	915.6 $[M+2H^+]^{2+}$	√	√
			610.7 $[M+3H^+]^{3+}$		
34.1	LQLQPFPQPQLPYPQPQPF	2262.2	1132.1 $[M+2H^+]^{2+}$	√	√
			755.1 $[M+3H^+]^{3+}$		
35.2	RQPVLPQQPPFSQQQQLVLLQQ	2599.4	868.0 $[M+3H^+]^{3+}$	√	X

36.2	QLQPFPQPQLPYPQPQLPYPQPQPF	2974.4	744.7 $[M+4H^+]^{4+}$ 992.7 $[M+3H^+]^{3+}$	X	√

The great majority of them belong to α-gliadins, less to γ-gliadins and only a few to ω-gliadins: this could be due to the sequence variability of these latter proteins, which might lead to the formation of many different peptides, each present in a small amount, with different sequences and length. Moreover α-gliadins are more abundant than γ-gliadins and, even more, of ω-gliadins.[28]

4.3.3 Marker peptides for gliadin amount

The real amount of gliadins in a given food is a parameter often difficult to determine on the intact proteins. As seen before, α-gliadin sequence is highly variable, but database search indicated that the N-terminal region (1-30) is conserved in all gliadin isoforms. Thus, peptides 1-24 and 3-24 (quite abundant in our digests) were evaluated as potential markers of the initial amount of α-gliadin, assuming that their presence was proportional to the initial number of α-gliadin molecules. Peptides derived from the enzymatic cleavage of the N-terminal region of gliadins are reported in Table 4.3.

Table 4.3 Markers peptides of the amount of gliadins in the food product and most abundant immunogenic and toxic peptides identified in the digested prolamin extracts (known immunogenic and toxic sequences are underlined), together with a indication of their relative abundance in the different types of wheats.

Marker peptides for gliadin amount	Gliadin Type	Relative amount (durum)	Relative amount (common)
VPVPQLQPQNPSQQQPQEQVPL	α-gliadin	**++**	**++**
VRVPVPQLQPQNPSQQQPQEQVPL	α-gliadin	**+++**	**+++**
VRVPVPQLQPQNPSQQQPQEQVPLVQQQQF	α-gliadin	**+**	**+**

Immunogenic peptides identified	Gliadin Type	Relative amount (durum)	Relative amount (common)
QLQPFPQPQLPY	α-gliadin	+++	+
QLQPFPQPQLPYPQPQPF	α-gliadin	+	+
LQLQPFPQPQLPY	α-gliadin	+	+
LQLQPFPQPQLPYPQPQPF	α-gliadin	++	+
QLQPFPQPQLPYPQPQLPYPQPQPF	α-gliadin	nd	+
QLQPFPQPQLPYPQPHLPYPQPQPF	α-gliadin	nd	++
LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF	α-gliadin	nd	+++
LPFPQQPQQPFPQPQ	γ-gliadin	trace	trace

Toxic peptides identified	Gliadin type	Relative amount (durum)	Relative amount (common)
SHIPGLEKPSQQQPLPL	LMW-glutenin	+	+
VRVPVPQLQPQNPSQQQPQEQVPLVQQQQF	α-gliadin	+	+
QNPSQQQPQEQVPLVQQQ	α-gliadin	+	+
VPVPQLQPQNPSQQQPQEQVPL	α-gliadin	++	++
VRVPVPQLEPQNPSQQQPQEQVPL	α-gliadin	+	+
VRVPVPQLQPQNPSQQQPQEQVPL	α-gliadin	+++	+++
VRFPVPQLQPQNPSQQQPQEQVPL	α-gliadin	+	+
PSSQVQWPQQQPVPQ	γ-gliadin	+	+
NMQVDPSGQVQWPQQQPF	γ-gliadin	+	+

+++: very abundant; ++: abundant; +: detectable; nd: not detectable.

4.3.4 Pathogenic peptides for people affected by the celiac disease

Some of the peptides identified in the digested extracts of wheat varieties contain epitopes or toxic sequences known in literature to elicit, respectively, the adaptive and the innate immune system of celiac patients. These peptides are shown in Table 4.3.

Immunogenic peptides contain one or more epitopes, whose sequence is recognized by HLA-DQ2 or HLA-DQ8 cells: the epitopes found in several identified peptides were PFPQPQLPY (glia- αI), PQPQLPYPQ (glia- αII),[29] PYPQPQLPY (glia- αIII)[30] and FPQQPQQPF (glia- αII).[31] The first two epitopes were contained in a peptide identified as immunodominant after wheat challenge in celiac people in the study of Tye-Din et al. (2010).[32] Very interestingly, the immunodominant HLA DQ8-restricted alpha-gliadin T cell epitope (QGSFQPSQ), proposed to be one of the major triggers of celiac disease[33] was not found in the digested extract, probably because it contains a cleavage site for pepsin and chymotrypsin, that break the peptide bond at the C-terminal of phenylalanine.

Some short sequences (PSQQ, QQQP, QPYP, QQPY) proposed in literature to be toxic through stimulation of the innate immune response via interleukin-15 production[34] were found in the digested prolamin extracts. From in vitro studies on culture from jejunum biopsy, it has emerged that α-, γ- and ω-gliadins have a decreasing toxicity and, since the N-terminal sequence is similar also in barley and rye prolamins, are probably these amino acid sequences that determine toxicity.[35]

The single gastric phase was also modelled, by digesting with pepsin only: in this case only some toxic peptides were generated (VRVPVPQLQPQNPSQQQPQEQVPLVQQQQF, VPVPQLQPQNPSQQQPQEQVPL), and no peptides containing immunogenic sequences were detected. Thus from these experiments it clearly results that it is the combined action of the gastric and intestinal phase which lead to an increased pathogenicity of the wheat for people affected by celiac disease.

4.3.5 Quantification of the identified peptides

Quantification of the peptides in the digested extracts was carried out by adding to the samples, before digestion, 5 µl of a 0.477 mM solution of LQLQPF(d_5)PQPQLPY, used as internal standard. Given the similarity of molecular weight and aminoacidic composition with the other immunogenic and toxic peptides, the response factor was considered to be 1 for all the peptides quantified.

In order to verify the ability of the peptides chosen for estimating the amount of gliadin to act as suitable predictors, different amount of a commercial standard of gliadin underwent simulated gastrointestinal digestion and the marker peptides were quantified. A good correlation was found between the gliadin amounts and these peptides, indicating their suitability to be used as marker for the general gliadin content (data not shown). The quantitative data concerning marker peptides for the amount of gliadin in the different varieties are reported in Figure 4.2.

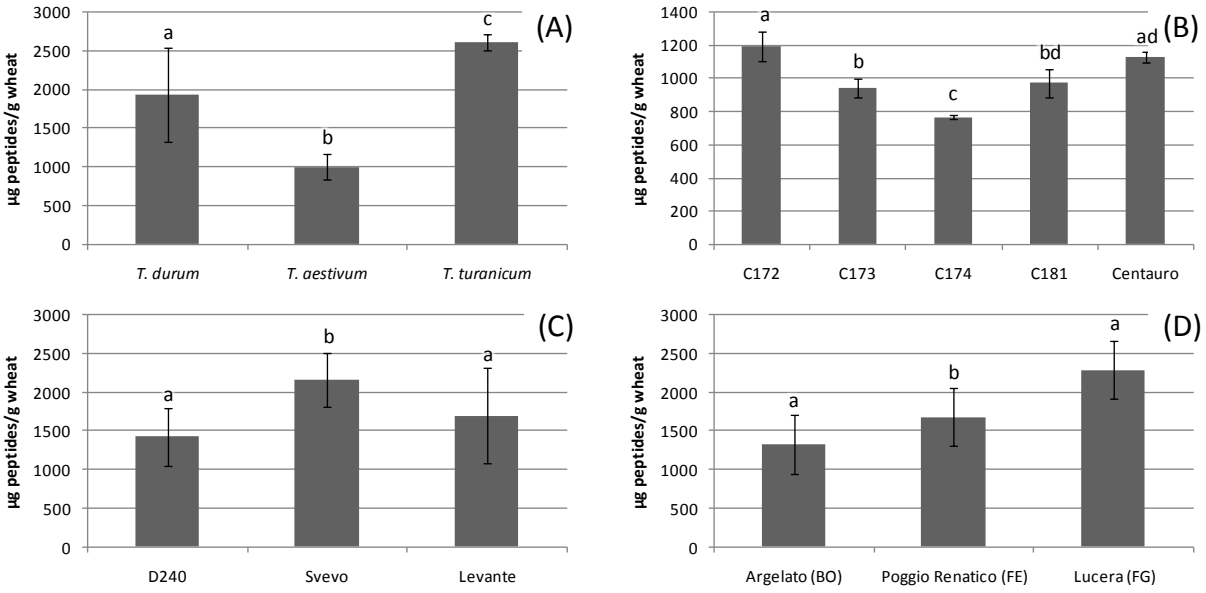

Figure 4.2 Total amount of the marker peptides for α-gliadins mediated for: species (A), *T. aestivum* variety (B), *T. durum* variety (C) and cultivation area (D). Significance levels are discussed in text.

The content of α-gliadins markers was found to be highly variable among the different samples, and significant differences (p<0.05) between the species were found (*T. turanicum*>*T.durum*>*T. aestivum*, Figure 4.2A). However, the α-gliadin content was calculated multiplying the moles of marker peptides per a medium MW of 31000 g/mol ranged from 24 to 34% of total prolamin content for *T. aestivum* samples, from 31 to 73% for *T. durum* and from 67 to 70% for *T. turanicum*. Among the *Triticum turgidum* spp *durum* varieties, the content of peptides marker for α-gliadin was significantly higher in Svevo (p<0.01) than in Levante and D240 (Figure 4.2C). Anyway, also the cultivation areas and/or farming practices clearly affect this parameter: in fact, the samples from Lucera (FG) showed the higher content of marker peptides for α-gliadin, while Argelato (BO) showed the lowest one (p<0.01, Figure 4.2D). Among the samples of *T. aestivum*, the highest amount of gliadin-derived peptides was observed for varieties C172 and Centauro, whereas the lowest for samples from varieties C173 and C174. Quite notably, samples 19, 20, 21 and 22 derive from cross-fertilization aimed at having a lower α-gliadin content (*T. aestivum* sample 23 was taken in the market to be used as reference). Of the four analyzed *T. aestivum* varieties, only sample 20 (variety C173) and 21 (variety C174) have reached this purpose (p<0.05, Figure 4.2B). It is interesting to note that the α-gliadin content determined has a positive correlation (0.887, p<0.01) with the total protein content and with the gliadin amount (0.900, p<0.01).

The quantitative data concerning pathogenic peptides for the people suffering of celiac disease are reported in Figure 4.3.

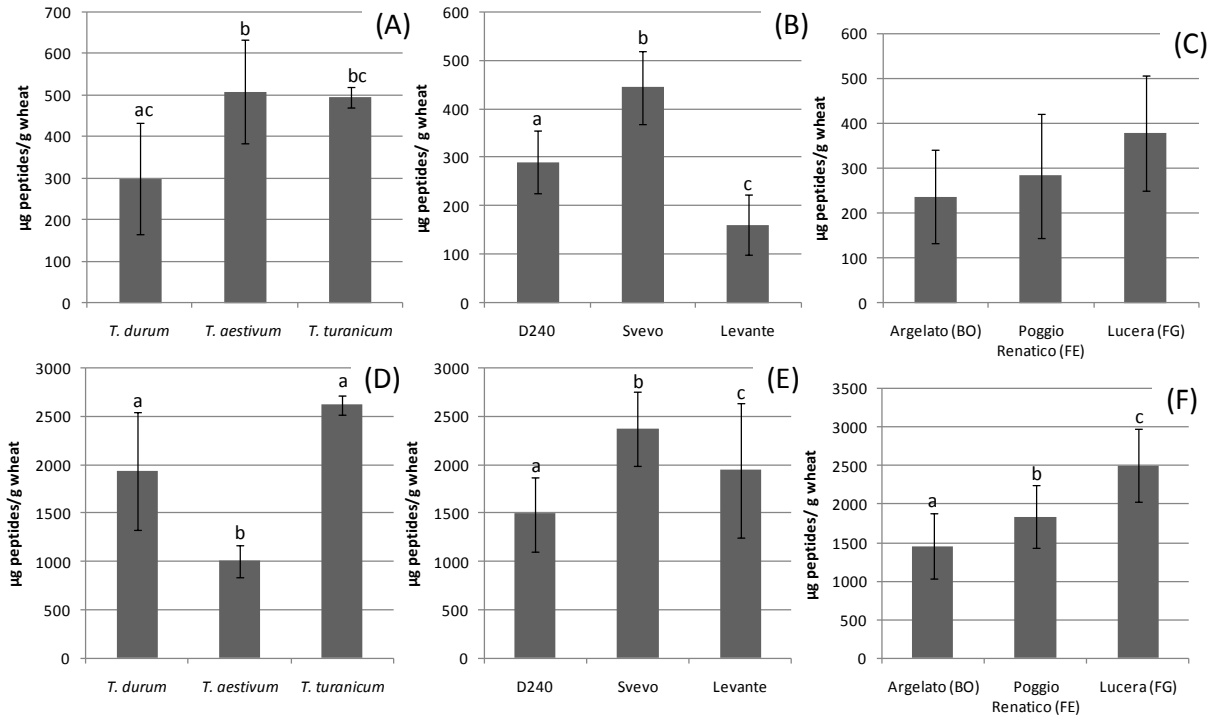

Figure 4.3 Total amount of immunogenic peptides mediated for: species (A), variety (B) and cultivation area (C). Total amount of toxic peptides mediated for: species (A), variety (B) and cultivation area (C). Significance levels are discussed in text.

As it can be seen, all wheat varieties generated immunogenic peptides upon digestion, but with qualitative and quantitative differences (Figure 4.3A): *T. durum* samples generated a significant lower (p<0.01) amount of immunogenic peptides than *T. aestivum*. Three of the immunogenic peptides identified (LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF, QLQPFPQPQLPYPQPQLPYPQPQPF and QLQPFPQPQLPYPQPHLPYPQPQPF) were found to be present only in *T. aestivum*, in which were the main immunogenic components found. The first peptide, also known as the 33-mer peptide, contains 6 epitopes partially overlapped and it has been the subject of many previous studies because is one of the major immunodominant peptides. According to our and previous results, this peptide was the most abundant among the immunogenic peptides in digests coming from *T. aestivum* samples. The second and the third peptide contain, respectively, 4 and 2 epitopes. Instead, the peptide (L)QLQPFPQPQLPY seems to be most abundant in digests coming from *T. turgidum* spp *durum* and *T. turanicum*.

The samples of *T. durum* with a lower content of immunogenic peptides belonged to the Levante variety, followed by D240 and Svevo (p<0.01, Figure 4.3B). If genetics seems to be the main determinant for the production of immunogenic peptides upon digestion, cultivation area and/or farming practices also seems to have a role: the immunogenic peptides content in fact reflects the trend showed for α-gliadin content (Argelato<Poggio Renatico<Lucera, Figure 4.3C). The amount of immunogenic peptides showed a positive correlation with the total protein content (0.571, p<0.01), with the global prolamine content (0.613, p<0.01) and with the marker for α-gliadin content (0.715, p<0.01).

As far as toxic peptides are concerned, all wheat varieties generated toxic peptides upon digestion, although their content was found to be highly variables among species (Figure 4.3D): samples from *T. aestivum* in fact

generated fewer peptides than *T. durum* and *T. turanicum* (p<0.01). The most represented toxic peptide in the digests was always found to be VRVPVPQLQPQNPSQQQPQEQVPL, followed by its digestion product VPVPQLQPQNPSQQQPQEQVPL. The variety with the higher amount of toxic peptides was Svevo, followed by Levante and D240 (p<0.01). Thus, again, genetics seems to be one of the most important determinant for the generation of toxic peptides, but also in this case cultivation area and/or farming practices seems to play a role, with an increasing content in the order for Argelato<Poggio Renatico<Lucera (p<0.01). Also in this case, the amount of toxic peptides shows a positive correlation with the total protein content (0.889, p<0.01), with the total gliadin amount (0.993, p<0.01) and with the peptides marker for α-gliadin amount (0.990, p<0.01).

So, the content of pathogenic peptides correlate both with the total protein content and with gliadin amount, but this correlation seems to be slightly weaker for immunogenic peptides. The correlation with peptides marker for α-gliadin content is stronger for toxic peptides because they belong to the N-terminal region of α-gliadin (as the marker for the gliadin amount), that is the most conserved. Instead, the higher variability in the region were are mostly present immunogenic peptides lead to a weaker, but clear, correlation. So, not only the protein amount is important, but also the type and composition of the nitrogen fraction, as demonstrated by the opposite trend for toxic and immunogenic peptides between *T. aestivum* and *T. durum* and, among *T. durum* varieties, between D240 and Levante.

4.4 CONCLUSIONS

In different wheat varieties the main peptides produced by simulated gastrointestinal digestion of the prolamin fraction were identified and quantified. Different wheat varieties led to variable outcomes in terms of amount and type of peptides produced. As far as pathogenic peptides related to celiac disease are concerned, these data are of high interest also from the epidemic point of view, since they indicate a high variability among the wheat samples, and this might have a different impact on celiac disease development. Although all wheat samples were found to generate pathogenic peptides upon digestion (thus no variety can be considered “safe” for celiac patients), peptide mixtures derived from the digestion of *T. aestivum* samples were found to contain less toxic peptides than *T. durum* and *T. turanicum*, while immunogenic peptides were less abundant in *T. durum* digests than in *T. aestivum* and *turanicum*. Among *T. durum* samples, on the other hand, a quite high variability was observed: D240 was the variety with the lower content of toxic peptides, while Levante was the one with a lower content of immunogenic peptides. Albeit not “safe” for celiac patients, the use of this variety, for example in the formulations of baby food, could reduce the exposure during the most critical period for the developing of the disease (Ivarsson et al., 2002).[36]

As far as cultivation area/farming practices are concerned, the impact is less evident, but somehow present: given the variety, which is the most important determinant, some areas of cultivation seems to promote a greater content in α-gliadins, toxic and immunogenic peptides in digests, whereas in digested wheat cultivated in other places less of these peptides seem to be present. Obviously, a more extensive study (in term of varieties and cultivation areas tested) is needed, besides a comparison between different years of cultivation. Moreover, further studies will be needed in order to better define in vivo the role of the identified peptides and to also define the best farming practices which can further lower wheat pathogenicity.

4.5 ACKNOWLEDGEMENTS

The “AGER – Agrifood and Research” consortium is gratefully acknowledged for funds provided to this work through the project “From seed to pasta”. Dr. Andrea Massi, from Società Produttori Sementi S.p.A., is also gratefully acknowledged for his useful suggestions during the draft of this manuscript.

---

[1] Osborne TR (1907) *The protein of the wheat kernel*. Publ. 84. Carnegie Inst., Washington DC.

[2] Stepniak D, Spaenij-Dekking L, Mitea C, Moester M, de Ru A, Baak-Pablo R, van Veelen P, Edens L, Koning F (2006) Highly efficient gluten degradation with a newly identified prolyl endoprotease: implications for celiac disease. *Am J Physiol-Gastr L* 291, 621-629.

[3] Shan L, Molberg O, Parrot I, Hausch F, Filiz F, Gray GM, Sollid LM, Khosla C (2002) Structural basis for gluten intolerance in celiac sprue. *Science* 297, 2275-2279.

[4] Trier JS (1998). Diagnosis of celiac disease. *Gastroenterology* 115, 211–216.

[5] Sollid LM, Lie BA (2005) Celiac disease genetics: current concepts and practical applications. *Clin Gastroenterol Hepatol* 3, 843-851.

[6] Gianfrani C, Auricchio S, Troncone R (2005) Adaptive and innate immune responses in celiac disease. *Immunol Lett* 99, 141-145.

[7] Cornel, HJ (1996) Coeliac disease: a review of the causative agents and their possibly mechanisms of action. *Amino Acids* 10, 1-19.

[8] De Ritis G, Auricchio S, Jones HW, Lew EJL, Bernardin JE, Kasarda DD (1988). “In vitro” (organ culture) studies of the toxicity of specific α-gliadin peptides in celiac disease. *Gastroenterology* 94, 41–49.

[9] Londei M, Ciacci C, Ricciardelli I, Vacca L, Quarantino S, Maiuri L (2005). Gliadin as a stimulator of innate responses in celiac disease. *Mol Immunol* 42, 913-918.

[10] Ciccocioppo R, Di Sabatino A, Corazza GR (2005). The immune recognition of gluten in celiac disease. *Clin Exp Immunol* 140, 408-416.

[11] Farrell RJ, Kelly CP (2002) Celiac sprue. *New Engl J Med* 346, 180-188.

[12] Van de Wal Y, Kooy I, Van Veelen P, Pena S, Mearin L, Papadopoulos G, Koning F (1998) Selective deamidation by tissue transglutaminase strongly enhances gliadin specific T-cell reactivity. *J Immunol* 161, 1585-1588.

[13] Molberg O, McAdam SN, Korner R, Quarsten H, Kristiansen C, Madsen L, Fugger L, Scott H, Norén O, Roepstorff P, Lundin KEA, 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, 713-717.

[14] Nilsen EM, Jahnsen FL, Lundin KE, Johansen FE, Fausa O, Sollid LM, Jahnsen J, Scott H, Brandtzaeg P (1998) Gluten induces an intestinal cytokine response strongly dominated by interferon gamma in patients with celiac disease. *Gastroenterology* 115, 551-563.

[15] Matysiak-Budnik T, Candalh C, Dugave C, Namane A, Cellier C, Cerf-Bensussan N, Heyman M (2003) Alterations of the intestinal transport and processing of gliadin peptides in celiac disease. *Gastroenterology* 125, 696-707.

[16] Sollid LM (2000) Molecular basis of celiac disease. *Annu Rev Immunol* 18, 53–81.

[17] Shewry PR, Tatham AS, Kasarda DD (1992) Cereal proteins and celiac disease. In *Celiac Disease*, ed. M Marsh, Blackwell: Oxford, pp. 305–348

[18] Tye-Din JA, Stewart JA, Dromey JA, Beissbarth T, van Heel DA, Tatham A, Henderson K, Mannering SI, Gianfrani C, Jewell DP, Hill AVS, McCluskey J, Rossjohn J, Anderson RP (2010). Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci Transl Med* 2, 41-51.

[19] Vader W, Kooy Y, Van Veelen P, De Ru A, Harris D, Benckhuijsen W, Peña S, Mearin L, Drijfhout JW, Koning F (2002). The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides. *Gastroenterology* 122, 1729–1737

[20] Clemente MG, De Virgilis S, Kang JS, Macatagney R, Musu MP, Di Pierro RM, Drago S, Congia M, Fasano A (2003). Early effects of gliadin on enterocyte intracellular signaling involved in intestinal barrier function. *Gut* 2, 218–23.

[21] Wieser H (1996) Relation between gliadin structure and celiac toxicity. *Acta Pædiatr* 412, 3-9

[22] Mamone G, Ferranti P, Rossi M, Roepstorff P, Fierro O, Malori A, Addeo F (2007) Identification of a peptide from α-gliadin resistant to digestive enzymes: implications for celiac disease. *J Chromatog B* 855, 236-241

[23] COMMISSION REGULATION (EC) No 41/2009 of 20 January 2009 concerning the composition and labelling of foodstuffs suitable for people intolerant to gluten

[24]Catassi C, Kryszak D, Bhatti B, Sturgeon C, Helzlsouer K, Clipp SL, Gelfond D, Puppa E, Sferruzza A, Fasano A (2010). Natural history of celiac disease autoimmunity in a USA cohort followed since 1974. *Ann Med* 42, 530-538

[25] Gobetti M, Rizzello GM, Di Cagno R, De Angelis M (2007) Sour-dough lactobacilli and celiac disease. *Food Microbiol* 24, 187-196.

[26] Shan L, Molberg O, Parrot I, Hausch F, Filiz F, Gray GM, Sollid LM, Khosla C (2002) Structural basis for gluten intolerance in celiac disease. *Science* 297, 2275-2279

[27] Cavatorta V, Sforza S, Aquino G, Galaverna G, Dossena A, Pastorello EA, Marchelli R (2010) In vitro gastrointestinal digestion of the major peach allergen Pru P 3, a lipid transfer protein: molecular characterization of the products and assessment of their IgE binding abilities. *Mol Nutri Food Res* 54, 1452-1457

[28] Wieser H (2000) Comparative investigations of gluten proteins from different wheat species. I. Qualitative and quantitative composition of gluten protein types. *Eur Food Res Technol* 211, 262-268.

[29] Arentz-Hansen H, Korner R, Molberg O, Quarsten H, Vader W, Kooy YM, Lundin KE, Koning F, Roepstorff P, Sollid LM, McAdam SN (2000) The intestinal T cell response to alpha-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *J Exp Med* 191, 603-612.

[30] Arentz-Hansen H, McAdam SN, Molberg O, Fleckenstein B, Lundin KE, Jorgensen TJ, Jung G, Roepstorff P, Sollid LM (2002) Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. *Gastroenterology* 123, 803-809.

[31] Spaenij-Dekking L, Koning F (2005) *The molecular basis of celiac disease.* In: Molecular Autoimmunity 3, 141-150.

[32] Tye-Din JA, Stewart JA, Dromey JA, Beissbarth T, van Heel DA, Tatham A, Henderson K, Mannering SI, Gianfrani C, Jewell DP, Hill AVS, McCluskey J, Rossjohn J, Anderson RP (2010) Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci Trasl Med* 2, 41-51.

[33] Van de Wal Y, Kooy YM, Van Veelen PA, Pena SA, Mearin LM, Molberg O, Lundin KE, Sollid LM, Mutis T, Benckhuijsen WE, Drijfhout JW, Koning F (1998) Small intestinal T cells of celiac disease patients recognize a natural pepsin fragment of gliadin. P *Natl Acad Sci USA* 95, 10050-10054.

[34] Cornell HJ, Wills-Johnson D (2001) Structure activity relationship in celiac-toxic gliadin peptides. *Amino Acids* 21, 243-253.

[35] Vader W, Kooy Y, Van Veelen P, De Ru A, Harris D, Benckhuijsen W, Peña S, Mearin L, Drijfhout JW, Koning F (2002) The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides. *Gastroenterology* 122, 1729–1737.

[36] Ivarsson A, Hernell O, Stenlund H, Åke Persson L (2002) Breast-feeding protects against celiac disease. *Am J Clin Nutr* 75, 914-921.

5 GENETIC AND ENVIRONMENTAL FACTORS AFFECTING PATHOGENICITY OF WHEAT AS RELATED TO CELIAC DISEASE

Based on: Prandi B, Mantovani P, Galaverna G, Sforza S (2014) Genetic and environmental factors affecting pathogenicity of wheat as related to celiac disease. Journal of Cereal Science 59(1): 62-69.

5.1 INTRODUCTION

Wheat is one of the most harvested cereals in the world and represents the most important source of complex carbohydrates in the human diet (http://www.fao.org/docrep/W8079E/w8079e0g.htm). Wheat is principally cultivated in North Africa, Western Europe, North America and middle-East of Asia.[1] The main species used for human consumption are *Triticum aestivum* (or bread wheat), that is mainly used for baked products, and *Triticum durum* Desf. usually utilized in pasta production and for some traditional breads of the Mediterrean basin.[2] In addition to direct consumption, changes in wheat starch and gluten content and composition have expanded its use also for food and non-food additives.[3] Gluten represents the 80% of the protein fraction in wheat kernel and plays a key role in bread production and in pasta industry, conferring elasticity to dough and firmness of pasta during the cooking process. Gluten is composed of an alcohol soluble fraction (gliadins) and an alkaline/acid soluble-fraction (glutenins) that needs denaturing and reducing condition for their extraction;[4] a correct balance between this two fractions is of outmost importance for a correct dough development.[5] Gliadins are small and monomeric proteins that contribute to the dough viscosity and extensibility, while glutenins are present in polymers of high molecular mass formed by the establishment of disulphide bonds between the single monomeric subunities and confers strength and elasticity to dough.[6] To improve the rheological properties of wheat flour, breeders selected varieties with genetic features promoting high protein and gluten content.[7]

As well as for its rheological properties, there is an increasing interest in studying gluten proteins because they are the main factor causing celiac disease, an autoimmune enteropathy. Besides the consumption of gluten, celiac disease needs a genetic predisposition due to an antigluten T-cell response associated with specific hystocompatibility antigens, namely HLA-DQ2.5 and HLA-DQ8, and the loss of the oral tolerance to gluten.[8] The immune response in celiac subjects is due both to toxic motives triggering the innate immune system and to epitopes influencing the adaptive system.[9] These peptides are generated in the gastrointestinal tract when gluten is partially proteolyzed by digestive enzymes. Gluten breakdown is not complete in human digestion because of the high content of proline whose cyclic structure hinders the access to the peptide bond: this led to the formation of long peptides involved in this autoimmune disorder.[10] Celiac disease is nowadays one of the most common chronic health disorder. In the recent past the prevalence of this pathology ranged from 0.5 to 1.26% of children and adolescents,[11] but in the last decades there was an increase of celiac cases even in the adulthood.[12] There are many hypothesis that could explain this trend such as the higher amount of gluten ingested, its quality, the reduction of the leavening time during the baking process or changes of the intestinal microbiota.[13, 14] Recent studies have demonstrated the presence of a different content of celiac disease epitopes in modern and old wheat varieties,[15] therefore it is questionable whether there is a correlation between breeding practices and the increasing of celiac disease incidence. Although current trend of breeding is to develop wheat varieties with higher protein and gluten content, it is alsoknown that the modest influence of evolutionary pressure on gliadins

and glutenins determinate their great variability among wheat cultivarsin comparison with functional proteins. The different contribution of the AA, BB and DD genomes in the production of epitopes-containing gliadins have already been demonstrated,[16] and from genetic studies large differences emerged in T cells stimulatory epitopes content in proteins from different hexaploid and tetraploid accessions.[17] These differences were also confirmed by T-cells proliferation tests and monoclonal antibodies competition assays.[18]

In a recent work from our group,[19] the quantification of celiac-related peptides in samples from various *Triticum* samples demonstrated strong differences among the varieties tested. Some samples, belonging to the same varieties and/or cultivation area, showed a lower α-gliadin content, and a smaller amount of peptides involved in adaptive and innate immune response. This observation can be useful in order to identify varieties with a lower content of pathogenic peptides, thus reducing the exposure of non celiac people (especially of young population, e.g. baby foods) and the probability of developing the disease.

In order to further confirm and expand those results, in this work, a set of durum wheat inbred lines and varieties was utilized to extract the prolamine fraction, in order to investigate the effect of genotype and environment on the content of peptides containing sequences involved in adaptive and innate immune response generated after simulated gastrointestinal digestion. A set of inbred lines and cultivars with low genetic diversity, together with a set of germplasm accessions representative of the Mediterranean area with higher genetic diversity, cultivated in different places, were analyzed for their ability to produce pathogenic peptides upon gastrointestinal digestion. The same samples were analyzed for total nitrogen content, and fractionated with the Osborne procedure in order to verify a possible correlation between the total protein amount and the gliadin/glutenin content.

5.2 MATERIAL AND METHODS

5.2.1 Plant Material

All the durum wheat samples were provided by Società Produttori Sementi S.p.A. (Argelato, Bologna, Italy). Two sets of samples were collected aiming to evaluate the role of the environment and of the genotype on the pathogenic peptides content.

Table 5.1 Samples of commercial varieties analyzed for pathogenic peptides production after in vitro digestion of the prolamine extract.

	Cultivation area	Argelato (BO) North Italy			Falconara (AN) Central Italy			Lucera (FG) South Italy		
	Replicate	1	2	3	1	2	3	1	2	3
Genotype	Type	ID number			ID number			ID Number		
Aureo	Cultivar	5914	5939	5971	6067	6088	6120	6216	6238	6262
F255	Inbred line	5911	5933	5952	6068	6093	6114	6218	6247	6273
F312	Inbred line	5910	5948	5959	6071	6096	6123	6208	6241	6272
F371	Inbred line	5924	5946	5965	6054	6079	6118	6211	6235	6271
Levante	Cultivar	5916	5925	5958	6058	6091	6106	6204	6249	6269

The set of durum sample was composed of:

a) a total number of 45 samples from five inbred line and cultivars (Aureo, F255, F312, F371, Levante; Table 5.1) cultivated in three locations well-representing the different Italian wheat cultivation areal (Argelato-North Italy, Falconara-AN Central Italy, Lucera-South Italy); for each sample, three replicates/location were collected.

b) Because of the modest genetic diversity among these five samples in terms of pedigree, a second set of 25 samples from a collection of durum wheat accessions were selected and analyzed for the content of peptides involved in celiac disease generated after in vitro digestion of the prolamine extract.

25 samples (Table 5.2) were selected from a collection of durum wheat accessions representative of the germplasm cultivated in the Mediterranean basin, with a suitable level of genetic variation and deeply characterized from a molecular point-of-view (Durum Panel).[20]

Table 5.2 Samples of the Durum Panel analyzed for pathogenic peptides production after in vitro digestion of the prolamine extract, classified by proximity in the phylogenetic dendrogram.

Name	Pedigree	IDUWUE genotype CODE	Origin
ICARDA AND ITALIAN ACCESSIONS FOR DRYLAND AREAS HAURANI/EITI/OMRABI			
YOUNES 1	Mrb3/Mna-1	03-187	ICARDA
AW12/BIT	AWAL/BIT	03-60	ICARDA
BLK2	CRANE/STK	03-133	ICARDA
OMSNIMA 1		03-169	ICARDA
GIDARA 2	ICD90-0179-ABL-0AP-2AP-0AP-3AP-0AP	03-141	ICARDA
ICARDA ACCESSIONS FOR TEMPERATE AREAS CHAM 1			
MESSAPIA	MEXA/CRANE//TITO	03-102	ITALIA
MARZAK		03-55	INRAMOROCCO
MOROCCO 1809=MAROUANE	INRAM.1809	03-50	INRAMOROCCO
H-MOULINE (MOR)/CHABA 88	HML/CHHB88	03-65	ICARDA
MOULSABIL-2		03-68	ICARDA
CIMMYT ACCESSIONS YAVAROS 79/KARIM/BITTERN			
KARIM =YAVAROS 79	JORI(D21563)/AA//FGO	03-64	ICARDA
MOROCCO 1805=NASSIRA	INRAM.1805	03-47	INRAMOROCCO
YASMINE		03-59	INRAMOROCCO
DUILIO	CAPPELLI//ANHINGA/FLAMINGO	03-91	ITALIA
BRONTE	BERILLO/LATINO	03-82	ITALIA
CIMMYT ACCESSIONS ALTAR 84/GALLARETA			
PORTO 5	CHEN/ALTAR_84//JORI_C_69	03-17	CIMMYT
CIMMYT-136	FOCHA_1/5*ALAS	03-21	CIMMYT
BOMBASI		03-36	IRTA-SPAIN
ASTIGI		03-31	IRTA-SPAIN
IRIDE	ALTAR_84/IONIO	03-97	ITALIA
ITALIAN ACCESSIONS FOUNDERS: CAPPELLI-VALNOVA-CRESO			
MEXICALI 75	GDOVZ469/3/JO_1//ND61.130/LDS	03-103	CIMMYT
ANGRE		03-120	ICARDA

BRAVADUR		03-81	DESERT DURUM
SENADUR		03-44	IRTA-SPAIN
COLORADO	P_92/932-2	03-86	DESERT DURUM

5.2.2 Reagent and solvents

Deionised water was obtained from a Millipore Alpha Q-Waters purification system (Billerica, MA, USA). Pepsin from porcine gastric mucosa, trypsin from porcine pancreas, α-chymotrypsin from bovine pancreas, sodium dihydrogen phosphate, acetonitrile, ethanol, Fmoc-glutamine(Trt)-OH, piperidine, diisopropylethylamine, dichloromethane, triisopropylsilane, tyrosine, phenylalanine, tris(hydroxymethyl)-aminomethane, urea, dithiothreitol and diethyl ether were purchased from Sigma-Aldrich (St. Louis, MO, USA). Hydrochloric acid (37% V/V), sulphuric acid (96%), phosphosulphuric acid, copper (I) oxide, sodium hydroxide, bromocresol green, methyl red, 1-propanol and dimethylformamide were purchased from Carlo Erba (Milan, Italy). Fmoc-leucine-OH, Fmoc-proline-OH, HBTU and Fmoc-tyrosine(tBu)-Wang resin were purchased from Novabiochem (Darmstadt, Germany). Formic acid and trifluoroacetic acid were purchased from Acros Organics (Geel, Belgium). Selenium, silicone antifoam and sodium sulphate were purchased from Thompson and Capper (Hardwick, UK). Boric acid was purchased from Fluka (Buchs, Switzerland). Hydrochloric acid 0.1 N and sodium chloride were purchased from AnalaR Normapur (Milan, Italy).

5.2.3 Synthesis of the internal standard

The peptide LQLQPF(d_5)PQPQLPY (isotopically labeled on the phenylalanine residues) was synthesized on solid phase according to Fmoc/t-butyl strategy using a Syro I Fully Automated Peptide Synthesizer (Biotage, Uppsala, Sweden). The peptide was cleaved from the Wang-resin using a TFA:TIS:H_2O (95:2.5:2.5) solution and purified using a semipreparative RP-HPLC-UV (λ=280 nm). The purified product was quantified by spectrophotometric method at 280 nm using an external calibration curve made with an equimolar solution of tyrosine and phenylalanine.

5.2.4 Osborne fractionation

Albumins, globulines and gliadins extraction was carried out as described by Lookhart and Bean.[21] Glutenins were extracted as described by Wieser et al.[22] In order to assess the protein content of the extracted fractions, a Q-bit fluorometer (Invitrogen, Grand Island, NY, USA) was used. The working solution was prepared by mixing 1 µl of the Quant-it reagent in 199 µl of Quant-it buffer; 2 µl of each extract were mixed with 198 µl of the working solution. The sample was left for 15 min in the dark before reading the protein content with the fluorometer.

5.2.5 Extraction and digestion

Many studies hinted at the prolamin fraction[23] as the major external trigger in celiac disease, which is mainly composed of gliadins. The term prolamine derives from the high content in proline and glutamine of this class of proteins. So, the prolamin fraction was extracted from 1 g of grounded wheat kernels treated with 10 ml of a 70% ethanol-aqueous solution for 2.5 hours. One milliliter of extract was spiked with 10 µl of internal standard peptide (0.5 mM) and dried under nitrogen flux. Samples were in vitro digested adding a solution containing the three main gastric and pancreatic endoproteases (pepsin, trypsin and chymotrypsin). For simulating the gastric phase,

each sample was reconstituted with 500 µl of HCl 10 mM (pH=2) and added with 40 µl of a pepsin solution (1 mg/ml). After three hours of incubation at 37 °C, the pH was set up to 7.2 by adding a phosphate buffer solution (100 mM) and the final solution was added with 40 µl of a trypsin and chymotrypsin solution (1 mg/ml). After 4 hour of incubation at 37°C samples were dried under nitrogen flux. The possible contribution of brush border proteases and post-absorption peptide processing was not considered in this study.

5.2.6 UPLC/ESI-MS analysis

In order to quantify the peptides in the digested mixtures, the dried digested samples were redissolved in 300 µl of HCOOH 0.1% and separated by a RP column (ACQUITY UPLC BEH 300 C18 1.7 µm 2.1*150 mm) in an UPLC/ESI-MS system (UPLC Acquity Waters with a single quadrupole mass spectrometer Waters Acquity Ultraperformance) using a gradient elution. Eluent A was a water solution of formic acid (0.1%) and acetonitrile (0.2%), and eluent B was an acetonitrile solution with formic acid (0.1%). The following protocol for the gradient elution was carried out: 0-7 min 100% eluent A; 7-50 min from 100% to 50% eluent A; 50-52.6 min 50% eluent A; 52.6-53 min from 50% to 0% eluent A; 53-58.2 min 0% eluent A; 58.2-59 min from 0% to 100% eluent A; 59-72 min 100% eluent A. The digested sample extracts were analyzed with UPLC/ESI-MS in the Full Scan mode (flow 0.2 ml/min; analysis time 72 min; column temperature 35°C; sample temperature 18 °C; injection volume 5 µl; acquisition time 7-58.2 min; ionization type positive ions; scan range 100-2000 m/z; capillary voltage 3.2 kV; cone voltage 30 V; source temperature 100°C; desolvation temperature 200°C; cone gas flow 100 l/h; desolvation gas flow 650 l/h), the characteristic ions of every peptide were extracted, obtaining eXtract Ion Chromatograms (XICs), in which the identified peptides and internal standard LQLQPF(d_5)PQPQLPY were integrated with the MassLynx software. The quantification value was obtained as the ratio peptide-area/standard-area multiplied by the moles of internal standard.

5.2.7 Total protein content

The total protein content (TPC) was determined according to the Kjeldhal method: 1 g of sample was added with 7 ml of sulphuric acid, 10 ml of phosphosulfuric acid, copper (I) oxide, selenium (0.2 mg/ml), sodium sulphate (100 mg/ml) and silicone antifoam (3.5 µg/ml). The mineralization step was carried out at 420 °C for one hour. The hydrolyzed sample was added with 50 ml of water and 100 ml of an aqueous solution of NaOH (32 %). Ammonia was condensed in 25 ml of aqueous solution of boric acid (4 %) added with bromocresol green and methyl red; the solution was titrated with 0.1 N HCl. The TPC was determined multiplying the obtained value per 5.7, a commonly used conversion factor allowing the estimation of the TPC starting from the total nitrogen content (http://www.fao.org/docrep/006/y5022e/y5022e03.htm).

5.3 RESULTS AND DISCUSSION.

5.3.1 Protein quantification

All the samples were first analyzed for total protein content (TPC) determined by the Kjeldhal method and the sum of protein classes determined in Osborne fractions (results are reported in Table 5.3).

Table 5.3 Total protein content (TPC) and protein distribution profile in commercial varieties and in-bred lines analyzed.

	TOTAL PROTEINS (%)		ALBUMINS (%)		GLOBULINS (%)		GLIADINS (%)		GLUTENINS (%)	
SAMPLE	AVERAGE	STD DEV	AVERAGE	STD DEV	AVERAGE	DEV STD	AVERAGE	DEV STD	AVERAGE	STD DEV
5910	13.37	0.06	1.16	0.09	1.06	0.03	3.91	0.18	6.75	0.34
5911	14.53	0.34	1.40	0.16	1.39	0.01	4.15	0.00	7.24	0.06
5914	15.67	0.03	2.25	0.06	1.85	0.01	6.83	0.25	8.86	1.42
5916	14.49	0.96	1.24	0.02	1.02	0.01	4.84	0.17	6.43	0.16
5924	13.25	0.23	1.34	0.01	1.16	0.08	4.25	0.18	6.03	0.13
5925	13.33	0.11	1.30	0.04	0.97	0.01	4.82	0.20	6.62	0.29
5933	13.49	0.11	1.66	0.13	1.37	0.02	5.14	0.17	5.22	0.30
5939	16.01	0.17	1.94	0.07	1.85	0.23	6.30	0.13	9.13	0.13
5946	13.69	0.17	1.29	0.06	1.19	0.08	4.18	0.09	5.27	0.14
5948	13.05	0.17	1.37	0.05	1.13	0.03	3.98	0.03	6.43	0.22
5952	13.69	0.06	1.52	0.04	1.31	0.13	4.20	0.35	8.12	0.35
5958	14.25	0.17	1.32	0.02	1.04	0.02	4.82	0.12	7.09	0.28
5959	13.45	0.06	1.38	0.03	1.12	0.07	4.16	0.03	6.59	0.18
5965	12.86	0.56	1.25	0.02	1.14	0.07	4.37	0.16	5.43	0.10
5971	15.45	0.17	1.88	0.27	1.70	0.08	7.43	0.48	11.54	0.49
6054	12.54	0.56	1.14	0.01	1.08	0.07	1.01	0.41	5.48	0.13
6058	10.30	0.34	1.18	0.01	0.86	0.01	3.96	0.15	6.26	0.17
6067	12.46	0.23	1.83	0.26	1.54	0.18	4.53	0.03	8.84	0.64
6068	10.70	0.00	1.79	0.02	1.69	0.15	4.64	0.44	6.93	0.84
6071	11.54	0.06	1.34	0.02	1.05	0.03	3.74	0.14	5.56	0.18
6079	11.54	0.40	1.23	0.04	1.08	0.08	4.50	0.12	4.91	0.30
6088	12.14	0.34	1.73	0.08	1.73	0.25	5.01	0.20	7.17	0.11
6091	11.82	0.11	1.30	0.01	0.90	0.01	4.44	0.06	5.75	0.14
6093	11.90	0.34	2.22	0.44	1.43	0.12	4.26	0.15	4.68	0.12
6096	11.70	0.06	1.31	0.02	1.08	0.03	4.08	0.10	6.23	0.21
6106	12.02	0.17	1.32	0.03	1.00	0.03	4.70	0.07	5.60	0.50
6114	12.18	0.06	2.06	0.09	1.77	0.08	5.11	0.25	8.42	0.14
6118	12.50	0.28	1.27	0.03	1.09	0.08	3.95	0.08	5.43	0.02
6120	11.58	0.23	1.74	0.11	1.84	0.14	4.89	0.79	8.16	0.37
6123	12.82	0.06	1.37	0.02	1.05	0.04	3.92	0.03	5.79	0.09
6204	10.62	0.34	1.21	0.01	0.84	0.02	3.64	0.02	5.08	0.33
6208	11.78	0.06	1.10	0.02	0.96	0.05	3.61	0.01	6.54	0.40
6211	11.54	0.06	1.28	0.05	1.09	0.08	4.09	0.27	5.58	0.16
6216	11.86	0.40	1.76	0.06	1.28	0.07	4.79	0.54	7.44	0.43
6218	12.30	0.11	1.85	0.06	1.62	0.22	5.02	0.41	5.36	1.21
6235	13.21	0.28	0.96	0.05	1.07	0.08	4.13	0.05	5.91	0.29
6238	11.94	0.06	1.69	0.04	1.38	0.13	5.15	0.95	6.48	0.48
6241	12.50	0.06	1.17	0.04	0.97	0.03	3.36	0.10	6.41	0.78
6247	11.98	0.11	2.03	0.13	1.44	0.05	5.34	0.91	5.33	0.42
6249	12.18	0.28	1.20	0.04	0.87	0.02	3.78	0.08	5.40	0.10
6262	12.42	0.40	1.90	0.32	1.41	0.05	4.58	0.42	10.11	0.45
6269	12.26	0.40	1.20	0.02	0.82	0.01	4.01	0.02	5.56	0.72

6271	12.54	ND	1.19	0.03	1.07	0.08	4.19	0.07	5.94	0.28
6272	12.42	1.75	1.29	0.03	1.06	0.02	3.53	0.33	6.80	0.03
6273	11.86	0.06	2.00	0.09	1.62	0.09	5.52	0.53	5.82	0.07

These data allowed to evaluate the influence of total protein amount and distribution on the production of peptides implicated in celiac disease.

5.3.2 Pathogenic peptides quantification

The causal agent of celiac disease resides mainly in the gliadin fraction of gluten, where several peptides were shown to elicit a strong and rapid T-cells response in nearly all celiac patients.[24] Thus all the samples were subjected to simulated gastrointestinal digestion and the peptides generated were identified by LC/MS methodologies (details in the experimental section). The peptides containing sequences eliciting adaptive immune system identified in the mixtures were LQLQPFPQPQLPY, QLQPFPQPQLPY and QLQPFPQPQLPY (where the N-terminal glutamine underwent cyclization), all containing the epitope DQ2.5-glia-α1a (PFPQPQLPY, region 56-68 of α-gliadin, Arentz-Hansen et al., 2000).[25] The peptides containing sequences eliciting the innate immune system identified in the sequences were VPVPQLQPQNPSQQQPQEQVPL, VRVPVPQLQPQNPSQQQPQEQVPL, VRFPVPQLQPQNPSQQQPQEQVPL (all from the N-terminal region of α-gliadins and containing the toxic sequences QQQP and PSQQ)[26] and NMQVDPSGQVQWPQQQPF (belonging to the N-terminal region of γ-gliadin and containing the toxic sequence QQQP). The sequences PSQQ and QQQP are redundant in gluten proteins and for this reason they are used also used for gluten quantization in food products. The peptide LGQQQPFPPQQPYPQPQPF, known to be elicit an innate immune response[27] was not detected in our digested extracts. Probably it was not generated with the digestion method adopted, it is likely that the use different gastrointestinal proteases combinations (with pancreatic exoproteases for example) could lead to the generation of this peptide.

Pathogenic peptides deriving from α-gliadin were strongly predominant compared to the few peptides deriving from γ-gliadin and none from ω-gliadins. This could be due to a combination of factors, such as the higher sequence variability of γ- and ω-gliadins, which might lead to the formation of many different peptides, each present in an amount lower to the detection limit, and the higher abundance of α-gliadins respect to γ- and ω-gliadin.

Since in vitro digestion method adopted required the use of only three gastrointestinal proteases (pepsin, chymotrypsin and trypsin), the generated peptides can be further proteolyzed in an in vitro system, where other enzymes are present, such as pancreatic exoproteases and brush border membrane enzymes. Anyway, the method used is suitable to quantify the pathogenic sequences present in samples gluten, allowing to obtain a hierarchy of varieties with a high content of peptides involved in celiac disease.

All these peptides were quantified with the internal standard method using an isotopic labeled peptide (LQLQPF(d_5)PQPQLPY). The characteristic ions of every peptide and of the internal standard were extracted from the Total Ion Chromatogram and the peaks were integrated using the MassLynx software. A representative chromatogram with the MS and MS/MS mass spectra is reported in Figure 5.1.

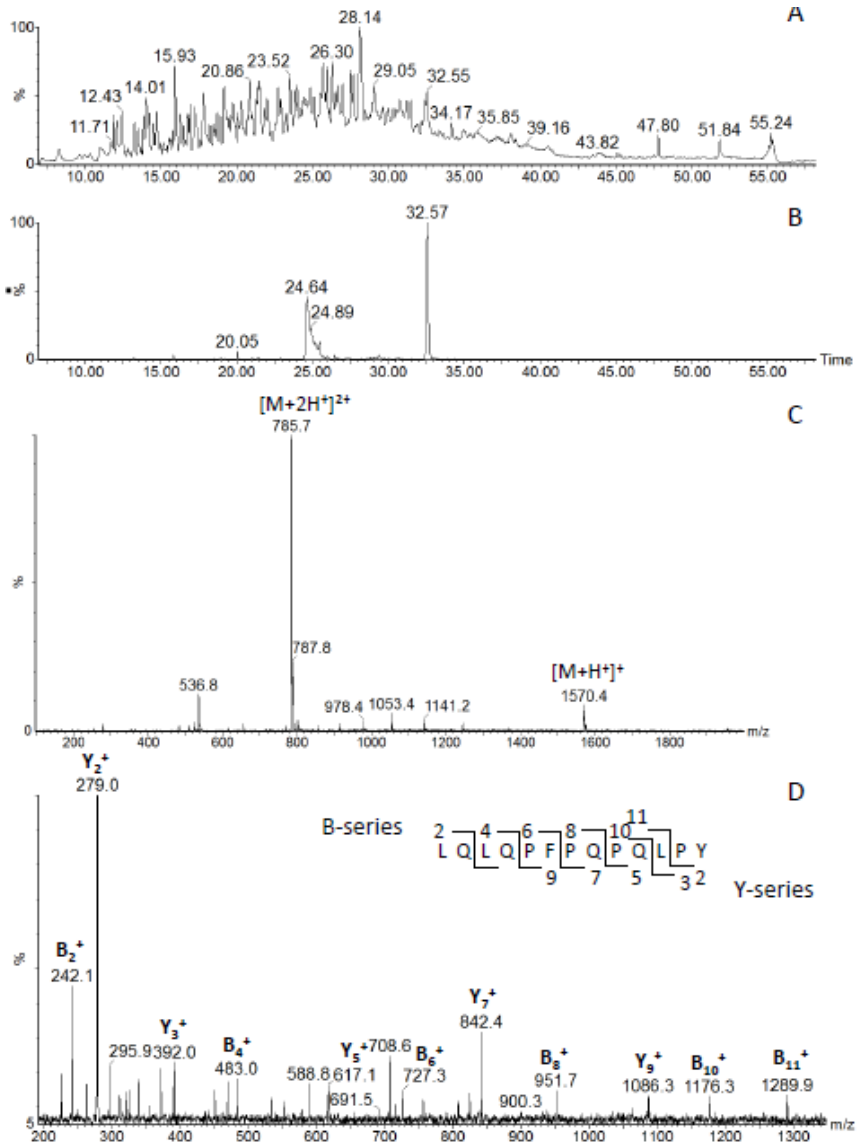

Figure 5.1 Total Ion Chromatogram of a wheat ethanolic extract underwent in vitro digestion (A). Extract Ion Chromatogram of the peptide LQLQPFPQPQLPY (m/z 785.7 for the doubly charged ion at 32.57 min, B). In C is reported the MS spectra of the peptide, while in D is shown the tandem mass spectra obtained from collision induced dissociation. The fragmentation is annotated with the standard B and Y fragments usually employed for peptides identification.

The quantification value was obtained as the ratio peptide area/internal standard area multiplied by the molar amount of internal standard, approximating a response factor of 1 for all peptides quantified. In the first set of samples analyzed (Table 5.1) the total content of immunogenic peptides ranged from a minimum of 84 µg/g for line F255 to a maximum of 548 µg/g for line F371 both cultivated in Lucera (FG, Southern Italy); the content of peptides containing short toxic sequences was approximately one fold higher, ranging from 1115 µg/g for the line F371 cultivated in Lucera (FG) to 3898 µg/g of variety Aureo cultivated in Argelato (BO, Northern Italy).

The total content of peptides elicitating adaptive and innate immune response, mediated for variety and cultivation area is reported in Figure 5.2.

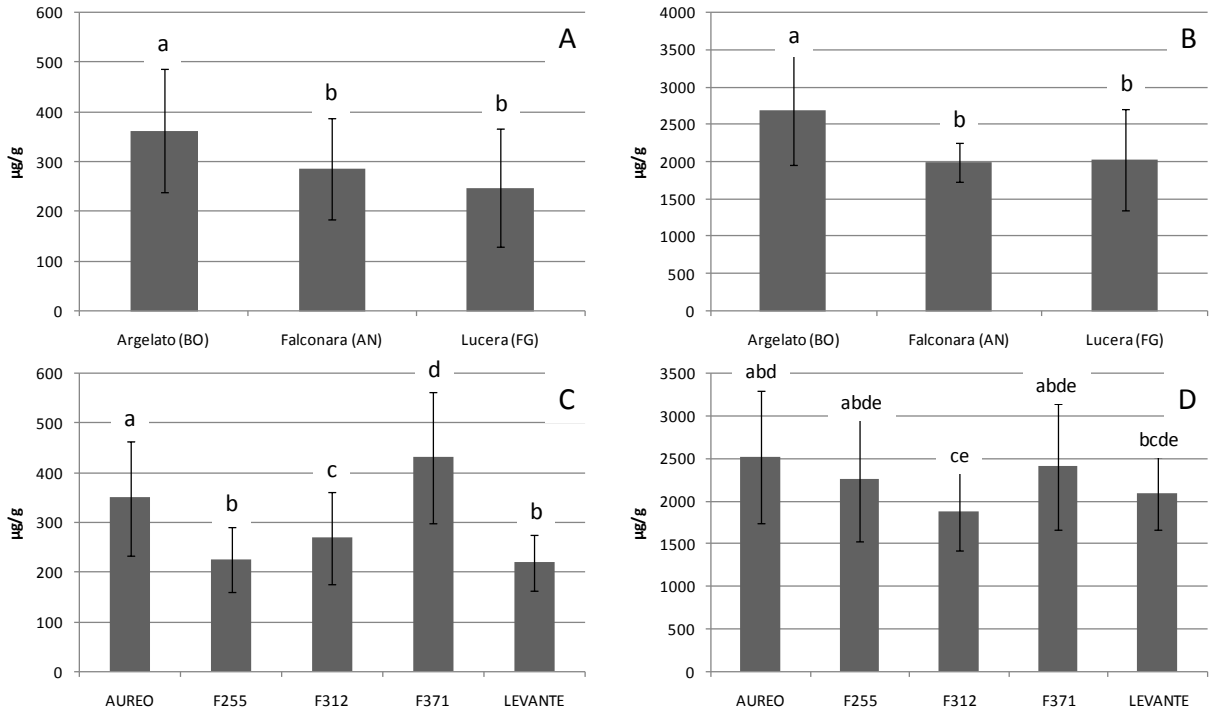

Figure 5.2 Content of immunogenic (A) and toxic (B) peptides (expressed in µg of peptide for gram of sample) mediated for cultivation area. Content of immunogenic (C) and toxic (D) peptides (expressed in µg of peptide for gram of sample) mediated for variety. Immunogenic peptides are defined as containing sequences involved in adaptive immune response in celiac disease, while toxic peptides are those elicitating innate immune response. Different letters on the bars correspond to statistically different values (two-way ANOVA); differences are considered significant with $p<0.05$ (test post hoc LSD).

Even if there is a wide intra-area and intra-line variability in the content of pathogenic peptides generated, significant differences were outlined both due to genotypes and environments. The highest variability among samples regards the content of immunogenic peptides, where each line differs from the others with the exception of F255 and Levante ($p<0.05$). These differences are less evident for peptides involved in innate immune response, where only line F371 seems to have a substantial lower content of these peptides elicitating innate immune response. The lower difference observed for peptides involved in innate immune response could be attributed to the high level of aminoacidic conservation in the N-terminal region of α- and γ-gliadins, from which the peptides identified derive. On the other hand, the gliadin region that generate the immunogenic peptides (approximately from 56th to 68th aminoacid residue) showed higher sequence variability so the genotypic effect can play an important role. The influence of the growing area seems less evident, in fact only samples from Argelato (BO) show a higher pathogenic peptides content. So this first sets of samples essentially confirmed the results already obtained recently, showing a high variability among the wheat samples with genetics as a main determinant. Anyway, this first set of samples carries a very low genetic diversity. Actually, wheat breeding programs have been carried out for many years crossing “elite x elite” germplasm with the final aim to obtain desired technological properties, resulting in a decrease of biodiversity. In order to deeply investigate the possibility to extend the variability in the production of peptides containing sequences involved in celiac disease during gastrointestinal digestion, a second set of 25 wheat accessions from a durum wheat panel were analyzed (Table 5.2). The innate response involved peptides content was quite variable among samples, ranging from 977

µg/g of cultivar Marouane to 2657 µg/g of cultivar Omsnima 1, but the range of variability of immunogenic peptides was even wider, ranging from 139 µg/g of cultivar Messapia to 776 µg/g of cultivar AW12/BIT. As shown in Table 5.2, samples could be divided into five groups on the basis of their genetic similarity. Concerning peptides involved in innate response anyway, the results of these analyses showed that the variability intra- and inter-group was quite limited, with the cluster of Icarda accessions, selected for temperate environment, showing the lowest content (Figure 5.3A and Figure 5.3C).

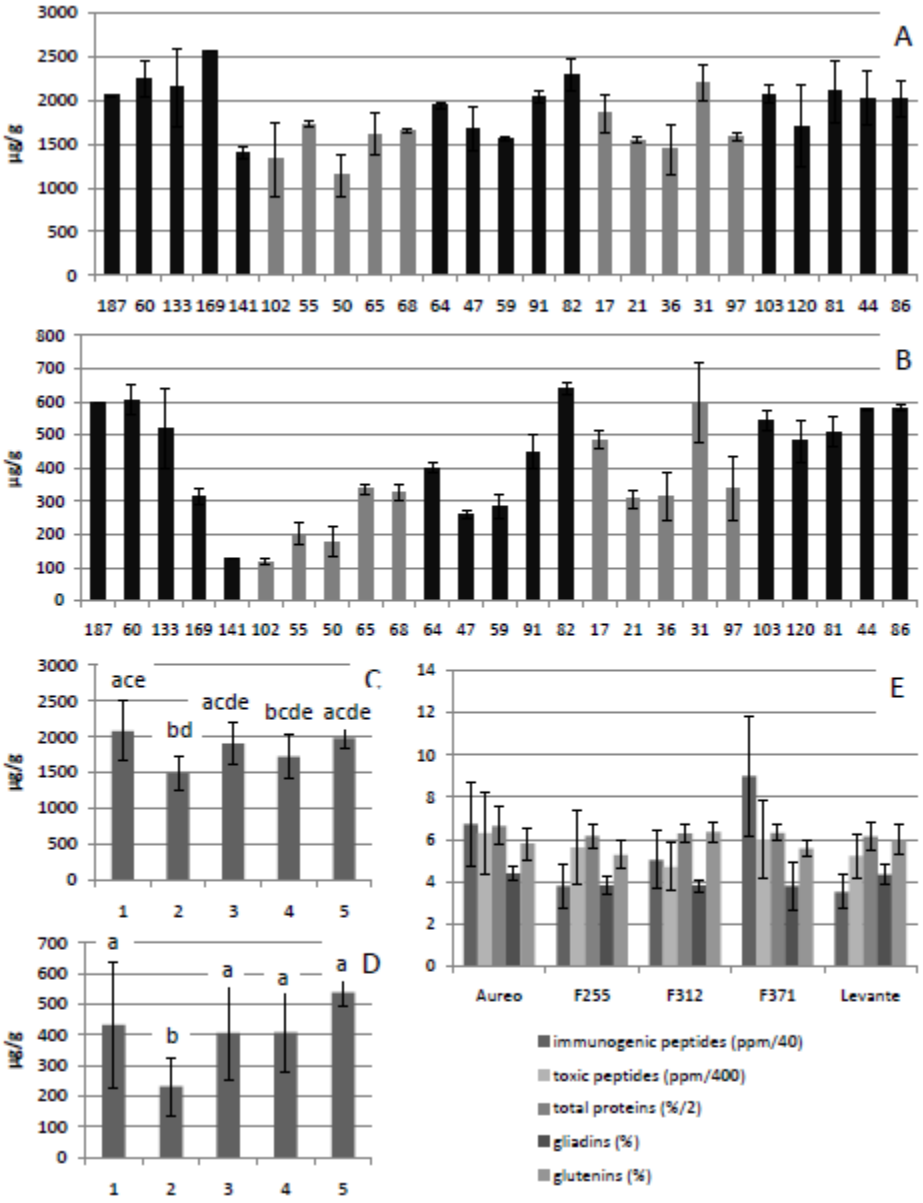

Figure 5.3 Total content of toxic (A) and immunogenic (B) peptides (expressed in µg of peptide for gram of sample) in samples from the Durum Panel. The different black and grey shadows correspond to the five groups reported in table 2. Total content of toxic (C) and immunogenic (D) peptides(expressed in µg of peptide for gram of sample) in samples from the Durum Panel, mediated for the five groups of genetic affinity. Different letters on the bars correspond to statistically different values (one-way ANOVA); differences are considered significant with $p<0.05$ (test post hoc LSD). Immunogenic peptides are defined as containing sequences involved in adaptive immune response in celiac disease, while toxic peptides are those elicitating innate immune response. In E is shown the content of immunogenic peptides (expressed as µg/g), toxic peptides (µg/g), total proteins (%), gliadins (%) and glutenins (%) of samples analyzed mediated for and variety. Immunogenic peptides are defined as containing sequences involved in adaptive immune response in celiac disease, while toxic peptides are those elicitating innate immune response. Values are divided for an appropriate scale factor in order to have the same order of magnitude.

However, the mean content was comparable with those found in the first set of inbred lines (Table 5.1), so, even if genetic diversity was higher, the variability in peptides eliciting innate immune response did not seem to increase.

On the other side, several accessions had a notably higher variability for immunogenic peptides with respect to the wheat inbred lines analyzed previously (Figure 5.3B and Figure 5.3D). In fact, even considering the high variability within each genetic cluster, group 2 has a significantly lower content of epitope containing peptides ($p<0.05$). So, in this case, the higher genetic diversity brings about also more variability in the expression of immunogenic peptides. This suggests the possibility of a breeding program aimed to decrease immunogenic peptides content in durum wheat.

5.3.3 Correlation with gluten and TPC

The content of peptides responsible for celiac disease was plotted against TPC, gliadins and glutenins content. A weak correlation was found between peptides involved in innate immune response and gliadins (Pearson coefficient 0.319, $p<0.01$) but no correlation was observed for glutenins; in fact the most abundant peptides (involved in innate response) identified belonged to gliadins, and only a small amount to glutenins. The correlation was higher for TPC, with a Pearson coefficient of 0.521 ($p<0,01$). For immunogenic peptides (involved in adaptive immune response), a small but significant correlation was observed only for TPC (Pearson coefficient 0.373, $p<0,01$) and not for gluten proteins.

These results indicate that a higher protein content could promote a higher presence of sequences triggering celiac disease. Anyway, as shown in Figure 5.3E the variability of peptides related to celiac disease is much more higher than the variability of protein distribution, meaning that the same protein or gluten content between two varieties will not necessarily result in the same generation of peptides involved in celiac immune response. This information could be useful for wheat breeders, in order to develop new varieties with good rheological properties (adequate amount of gluten and gliadin/glutenin ratio) but limiting the production of peptides implicated in celiac disease.

5.4 CONCLUSIONS

The production of peptides implicated in celiac disease during simulated gastrointestinal digestion of prolamine extract was assessed in different commercial and pre-commercial durum wheat varieties. A consistent variability in the content of immunogenic peptides was found, while for innate response involved peptides the differences were less evident. The growing area seems to have a reduced impact with respect to genetic factors. The great genotypic effect was confirmed by the analysis of elite germplasm: parallel to a higher genetic variability there is an increase of differences in peptides production upon digestion. Interestingly, the TPC and gluten proteins shows a reduced variability compared to peptides production, leading to a possible selection of wheat lines with good protein and gluten content but limiting the increase of pathogenic peptides production, paving the way to the use of wheat varieties less stressful for the immune system and thus useful for the prevention of celiac disease.

The results here reported confirmed that genetics plays the major role in determining the pathogenicity of different wheat varieties in celiac disease. The loose relationship between crude protein content and gluten

content with the amount of pathogenic peptides in wheat implies that less pathogenic varieties might potentially be developed at the same time preserving the same amount of gluten proteins. Less pathogenic varieties, albeit still impossible to be consumed by people already having celiac disease, on the other side could be useful for the prevention of the pathology in people genetically predisposed, but still healthy.

5.5 ACKNOWLEDGEMENTS

The AGER Foundation is gratefully acknowledged for its financial support to the project “From seed to pasta.”

---

[1] Leff B, Ramankutty N, Foley JA (2004)Geographic distribution of major crops across the world. *Global Biogeochem Cy* doi:10.1029/2003GB002108

[2] Peña RJ (2002) Wheat for bread and other foods. http://www.fao.org/docrep/006/y4011e/y4011e0w.htm

[3] Atchison J, Head L, Gates A (2010) Wheat as food, wheat as industrial substance; comparative geographies of transformation and mobility. *Geoforum* 2, 236-246

[4] Van den Broeck HC, America AHP, Smulders MJM, Bosch D, Hamer RJ, Gilissen LJWJ, Van der Meer IM (2009) A modified extraction protocol enables detection and quantification of celiac disease-related gluten proteins from wheat. *J Chromatogr B* 877, 975-982

[5] Uthayakumaran S, Gras PW, Stoddard FL, Bekes F (1999) Effect of varying protein content and glutenin-to-gliadin ratio on the functional properties of wheat dough. *Cereal Chem* 76, 389-394

[6] Wieser H (2007) Chemistry of gluten proteins. *Food Microbiol* 24, 115-119

[7] Šramková Z, Gregová E, Šturdík E (2009) Genetic improvement of wheat- a review. *Nova Biotechnologica* 9, 27-51

[8] Schuppan D (2000) Current Concepts of Celiac Disease Pathogenesis. *Gastroenterology* 119, 234-242

[9] Gianfrani C, Auricchio S, Troncone R (2005) Adaptive and innate immune responses in celiac disease. *Immunol Lett* 99, 141-145.

[10] Hausch F, Shan L, Santiago NA, Gray GM, Khosla C (2002) Intestinal digestive resistance of immunodominant gliadin peptides. *Am J Physiol-Gastr L* 283, 996-1003

[11] Rostom A (2004) Incidence and prevalence of celiac disease. NIH consensus development conference on celiac disease.

[12] Catassi C, Kryszak D, Bhatti B, Sturgeon C, Helzlsouer K, Clipp SL, Gelfond D, Puppa E, Sferruzza A, Fasano A (2010) Natural history of celiac disease autoimmunity in a USA cohort followed since 1974. *Ann Med* 42, 530–538

[13] Gobbetti M, Rizzello GM, Di Cagno R, De Angelis M (2007) Sourdough lactobacilli and celiac disease. *Food Microbiol* 24, 187–196

[14] Ivarsson A, Persson LA, Nyström L, Ascher H, Cavell B, Danielsson L, Dannaeus A, Lindberg T, Lindquist B, Stenhammar L, Hernell O (2000) Epidemic of coeliac disease in Swedish childrens. *Acta Paediatr* 89, 165–171

[15] Van den Broeck HC, de Jong HC, Salentijn EMJ, Dekking L, Bosch D, Hamer RJ, Gilissen LJWJ, Van der Meer IM, Smulders MJM (2010) Presence of celiac disease epitopes in modern and old hexaploid wheat varieties: wheat breeding may have contributed to increased prevalence of celiac disease. *Theor Applied Genet* 121, 1527–1539

[16] Molberg O, Uhlen AK, Jensen T, Flaete NS, Fleckenstein B, Arentz-Hansen H, Raki M, Lundin KE, Sollid LM (2005) Mapping of gluten T-cell epitopes in the bread wheat ancestors: implications for celiac disease. *Gastroenterology* 128, 393-401

[17] Van Herpen TWJM, Goryunova SV, Van der Schoot J, Mitreva M, Salentijn E, Vorst O, Schenk MF, Van Veelen PA, Koning F, Van Soest LJM, Vosman B, Bosch D, Hamer RJ, Gilissen LJWJ, Smulders MJM (2006) Alpha-gliadin genes from the A, B, and D genomes of wheat containdifferent sets of celiac disease epitopes. *BMC Genomics*, doi:10.1186/1471-2164-7-1

[18] Spaenij–Dekking L, Kooy–Winkelaar Y, Van Veelen P, Drijfhout JW, Jonker H, Van Soest H, Smulders MJM, Bosch D, Gilissen LJWJ, Koning F (2005) Natural variation in toxicity of wheat: potential for selection of nontoxic varieties for celiac disease patients. *Gastroenterology* 129, 797–806

[19] Prandi B, Bencivenni M, Faccini A, Tedeschi T, Dossena A, Marchelli R, Galaverna G, Sforza S (2012) Composition of peptide mixtures derived from simulated gastrointestinal digestion of prolamins from different wheat varieties. *J Cer Sci* 56, 223-231

[20] Maccaferri M, Sanguineti MC, Noli E, Tuberosa R (2005) Population structure and long-range linkage disequilibrium in a durum wheat elite collection. *Mol Breeding* 15, 271-290

[21] Lookhart G, Bean S (1995) Separation and characterization of wheat protein fractions by High-Performance Capillary Electrophoresis. *Cereal Chem* 72, 527-532

[22] Wieser H, Antes S, Seilmeier W (1998) Quantitative determination of gluten protein types in wheat flour by reversed-phase high-performance liquid chromatography. *Cereal Chem* 75, 644-650

[23] Sollid LM (2000) Molecular basis of celiac disease. *Annu Rev Immunol* 18, 53–81

[24] De Vita P, Ficco DBM, Luciani A, Vicentini O, Pettoello-Mantovani M, Silano M, Maiuri L, Cattivelli L (2012) A ω-secalin contained decamer shows a celiac disease prevention activity. *J Cer Sci* 55, 234-242

[25] Arentz-Hansen H, Körner R, Molberg QH, Vader W, Kooy YM, Lundin KEA, Koning F, Roepstorff P, Sollid LM, McAdam SN (2000) The intestinal T cell response to α-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *J Exp Med* 191, 603-612

[26] Cornell HJ (1996) Coeliac disease: a review of the causative agents and their possibly mechanisms of action. *Amino Acids* 10, 1-19

[27] Picarelli A, Di Tola M, Sabbatella L, Anania MC, Di Cello T, Greco R, Silano M, De Vincenzi M (1999) 31-43 amino acid sequence of the alpha-gliadin induces anti-endomysial antibody production during in vitro challenge. *Scand J Gastroentero* 34, 1099-1102

6 LC/MS ANALYSIS OF PROTEOLYTIC PEPTIDES IN WHEAT EXTRACTS FOR DETERMINING THE CONTENT OF THE ALLERGEN AMYLASE/TRYPSIN INHIBITOR CM3: INFLUENCE OF GROWING AREA AND VARIETY

Based on: Prandi B., Faccini A., Tedeschi T., Galaverna G., Sforza S. (2013) LC/MS analysis of proteolytic peptides in wheat extracts for determining the content of the allergen amylase/trypsin inhibitor CM3: Influence of growing area and variety. Food Chemistry 140(1-2):141-6.

6.1 INTRODUCTION

Food allergy are defined as “an adverse health effect arising from a specific immune response that occurs reproducibly on exposure to a given food”[1] and affect about 6-8% of children[2] and 1-2% of the general population.[3] Wheat is one of the three most harvested cereals worldwide[4] and the rate of sensitization is 0.4-1.3% in children and 0.2-0.9% among adults.[5, 6] Food allergies to wheat are related to a wider spectrum of proteins, ranging from ω_5-gliadins (Battais et al., 2003), to α-amylase/trypsin inhibitors and non specific lipid transfer proteins.[7] If the sensitization occurs through the gastrointestinal tract, the most frequent adverse reaction is wheat dependent exercise induced anaphylaxis. Instead, if the sensitization occurs through the respiratory tract, it leads to a frequent occupational disease, the baker’s asthma.[8] While ω_5-gliadins are principally involved in wheat dependent exercise induced anaphylaxis,[9] the main trigger of baker’s asthma seems to be a group of proteins called α-amylase/trypsin inhibitors.[10]

The α-amylase/trypsin inhibitors belong to the prolamins superfamily:[11] they are cysteine-rich proteins and have a tridimensional structure rich in α-helices that confers them a certain stability to proteolysis and thermal processing.[12] They are involved in the plant defense from insects and microorganisms: they can complex exogenous proteases such as α-amylases and trypsin,[13] without inhibiting the activity of the endogenous enzymes of the germinating kernel. They are soluble in dilute salt solutions, but also in chloroform/methanol mixtures, thus being often called CM proteins.[14]

CM3 is a recognized allergen involved in baker’s asthma (Tri a 30): it has been demonstrated that it can bind IgE in sera of patients with atopic dermatitis[15] or with baker’s asthma[16], food allergy to wheat[17] and gave positive results in skin prick tests in patients with baker’s asthma.[18] In addition to this, it has been recently discovered that members of the non-gluten alpha-amylase/trypsin inhibitor family are potent activators of various innate immune cells such as dendritic cells and macrophages through activation of the Toll-Like Receptor 4. These findings defined cereal alpha-amylase/trypsin inhibitors as novel contributors to celiac disease.[19] Anyway, the actual content of this protein in different wheat varieties has never been studied in detail: this information might be essential in order to assess if the different varieties can bring about different levels of risk for allergic subjects and celiac patients.

In this work the α-amylase/trypsin inhibitor Tri a 30 in durum wheat samples was identified through in-gel digestion and LC/MS analysis of the tryptic digests. For the quantification of this protein, two marker peptides generated from enzymatic cleavage of the salt soluble extract were identified by L/ESI-MS/MS: one of these peptides was synthesized in the isotopically labeled form and used as internal standard for the quantification of

its analogue in the digested extracts by UPLC/ESI-MS, in turn allowing for protein quantification. The protein content in samples belonging to different varieties and coming from different cultivation areas was determined.

6.2 MATERIALS AND METHODS

6.2.1 Samples

Wheat samples (18, Table 6.1) were obtained from Società Produttori Sementi SpA (Bologna, Italy).

Table 6.1 List of the analysed *Triticum turgidum* spp *durum* samples and of their total proteins and salt soluble proteins content.

Sample	Variety	Area of production	TPC %, Kjeldhal	Globulin content %, fluorimetric	CM 3 content mg/g
1	D240	Argelato (BO)	10.34±1.30	2.72±0.05	1.07±0.01
6	D240	Argelato (BO)	9.66±0.11	2.72±0.04	1.10±0.09
9	D240	Poggio Renatico (FE)	11.73±1.02	2.90±0.05	0.78±0.04
12	D240	Poggio Renatico (FE)	11.54±1.19	2.91±0.03	0.78±0.05
13	D240	Lucera (FG)	13.57±0.45	2.93±0.03	0.35±0.02
18	D240	Lucera (FG)	14.85±0.23	3.02±0.03	0.43±0.01
3	Levante	Argelato (BO)	11.61±0.28	2.77±0.03	0.74±0.05
5	Levante	Argelato (BO)	10.22±0.56	2.69±0.05	0.75±0.003
8	Levante	Poggio Renatico (FE)	13.45±0.51	2.80±0.03	0.62±0.04
11	Levante	Poggio Renatico (FE)	12.46±0.45	2.99±0.02	0.50±0.03
15	Levante	Lucera (FG)	15.69±0.17	3.16±0.03	0.34±0.02
16	Levante	Lucera (FG)	17.65±0.45	3.41±0.06	0.22±0.02
2	Svevo	Argelato (BO)	13.65±1.69	3.00±0.03	0.68±0.04
4	Svevo	Argelato (BO)	12.38±0.23	2.68±0.03	0.96±0.03
7	Svevo	Poggio Renatico (FE)	14.21±0.45	2.78±0.03	0.53±0.04
10	Svevo	Poggio Renatico (FE)	14.57±0.40	3.12±0.04	0.59±0.06
14	Svevo	Lucera (FG)	15.21±0.40	3.07±0.05	0.45±0.01
17	Svevo	Lucera (FG)	18.88±0.51	3.19±0.03	0.36±0.01

All the samples belong to the specie *Triticum turgidum* spp *durum* and the ancestors of origin are respectively: Syndiouk/Mahmoudi//Langdon 341 for D240, G80/Piceno//Ionio for Levante and Linea Cimmyt/Zenit for Svevo.

6.2.2 Reagent and solvents

Sodium chloride was purchased from AnalaR Normapur (Milan, Italy). Sulphuric and phosphosulphuric acids, copper (I) oxide, hydrochloric acid (37%) and glacial acetic acid, dimethylformamide, methyl red and bromocresol green were purchased from Carlo Erba (Milan, Italy), Kjeltabs ST and Antifoams Tablets were purchased from Thompson & Capper (Runcorn, UK). Boric acid, ammonium bicarbonate, Fmoc-arginine(pbf)-Wang resin and Fmoc-isoleucine-OH were purchased from Fluka (Buchs, Switzerland). XT sample buffer (4×), XT reducing agent (20×), XT MES running buffer (20×), SDS-PAGE standards broad range, Criterion XT Precast Gel (12% Bis-Tris) and Coomassie Brilliant Blue R-250 were purchased from Biorad (Hercules, CA, USA). Acetonitrile, methanol, trypsin from porcine pancreas, α-chymotripsin from bovine pancreas, pepsin (1:10000) from porcine gastric mucosa, sodium phosphate monobasic, iodoacetamide, DL-dithiothreitol, Fmoc-serine(OtBu)-OH, piperidine,

diisopropylethylamine, dichloromethane, triisopropylsilane and diethyl ether were purchased from Sigma-Aldrich (St. Louis, MO, USA). Formic acid and trifluoroacetic acid were purchased from Acros Organics (Geel, Belgium). Fmoc-phenylalanine-OH, Fmoc-leucine-OH, Fmoc-proline-OH and Fmoc-aspartic acid(OtBu)-OH were purchased from Novabiochem (Darmstadt, Germany). Fmoc-alanine(3,3,3-d_3)-OH was purchased from Isotech (Miamisburg, OH, USA). Fmoc-valine-OH was purchased from Advanced Biotech Italia (Monza Brianza, Italy). Fmoc-glutamine(Trt)-OH was purchased from Bachem (Bubendorf, Switzerland).

6.2.3 Salt soluble extracts

500 mg of ground wheat kernels were extracted with 10 ml of a 0.5 M NaCl aqueous solution for 2 hours and 30 minutes at room temperature. The mixture centrifuged at 2486 g for 15 min at 10°C and the supernatant stocked at -20°C.

6.2.4 Protein quantification

In order to assess the protein content of the salt soluble extract, a Q-bit fluorometer (Invitrogen, Grand Island, NY, USA) was used. The working solution was prepared by mixing 1/200 of the Quant-it reagent with 199/200 of Quant-it protein buffer; then, 2 µl of each salt soluble extract was mixed with 198 µl of working solution. The sample was left for 15 min in the dark before reading the protein content with the fluorometer. The total protein content was determined according to the Kjeldhal method.

6.2.5 Preparative SDS-PAGE of the salt soluble extract

800 µl of a salt soluble extract were dried under nitrogen flux and reconstituted with 500 µl of reducing sample buffer (pH 6,8; 1% SDS; 12,5% glycerol; 0,005% bromophenol blue; 2,5% 2-mercaptoethanol); after denaturation at 95°C for 5 min, the samples were loaded into a Criterion bis-tris precast gels 12%. The running buffer was 25 mM Tris, 192 mM glycine and 0.1% SDS and the potential applied was 150V. The gel was finally stained with Coomassie Brilliant Blue.

6.2.6 In gel digestion

The band around 14400 Da was excised and chopped into small pieces. The gel was destained and the reduction, alkylation and digestion of the band was carried out as described by the in-gel digestion protocol of the Biological Mass Spectrometry Laboratory (Ontario Wide Protein Identification Facility, http://www.biochem.uwo.ca/wits/bmsl/in-gel_digestion.html), adapting the volumes to the preparative scale. The solution containing the peptides extracted from the gel was filtered through a syringe-driven filter units with a pore size of 0.2 µm (Millipore, Billerica, MA, USA). The salts were removed from the solution using ZipTip pipette tips (Millipore, Billerica, MA, USA) according to the instruction of the manufacturer.

6.2.7 LTQ-OrbiTrap analysis

In order to confirm the identity of the band excised from the gel as CM3 protein, the digested samples were analysed by HPLC-LTQ-ORBITRAP (Thermo scientific, Waltham, MA, USA) using a C18 column and a gradient elution; eluent A was water with 0.1% acetonitrile and 0.1% formic acid and eluent B was acetonitrile with 0.1% formic acid (gradient: 0-4 min from 100% A to 95% A, 4-60 min from 95% A to 50% A, 60-62 min from 50% A to 10% A, 62-72 min 10% A, 72-74 min from 10% A to 95% A, 74-90 min 95% A). The analysis parameters were:

flow 5 µl/min; analysis time 90 min; column temperature 30°C; sample temperature: 10°C; injection volume 5 µl; acquisition time 0-75 min; ionization type positive ions; scan range 200-1800 m/z; source voltage 3.5 kV; capillary voltage 35 V; source temperature 275°C.

6.2.8 Enzymatic digestion of the extracts

1 ml of each extract was dried under nitrogen flux and reconstituted with 500 µl of HCl 10 mM (pH 2); 20 µl of a 1 mg/ml pepsin solution were added and the mixture was incubated for 3 hours at 37°C. The pH was set to 7.2 with 300 µl of NaH_2PO4 100 mM and the samples were added with 20 µl of a 1 mg/ml chymotrypsin solution and 20 µl of a 1 mg/ml trypsin solution before being incubated for 4 hours at 37°C. The samples were dried under nitrogen flux and reconstituted with 300 µl of a 0.1% formic acid solution, centrifuged at 16602 g for 10 min (4°C) and the supernatant was saved for LC/MS analysis. For the quantification of the marker peptides, 250 µl of supernatant were spiked with 5 µl of internal standard FIA(d_3)LPVPSQPVDPR solution (0.214 mM).

6.2.9 HPLC/ESI-MS/MS analysis

In order to confirm the aminoacidic sequence of the marker peptides, the samples were separated by a RP column (JUPITER 5 µm C18 300 Å 250*2 mm) in an HPLC/ESI-MS/MS (HPLC Waters Alliance 2695 with a triple quadrupole mass spectrometer Waters 4 Micro), using a gradient elution. Eluent A was water with 0.1% formic acid and 0.2% acetonitrile, eluent B was acetonitrile with 0.1% formic acid; gradient: 0-12 min 100% A, 12-77 min from 100% A to 50% A, 77-81 min 50% A, 81-82 min from 50% A to 0% A, 82-90 min 0% A, 90-91 min from 0% A to 100% A, 91-110 min 100% A. The samples were first analyzed in Full Scan mode, to identify the characteristic ions and the retention time of the marker peptides, and then in Daughters Scan modality using a collision energy of 25 eV. HPLC/ESI-MS/MS parameters were: flow 0.2 ml/min; run time 110 min; column temperature 35°C; injection volume 40 µl; acquisition time 7-90 min; ionization type positive ions; scan range 100-2000 m/z; capillary voltage 3.2 kV; cone voltage 35 V; source temperature 100°C; desolvation temperature 150°C; cone gas flow 100 l/h; desolvation gas flow 650 l/h. The peptide sequences were assigned on the basis of the mass spectra obtained.

6.2.10 Synthesis of the internal standard

The peptide FIA(d_3)LPVPSQPVDPR (isotopically labeled on the alanine residue) was synthesized on solid phase according Fmoc/t-butyl strategy using a Syro I Fully Automated Peptide Synthesizer (Biotage, Uppsala, Sweden). The peptide was cleaved from the resin using a TFA:TIS:H_2O (95:2.5:2.5) solution, precipitated with diethyl ether and purified using a semipreparative RP-HPLC-UV (λ=254 nm). The purified product was quantified via spectrophotometric method at 257 nm using an external calibration curve made with a phenylalanine solution at different dilutions.

6.2.11 UPLC/ESI-MS analysis

In order to quantify the marker peptides for the α-amylase/trypsin inhibitor CM3, the samples were separated by a RP column (ACQUITY UPLC BEH 300 C18 1.7 µm 2.1*150 mm) in an UPLC/ESI-MS system (UPLC Acquity Waters with a single quadrupole mass spectrometer Waters Acquity Ultraperformance) using a gradient elution. Eluent A was water with 0.1% formic acid and 0.2% acetonitrile, eluent B was acetonitrile with 0.1% formic acid; gradient: 0-7 min 100% A, 7-50 min from 100% A to 50% A, 50-52,6 min 50% A, 52,6-53 min from 50% A to 0%

A, 53-58,2 min 0% A, 58,2-59 min from 0% A to 100% A, 59-72 min 100% A. The digested sample extracts were analysed with UPLC/ESI-MS in the Full Scan mode (flow 0.2 ml/min; analysis time 72 min; column temperature 35°C; sample temperature 18°C; injection volume 5 µl; acquisition time 7-58,2 min; ionization type positive ions; scan range 100-2000 m/z; capillary voltage 3.2 kV; cone voltage 30 V; source temperature 100°C; desolvation temperature 200°C; cone gas flow 100 l/h; desolvation gas flow 650 l/h). The characteristic ions were extracted, obtaining eXtract Ion Chromatograms (XICs), in which the areas of the identified peptides and internal standard FIA(d_3)LPVPSQPVDPR were integrated with the MassLynx software. The quantification value was obtained as the ratio peptide area/internal standard area multiplied by the moles of internal standard.

6.3 RESULTS AND DISCUSSION

6.3.1 Identification of CM3 peptides in the salt soluble extracts

All the 18 samples analyzed belonged to *Triticum turgidum* spp *durum* species but to different varieties (D240, Levante and Svevo) and cultivation areas (Argelato, Poggio Renatico, Lucera). Two replicates were sampled for each combination variety/area, as described in Table 6.1. Salt soluble extracts of durum wheat samples were obtained as detailed in the experimental section. In order to detect the presence of the α-amylase/trypsin inhibitor CM3 in the extracts, a bottom up proteomic approach was applied. First the proteins in the salt soluble extract were separated by monodimensional electrophoresis, the band around 14000 Da, the likely CM3 protein, was excised and tryptically digested, and the tryptic mixture was analyzed by µHPLC-LTQ-OrbiTrap (Figure 6.1).

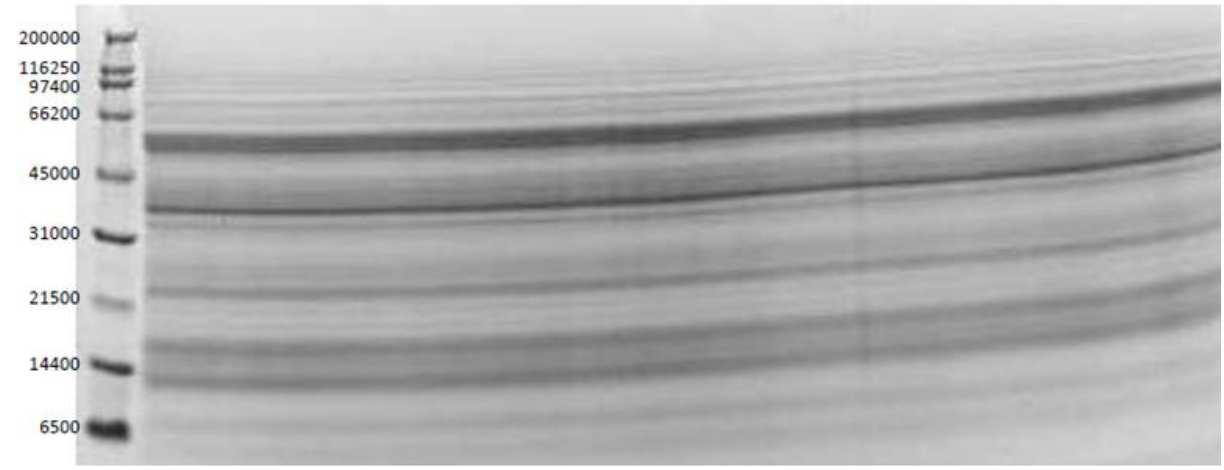

Figure 6.1 Preparative SDS-PAGE of a durum wheat salt soluble extract. The band around 14.4 kDa was digested with trypsin and analyzed with LTQ-OrbiTrap.

The peptides generated from the enzymatic cleavage were identified by the mass spectra obtained from the MS/MS fragmentation in the OrbiTrap system, and the corresponding proteins were identified by comparing the proteolytic peptides with those obtained by *in silico* digestion of the protein database. The highest score (189.14) was obtained for the α-amylase/trypsin inhibitor CM3 with a coverage of 48.21% (amino acids preceding the cleavage sites are in bold, the peptides are underlined):

SGSCVPGVAF**R**TNLLPHC**R**DYVLQQTCGTFTPGS**K**LPEWMTSASIYSPG**K**PYLA**K**LYCCQELAEISQQCRCEAL**R**Y
FIALPVPSQPVDP**R**SGNVGESGLIDLPGCP**R**EMQWDFV**R**LLVAPGQcNLATIHNV**R**YCPAVEQPLWI

6.3.2 Identification of marker peptides for CM3 quantification

After having confirmed the presence of CM3 protein in the extracts, its quantification was assessed. The direct quantification of a single protein in a complex matrix is usually difficult to obtain: high molecular weight compounds as proteins are difficult to ionize in LC/MS systems, and the quantification obtained from UV detection can be affected by co-elution of chromophoric compounds. Also non-chromatographic methods show several problems: electrophoretical bands show not high resolution and are thus affected by the co-migration of proteins with similar molecular weight, whereas the immunochemical methods are linked to the possibility of obtaining antibodies to the specific target protein. A different approach is related to the enzymatic cleavage of the protein, in order to obtain the corresponding derived peptides, easier to quantify in an LC/MS system. Thus, the enzymatic break-down of all proteins contained in the whole salt soluble extracts was achieved by using a pepsin-trypsin/chymotrypsin two step digestion, as the use of the only trypsin (though overnight) did not gave satisfactory results (data not shown). The digested mixture was analyzed by LC/ESI-MS/MS and among the peptides generated, two were identified as the sequences FIALPVPSQPVDPR and its related digestion product IALPVPSQPVDPR (region 77-90 of the CM3 protein) (Figure 6.2).

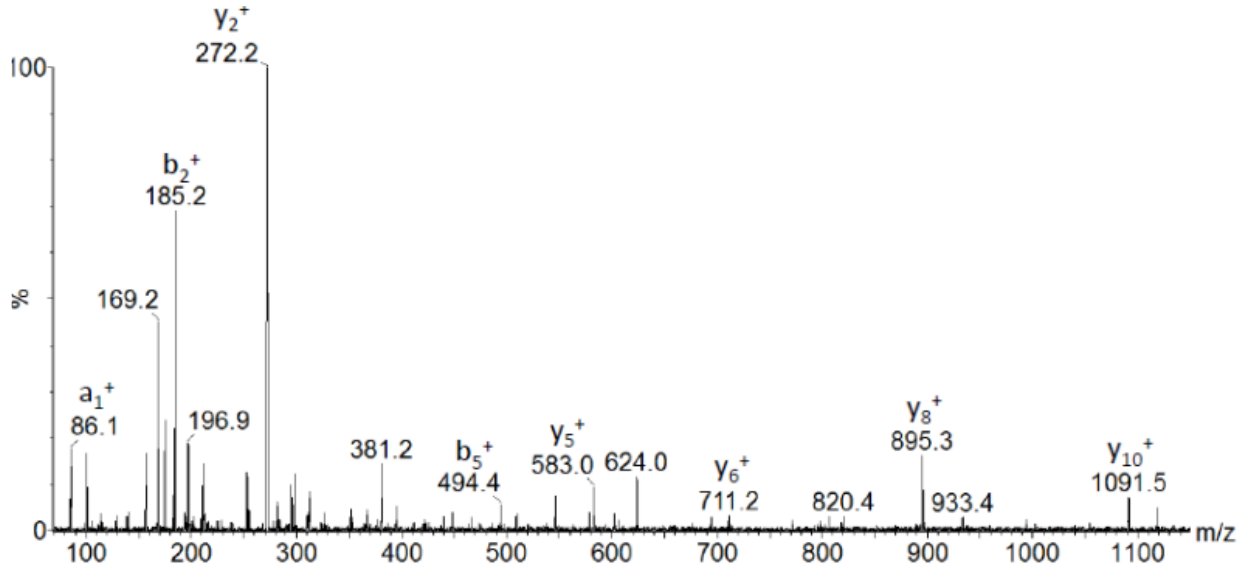

Figure 6.2 MS/MS spectra of the ion 694.9 m/z, $[M+2H^+]^{2+}$ of the peptide IALPVPSQPVDPR. Matching fragments are labeled with the corresponding ion.

In order to ensure that the break-down of the protein was quantitative, and thus there was an equimolar ratio between the marker peptides and the protein of origin, samples were taken during the digestion process at different times. The chromatographic peaks due to the peptides IALPVPSQPVDPR and FIALPVPSQPVDPR in the UPLC/ESI-MS system (Figure 6.3) were integrated with MassLynx software and plotted against the digestion time.

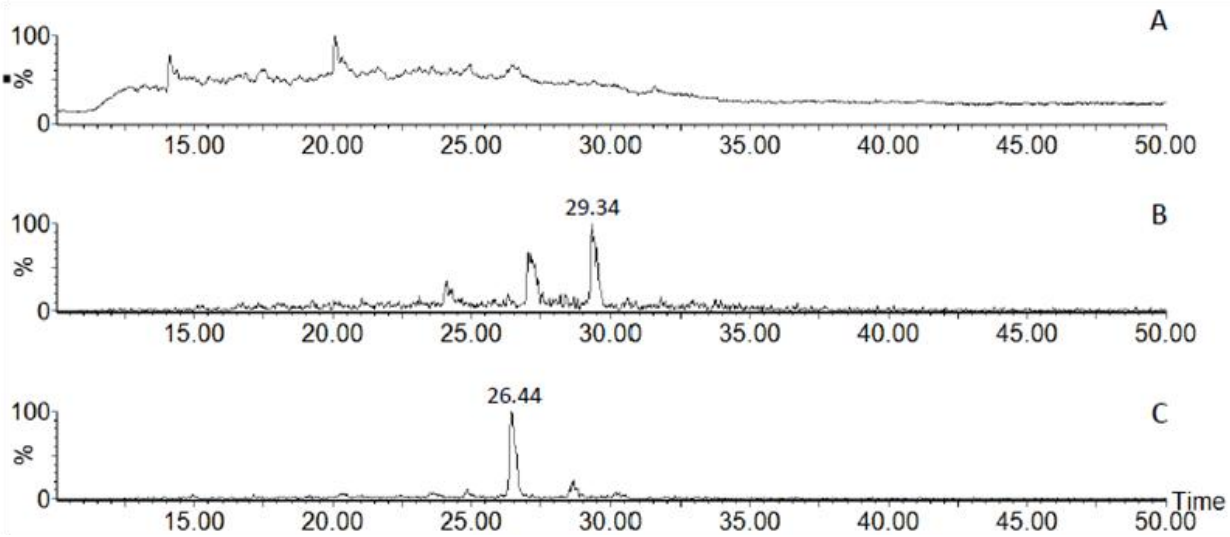

Figure 6.3 Total Ion Chromatogram of a salt soluble extract of durum wheat underwent to enzymatic cleavage with pepsin-trypsin/chymotrypsin (A), eXtract Ion Chromatogram of the ions 768.4+512.6 m/z ($[M+2H^+]^{2+}$ and $[M+3H^+]^{3+}$ of the peptide FIALPVPSQPVDPR) at Rt=26.44 min (B) and 694.9+463.6 m/z ($[M+2H^+]^{2+}$ and $[M+3H^+]^{3+}$ of the peptide IALPVPSQPVDPR) at Rt=29.34 min (C).

As shown in Figure 6.4, the enzymatic cleavage applied allows to obtain the exhaustive break-down of the protein, as indicated by the plateau phase reached at the end of the trypsin/chymotripsin digestion, suggesting that no more intact protein was anymore present.

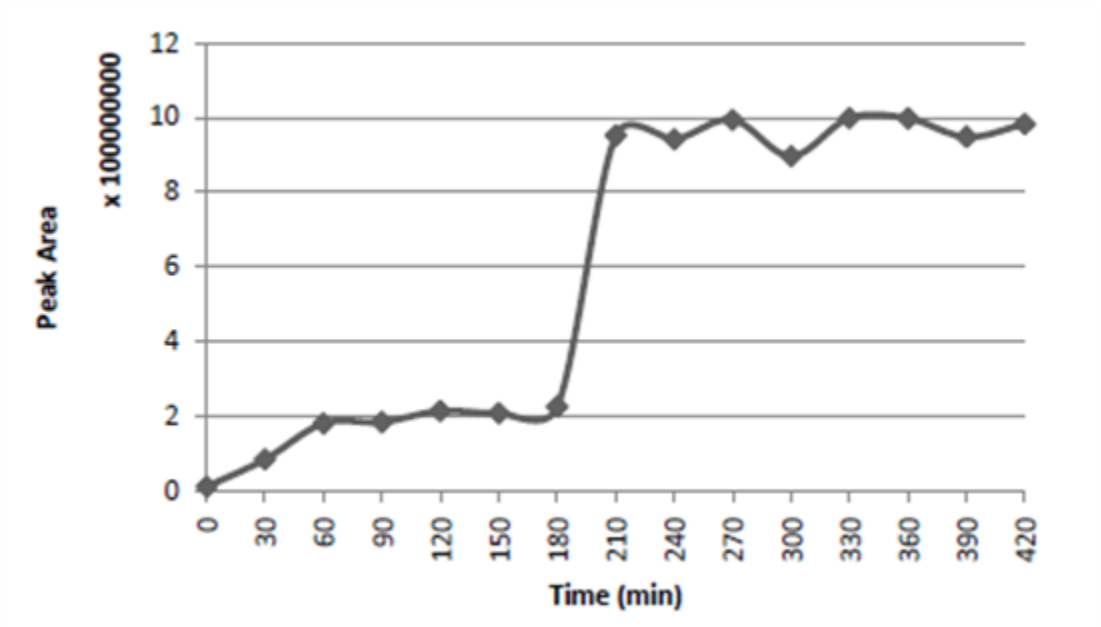

Figure 6.4 . Sum of the areas of the peptides FIALPVPSQPVDPR and IALPVPSQPVDPR, plotted against the digestion time (minutes). 0-180 min: pepsin; 180-420 min: trypsin/chymotripsin.

6.3.3 Quantification of the α-amylase/trypsin inhibitor CM3

For the exact quantification of the marker peptides an isotopically labeled peptide analogue of one of the two marker peptides, FIA(d_3)LPVPSQPVDPR was synthesized and used as internal standard, using an alanine with

three hydrogens replaced by three deuterium atoms. The analysis was carried out in a single stage UPLC/ESI-MS system, as detailed in the experimental section. After having obtained the molar amount of peptides, and considering an equimolar ratio between the CM3 protein and the marker peptides, the concentration (mg/g) of CM3 in the different durum wheat samples analyzed was obtained. Results are reported in Table 6.1, expressed as mg of CM3 protein per gram of ground whole wheat. These results showed a high variability among samples analyzed, with amounts ranging from 0.22 mg/g to 1.11 mg/g. The results, mediated for cultivation area and for variety are shown in Figure 6.5.

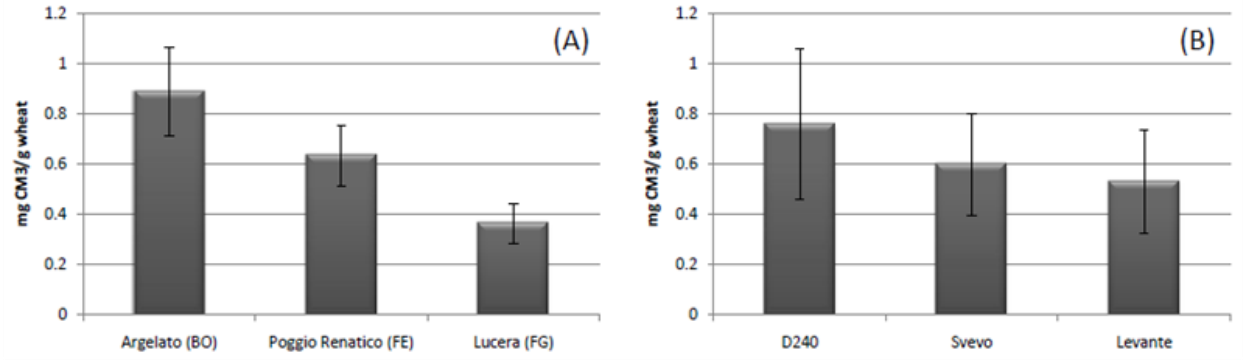

Figure 6.5 Amount of the α-amylase/trypsin inhibitor CM3, expressed as mg of protein per gram of wheat, mediated for the different cultivation areas (A) and varieties (B).

Strong differences ($p<0.01$) are present on the basis of the cultivation area (Argelato>Poggio Renatico>Lucera). This is in agreement with the defensive function of this protein *in vivo*, which thus can be influenced by environmental factors, such as insect or other pests infestations and climatic conditions. These differences are less evident, but however significant ($p<0.05$), among the varieties tested, with D240 showing the highest content, followed by Svevo and Levante; this implies a role of a genetic predisposition to the expression of the α-amylase/trypsin inhibitor CM3. The role of environmental factor (such as shadowing, altitude or storage conditions), as well as of intrinsic factor (genotype and ripening) in affecting the expression of allergen coding genes has been already demonstred for apple[20, 21] and for peach[22]. In this work we demonstrated that also for the wheat allergen CM3 the growing condition, besides the genotype, are of outmost importance in determining the allergenic potential of the product.

6.4 Conclusions

The content of the amylase/trypsin inhibitor CM3, a common wheat allergen also implied in celiac disease immunological response, has been exactly quantified for the first time in different wheat varieties, outlining large differences mostly due to the cultivation environment, and partly to the genetic asset. These results are of outmost importance for determining the environmental condition that can decrease the CM3 content, thus reducing the stimulation of innate celiac response. Moreover, given its importance as common wheat allergen, these results demonstrate that different varieties and agronomical practices (but mostly the latter) can profoundly affect its content in wheat, with direct consequences for wheat-allergic subjects and partially reducing the risks for the sensitized people/professional workers.

6.5 ACKNOWLEDGEMENTS

The AGER Foundation is gratefully acknowledged for its financial support to the project “From seed to pasta.”

---

[1] Boyce JA, Assa’ad A, Burks AW, Jones SM, Sampson HA, Jones SM, Sampson HA, Wood RA, Plaut M, Cooper SF, Fenton MJ (2010) Guidelines for the diagnosis and management of food allergy in the United States; report of the NIAID-sponsored Expert Panel. *J Allergy Clin Immun* 126, S1-S58

[2] Bock SA (1987) Prospective appraisal of complaints of adverse reactions to foods in children during the first 3 years of life. *Pediatrics* 79, 683-688

[3] Young E, Stoneham MD, Petruckevitch A, Barton J, Rona R (1994) A population study of food intolerance. *Lancet* 343, 1127-1130

[4] Leff B, Ramankutty N, Foley JA (2004) Geographic distribution of major crops across the world. *Global Biogeochem Cy* 18, GB1009

[5] Emmett SE, Angus FJ, Fry JS, Lee PN (1999) Perceived prevalence of peanut allergy in Great Britain and its association with other atopic conditions and with peanut allergy in other household members. *Allergy* 54, 380-385

[6] Venter C, Pereira B, Grundy J, Clayton CB, Arshad SH, Dean T (2006) Prevalence of sensitization reported and objectively assessed food hypersensitivity amongst six-year-old children: a population-based study. *Pediatr Allergy Immu* 17, 356-363

[7] Pastorello EA, Farioli L, Conti A, Pravettoni V, Bonomi S, Iametti S, Fortunato D, Scibilia J, Bindsley-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 Imm* 144, 10-22

[8] Palosuo K (2003) Update on wheat hypersensitivity. *Curr Opin Allergy Clin Immunol* 3, 205–209

[9] Morita E, Matsuo H, Mihara S, Morimoto K, Savage AW, Tatham AS (2003) Fast ω-5 gliadin, is a major allergen in wheat dependent exercise induced anaphylaxis. *J Dermatol Sci* 33, 99-104

[10] Tatham AS, Shewry PR (2008) Allergens in wheat and related cereal. *Clin Exp Allergy* 38, 1712-1726

[11] Breitender H, Radauer C (2004) A classification of plant food allergens. *J Allergy Clin Immun* 113, 821-830

[12] Breitender H, Mills ENC (2005) Plant food allergens-structural and functional aspects of allergenicity. *Biotechnol Adv* 23, 395-399

[13] Svensson B, Fukuda K, Nielsen PK, Bønsager BC (2004) Proteinaceous α-amylase inhibitors. *Biochim et Biophys Acta* 1696, 145– 156

[14] Shewry PR, Lafiandra D, Salcedo G, Aragoncillo C, Garcia-Olmedo F, Lew EJL, Dietler MD, Kasarda DD (1984) N-terminal amino acid sequence of chloroform/methanol-soluble proteins and albumins from endosperm of wheat, barley and related species. *FEBS Lett* 175, 359-363

[15] Kusaba-Nakayama M, Ki M, Iwamoto M, Shibata R, Sato M, Imaizumi K (2000) CM3, one of the wheat α-amylase inhibitor subunits, and binding of IgE in sera from Japanese with atopic dermatitis related to wheat. *Food Chem Toxicol* 38, 179-185

[16] Sander I, Rozynek P, Rihs HP, van Kampen V, Chew FT, Lee WS, Kotschy-Lang N, Merget R, Brüning T, Raulf-Heimsoth M (2011) Multiple wheat flour allergens and cross-reactive carbohydrate determinants bind IgE in baker's asthma. *Allergy* 66, 1208-1215

[17] Larrè C, Lupi R, Gombaud G, Brossard C, Branlard G, Moneret-Vautrin DA, Rogniaux H, Denery-Papini S (2011) Assessment of allergenicity of diploid and hexaploid wheat genotypes: identification of allergens in the albumin/globulin fraction. *J proteomics* 74, 1279-1289

[18] Armentia A, Sanchez-Monge R, Gomez L, Barber D, Salcedo G (1993) In vivo allergenic activities of eleven purified members of a major allergen family from wheat and barley flour. *Clin Exp Allergy* 23, 410-415

[19] Junker Y, Zeissig S, Kim SJ, Barisani D, Wieser H, Leffler DA, Zevallos V, Libermann TA, Dillon S, Freitag TL, Kelly CP, Schuppan D (2013) Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. J Exp Med 209(13), 2395-2408

[20] Botton A, Lezzer P, Dorigoni A, Barcaccia G, Ruperti B, Ramina A (2008) Genetic and environmental factors affecting allergen-related gene expression in apple fruit (Malus domaestica L. Borkh). *J Agric Food Chem* 56, 6707-6716

[21] Sancho AI, Foxall R, Browne T, Dey R, Zuidmeer L, Marzban G, Waldron KW, van Ree R, Hoffmann-Sommergruber K, Laimer M, Mills ENC (2006) Effect of post-harvest storage on the expression of the apple allergen Mal d 1. *J Agric Food Chem* 54, 5917-5923

[22] Botton A, Andreotti C, Costa G, Ramina A (2009) Peach (Prunus persica L. Batsch) allergen encoding genes are developmentally regulated and affected by fruit load and light radiation. *J Agric Food Chem* 57, 724-734.

7 ANALYTICAL TOOLS FOR GLUTEN DETECTION: A COMPARISON BETWEEN RT-PCR, ELISA AND MASS SPECTROMETRY

The work described in this PhD chapter was performed at ILVO T&V (Instituut voor Landbouw- en Visserijonderzoek, Technologie en Voeding) located in Melle, Belgium. The abroad experience lasted from June 2013 to December 2013.

7.1 INTRODUCTION

Gluten is the major external trigger for celiac disease and the only therapy for celiac patients at the moment is the strictly avoidance of food containing gluten or related proteins from barley and rye in a life-long diet. Thus, obviously, celiac patients cannot eat common/durum wheat, triticale, spelt, kamut (all gluten containing cereals, gliadins), barley (hordeins) and rye (secalins) but also oat consumption is still debated. It seems that not all celiac patients react to oat (avenins), but anyway the contamination with wheat is difficult to avoid during harvesting and storage.[1, 2 3, 4] The "safe" gluten threshold for celiac people, that is the amount of gluten that can be ingested by celiac patients without developing intestinal mucosa inflammation and villous shortening, is not well defined yet. Several *in vivo* challenges were performed to set the minimum level of gluten intake necessary to develop symptoms, analyzing small-bowel biopsy in coeliac patients taking small amounts of gluten. The study designs and the duration and the amount of gluten load have been highly variable among the different studies and the results suggest limit values between 10 mg and 500 mg of daily gluten assumption.[5] Thus it is important that gluten can be quantified in a robust and accurate way and clearly labelled using this info. The conditions for the use of the terms related to the absence of gluten is therefore laid down at Community level (COMMISSION REGULATION (EC) No 41/2009 of 20 January 2009 concerning the composition and labelling of foodstuffs suitable for people intolerant to gluten). Gluten-free foods (max **20 ppm** of gluten) and very low gluten foods (max **100 ppm** of gluten) are foodstuffs for particular nutritional uses which have been specially formulated, processed or prepared to meet the dietary needs of people intolerant to gluten. These products can be obtained using ingredients free of gluten or they can be technologically processed to eliminate gluten proteins initially present (even if these processes are difficult and expensive, like gluten free wheat starch). The threshold of 20 ppm for gluten free foods has been set because it has been demonstrated that also a prolonged exposure to this gluten concentration in foods is not likely to trigger an immunological response in celiac patients. Technological treatments for gluten removal from wheat, rye, barley and oats derived ingredients can perhaps be not sufficient to remove gluten completely. Considering that most celiac patients can eat small amount of "very low gluten" foods without immunological reactions, a second threshold of 100 ppm has been set for ingredients that have been specifically processed to reduce the gluten content, in order to maintain the consistency and/or texture of the food and a wider range of products for celiacs.

The very low level of gluten allowed in gluten free products makes it necessary to have detection methods that have a high sensitivity, specificity and accuracy. Methods for gluten detection are essentially of two types: DNA-

based methods (Real Time PCR or RT-PCR) and protein-based methods (mostly Enzyme-Linked Immunosorbent Assay or ELISA). Recently, a new detection methodology, based on mass spectrometry, is developing.

RT-PCR methods are based on the amplification of specific genes present in wheat, barley and rye and mostly use a probe (Figure 7.1) with a reporter fluorescent dye (R) and a Quencher (Q). After the denaturation of the double strand DNA and the annealing of the primers and the probe, the exonucleasic activity of Taq polymerase cleaves the hybridized probe and releases the reporter dye from the probe. As a consequence, the fluorescence signal increases during the amplification cycles. An alternative to TaqMan probes is SYBR-Green, a fluorescent dye that non-specifically binds to double stranded DNA, thus increasing the fluorescent signal during the amplification process. Once a correct fluorescence threshold is fixed, the number of cycles necessary to reach that value will be inversely proportional to the Log10 of gene copies present in the sample. RT-PCR can be semi-quantitative (Ct<35 is positive; 35<Ct<40 is low concentration and Ct>40 is negative) or also quantitative, using a standard as calibrator and to construct a calibration curve. In this latter case, the DNA concentration is plotted against the cycle threshold and will have a negative slope (between -3.2 and -3.6). Anyway, the use of qPCR to indicate the presence of gluten is debated, because the presence of wheat DNA cannot directly be related to gluten presence. The target of the qPCR assays are not gluten proteins, but wheat genes that are not coding for gluten proteins, so this is an indirect method. Most qPCR methods have a LOQ < 10 ppm.

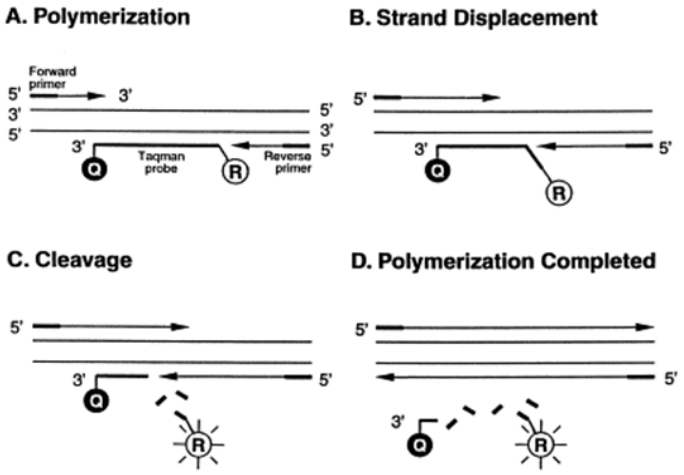

Figure 7.1 General scheme of a Real Time Polymerase chain reaction with a fluorescent probe as detection method.[6]

ELISA-based method for gluten detection are widely used for quality assurance and one of them (RIDASCREEN GLIADIN) is the only gluten detection method that has been approved (until now) by the AOAC (Association of Analytical Communities) and the Codex Alimentarius. ELISA tests exploit the use of specific antibodies directed toward short gluten amino acidic sequences. ELISA assays can be of the sandwich type or of the competitive type (Figure 7.2). In both cases an ELISA assay can be quantitative using suitable gluten standards (with well-known concentrations in ppm) to construct a calibration curve.

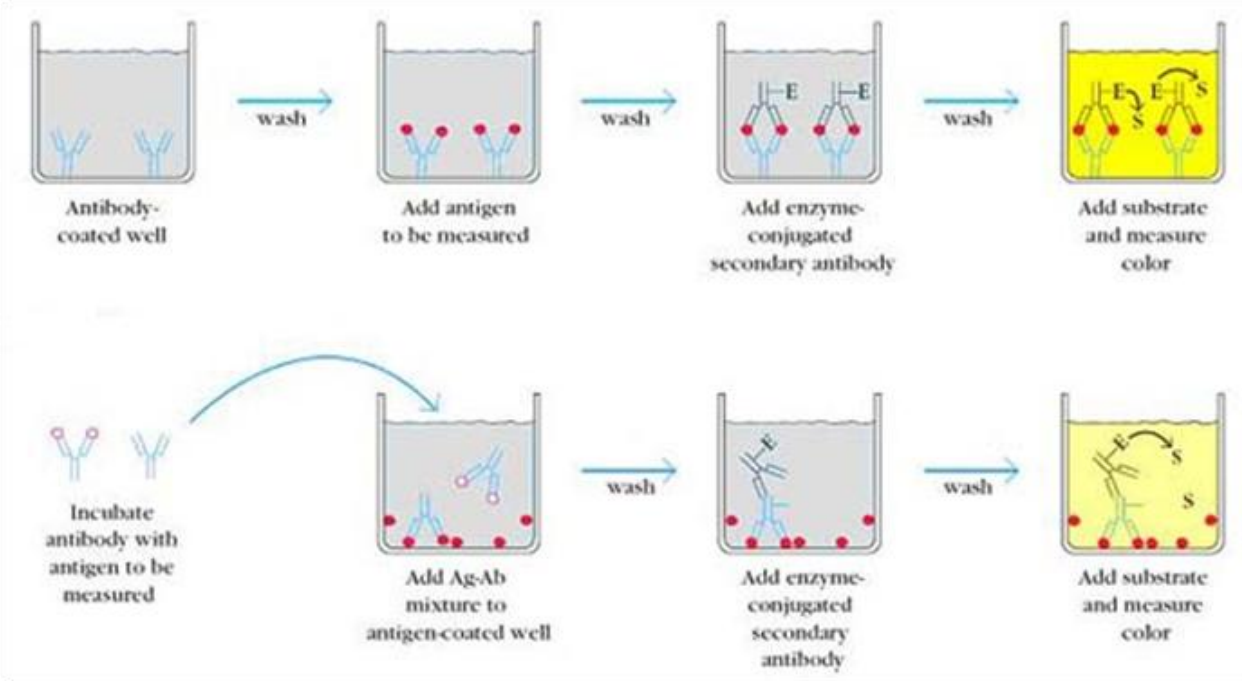

Figure 7.2 Schematic principles of sandwich ELISA (above) and competitive ELISA (below).[7]

In sandwich type ELISA the microwells are coated with antibodies specific to the target antigen (gliadin sequences); when sample is added to the well, gluten epitopes will bind with the fixed antibodies. After this step, other antibodies specific for gliadin (but conjugated with an enzyme) are added to the well and will bind the absorbed antigens. After adding the chromophoric substrate of the enzyme, only positive samples will develop a colour measurable spectrophotometrically. In competitive ELISA the sample is incubated with specific gliadin antibodies; the mixture is then added to a well precoated with gliadin antigens, where only free antibodies will bind (inversely proportional to gluten concentration in sample). A secondary antibody (conjugated with a chromophoric enzyme) is added to the well and the detection will be spectrophotometrically through addition of a chromophoric substrate. Many kits are available on the market, and they have a limit of detection (LOD) in the range of 0.3-10 ppm of gluten and limit of quantification (LOQ) of 0.8-50 ppm. Immunochemical methods are rapid, sensitive and do not require expensive equipment but, depending on the food matrix, cross reactivity can lead to false positives or to irreproducible results (especially if different kits are used).

In recent years mass spectrometry based detection methods have been increasingly developed, due to their high specificity and sensitivity. Anyway, the most important feature of mass spectrometry based methods is their ability to detect specific compounds in complex mixtures on account of their specific and characteristic fragmentation patterns. Moreover, as previously explained, mass spectrometry can be hyphenated with liquid chromatography, in order to achieve a better resolution of complex food samples and thus a lower interference from matrix compounds. Since entire proteins are often difficult to ionize in a reproducible way, MS methods take advantage of proteolytic enzymes that specifically cleave the protein in shorter peptides, hence making mass spectrometric analysis more feasible. The more frequently used approach is the reduction of the disulphide bridges of the protein with β-mercaptoethanol or dithiothreitol followed by the alkylation of the sulphidrilic groups by iodoacetic acid or iodoacetamide, thus achieving the complete denaturation of the protein that leads to a better accessibility

of the peptidic bonds by the proteolytic enzymes. The universally used enzyme for proteomic analysis is trypsin, due to its high specificity (it cleaves the peptidic bond at the C-term of basic amino acids, lysine and arginine). Moreover, the presence of a basic amino acid at the C-term of the generated peptides leads to a very good ionization in mass spectrometry (positive ions). Finally, the specific proteolytic peptides (marker for the allergenic protein) are quantified using an external calibration curve constructed using a standard peptide or with the isotopically labelled internal standard method, as shown in Figure 7.3.

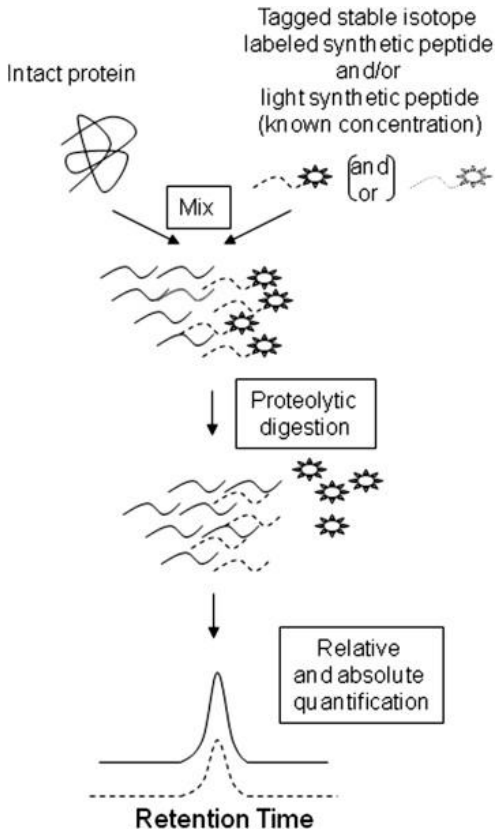

Figure 7.3 Absolute protein quantification using a specific marker peptide: the sample is spiked with the labelled standard peptide that will be used to quantify the peptides generated after enzymatic digestion of the sample.[8]

Given the complexity of food samples (especially if highly processed) and the trace amounts in which allergens need to be detected, usually MS/MS methods are employed: the most used is Multiple Reaction Monitoring (MRM), in which the selected m/z of the marker peptide is fragmented and one fragment ion is used for the quantification, while two other fragment ions are used for confirmation.

Mass spectrometry based methods for gluten detection have already been developed for wine, in which gluten is sometimes used as technological adjuvant for clarification;[9] this matrix is particularly difficult to analyze by ELISA due to the high presence of interfering compounds such as sugars and polyphenols, but anyway the high specificity and sensitivity of MS allows an accurate detection. A method of detection that quantifies six immunogenic gluten peptides was recently developed, and applied to a wide variety of food and non-food products submitted to simulated gastrointestinal digestion.[10] LOD and LOQ of this method range respectively between 1-30 ppb and between 10-100 ppb, depending on the peptide used for the detection. Mass spectrometry

is highly sensitive and can give an absolute identification and quantification of the allergen, but it is quite expensive and requires trained operators.

7.2 MATERIALS AND METHODS

7.2.1 Reagents and solvents

DNA extraction. Ultrapure water was purchased from Biochrom (Berlin, Germany), the kit for DNA extraction (DNAeasy plant mini kit) was purchased from QIAgen (Venlo, Netherlands). Ethanol was purchased from Merck (Darmstadt, Germany).

Polymerase Chain Reaction. PinB forward and reverse primers and PinB probe were purchased from Eurogentec (Seraing, Belgium). TaqMan Universal PCR Mastermix was purchased from Bioconnect (Applied Biosystems, Foster City, CA, USA). SureFood Allergen QUANT Gluten was purchased from R-Biopharm (Darmstadt, Germany), as the SureFood QUANTARD Allergen 40.

Enzyme Linked Immunosorbent Assay. Seven different ELISA kits were used, according with the instruction provided.

Simulated gastrointestinal digestion. Deionised water was obtained from a Millipore Alpha Q-Waters purification system (Billerica, MA, USA). Potassium thiocyanate, sodium hydrogen carbonate, ammonium chloride, potassium phosphate monobasic, D-glucose, D-glucuronic acid, D-(+)-glucoseamine hydrochloride, uric acid, mucin from porcine stomach, was purchased from α-amylase from hog pancreas, pepsin from porcine gastric mucosa, pancreatin from porcine pancreas, bile from bovine, lipase from porcin pancreas, potassium chloride, sodium dihydrogen phosphate, sodium sulfate, were purchased from Sigma Aldrich (St. Louis, MO, USA). Sodium chloride, calcium chloride dehydrate, magnesium chloride, albumin bovine, urea, acetonitrile, hydrochloric acid (37% V/V), formic acid, sodium chloride were purchased from Merck (Darmstadt, Germany). Urea was purchased from GibCo-BRL (Life Technologies, Carlsbad, CA, Stati Uniti).

7.2.2 DNA extraction

The DNA extraction was carried out using the Qiagen DNeasy Plant Mini Kit, a commercial kit from Qiagen especially developed for vegetal matrices. Usually DNA extraction is performed on young shoots because the cell density is higher and a greater amount of DNA can be extracted. The food matrices tested instead were ground kernels, with a starchy endosperm poor in DNA. So, the amount of starting material was increased from 20 to 100 mg. 100 mg of each sample were weighted in a 1.5 ml eppendorf tube (100 µl of HPLC water as negative control) in duplicate. The SureFood Allergen 40 was used as positive control, but the extraction was performed in a different day, in order to avoid cross-contamination of the samples. Then, 4 µl of RNase A Stock Solution and 400 µl of AP1 extraction buffer were added to lyse the cells and destroy RNA; samples were vortexed and incubated for 30 minutes at 65°C to lyse the cells. Then, 130 µl of AP2 precipitation buffer were added to remove proteins and polysaccharides; samples were vortexed and incubated 5 minutes on ice. Samples were centrifuged

at 20000*g* for 15 minutes and the supernatant was loaded onto a QIAshredder spin column placed in a 2 ml collection tube. Samples were centrifuged for 2 minutes at 20000*g* to remove cell debris and precipitates, obtaining a clear lysate. The flow-through was transferred into a new 1.5 ml eppendorf tube without disturbing the pellet and 1.5 volumes of AP3 binding buffer (diluted in ethanol 96% in a ratio 1:2) were added. After vortexing, 650 µl of the mixture were transferred onto a DNeasy Mini spin column placed in a 2 ml collection tube. Samples were centrifuged for 1 minute at 6000*g* and the flow-through was discarded. This step was repeated with the remaining amount of sample, modifying the centrifugation step (2 minutes at 14000*g*). The spin column was transferred into a new 2 ml collection tube and 500 µ of AW wash buffer were added. Samples were centrifuged for 1 min at 6000*g* and the flow-through was discarded. This step was repeated modifying the centrifugation step (2 minutes at 20000*g*). The spin column was transferred into a new 1.5 ml eppendorf tube and 100 µl of AE elution buffer were added. After 5 minutes of incubation at RT, samples were centrifuged for 2 minutes at 6000*g*. This step was repeated and the eluted DNA was stored at -20°C.

7.2.3 DNA quantification

The DNA concentration was quantified using the NanoDrop spectrophotometer (Thermo scientific, DE, USA), using the AE elution buffer of the DNA extraction kit as blank. The DNA concentration is determined from the absorbance at 260 nm (wavelength of maximum absorption of nucleic acids), while DNA purity is assessed making the ratio between the absorbance at 260 nm and 280 nm (wavelength of maximum absorption of proteins and phenols). Samples were subsequently diluted with HPLC-water to obtain a concentration of 10 ng/µl.

7.2.4 Polymerase Chain Reaction – Piknova et al. (2008)

The method was developed by Piknova and co-workers in 2008.[11] The primers were agcacttctcccgaacctca (Pinb Forward, 5 µM), gatggagcgatgtttcacaa (Pinb Reverse, 5 µM) and the TaqMan probe was FAM-ctcacagccgcccttccacca-TAMRA
(Pinb Probe, 5 µM). The mastermix was prepared using the ratio PCR mix : Pinb F : Pinb R : Pinb P : HPLC-water = 5 : 1 : 1 : 1 : 12 and calculating 20 µl of mastermix for every well. Then, 20 µl of mastermix were aliquoted in every well of the microplate and 5 µl of sample DNA were added (5 µl of HPLC-water as negative control).
The PCR program was: 5 min initial denaturation step at 95°C→ 50 cylces of denaturation step (15 sec at 95°C) followed by an annealing step (60 sec at 60°C).

7.2.5 Polymerase Chain Reaction – SureFood Quant Gluten

The Gluten Master Mix was prepared by adding 0.1 µl of Taq Polymerase to 19.9 µl of Gluten Reaction Mix (for every sample). In the same way, the Gluten Inhibition Mix was prepared by adding 0.1 µl of Taq Polymerase to 19.9 µl of Inhibition Control Mix (for every sample). Both Master Mixes were mixed well and centrifuged shortly before use. The Standard DNA was diluted with the Dilution Buffer as reported in Table 7.1.

Table 7.1 Preparation of the Standard DNA dilutions.

Standard	Dilutions	Copy number per µl	Final copy number per reaction
S1	45 µl Dilution Buffer + 5 µl Standard DNA	10000	50000
S2	45 µl Dilution Buffer + 5 µl DNA of S1	1000	5000
S3	45 µl Dilution Buffer + 5 µl DNA of S2	100	500
S4	45 µl Dilution Buffer + 5 µl DNA of S3	10	50

20 µl of the gluten mastermix were pipetted into appropriate wells and 5 µl of sample DNA or SureFood QUANTARD Allergen 40 DNA or standard dilutions were added into the designated wells. The same procedure was repeated for the Inhibition Control Mix. The microplate was centrifuged at 3000 *g* for 1 min and the PCR was run with the following parameters: initial denaturation step at 95°C→45 cycles of denaturation step (15 sec at 95°C) followed by an annealing step (30 sec at 60°C).

7.2.6 Simulated gastrointestinal digestion

The samples were digested as described by Versantvoort et al (2005),[17] adapting the volume of digestive juices to the smallest amount of sample. Briefly, 450 mg of sample were incubated 5 minutes with 600 µl of saliva; after this phase, 2.4 ml of gastric juice were added and the sample was incubated for 2 hours. To set the pH for the intestinal phase, 400 µl $NaHCO_3$ 1 M were added; after, 2.4 ml of duodenal juice and 1.2 ml of bile were added and the sample was incubated for 2 hours. All the digestion steps were carried out at 37°C. At the end of the digestion, 58.3 µl of HCl 37% were added and the sample was centrifuged at 8965 g at 4°C for 45 minutes, in order to precipitate insoluble compounds and undigested proteins. Prior to LC-MS analysis, all samples were filtrated with a cut off of 0.45 µm. For the calibration curves, 2.25 g of blank matrix (rice, bread mix, pasta or muesli) were spiked with 23 mg of pure wheat gluten in order to obtain a 10000 ppm standard solution. The standard was then digested and serial dilutions were made with the respective blank matrices undergone digestion, in order to obtain a calibration curve in matrix. The calibration curves were made in the range 0-300 ppm with the following points: 0, 5, 10, 20, 50, 100, 200 and 300 ppm.

7.2.7 UPLC/ESI-MS/MS analysis

Sample were separated by a RP column (ACQUITY UPLC BEH 300 C18 1.7 µm 2.1*150 mm equipped with a ACQUITY UPLC BEH C18 VanGuard Pre-column, 300Å, 1.7 µm, 2.1 mm X 5 mm) in an UPLC/ESI-MS/MS system (UPLC Acquity Waters with a triple quadrupole mass spectrometer Waters XEVO-TQS) using a gradient elution. Eluent A was water with 0.1% formic acid and 0.2% acetonitrile, eluent B was acetonitrile with 0.1% formic acid; gradient: 0-7 min 100% A, 7-50 min from 100% A to 50% A, 50-52,6 min 50% A, 52,6-53 min from 50% A to 0% A, 53-58,2 min 0% A, 58,2-59 min from 0% A to 100% A, 59-72 min 100% A. The digested sample extracts were analysed with UPLC/ESI-MS (flow 0.2 ml/min; analysis time 72 min; column temperature 35°C; sample temperature 18°C; injection volume 5 µl; acquisition time 7-58,2 min; ionization type positive ions; capillary voltage 3.2 kV; cone voltage 60 V; source temperature 150°C; desolvation temperature 350°C; cone gas flow 100 l/h; desolvation gas flow 650 l/h; injection volume 2 µl)

- ✓ in the Full Scan mode (scan range 100-2000 m/z)
- ✓ in the Daughters Scan mode (scan range and collision energy depending on the ion to be fragmented)
- ✓ in the Multiple Reaction Monitoring mode (Scan range 100-2000 m/z, Collision Energy 28, quantification transition 717.6→244.0, confirmation transitions 717.6→354.0 and 717.6→468.4).

7.3 RESULTS AND DISCUSSION

In the first part of the work, both gluten containing foods and naturally gluten-free foods (Table 7.2) were analyzed for gluten content using PCR methods (qualitative and quantitative) and the results were compared with those arising from ELISA detection. The qualitative PCR methods was taken from the literature,[11] while for the quantitative analysis a commercially available kit (SureFood Quant Gluten) was used. Several ELISA kits for gluten quantification are available on the market. For this experiment, the RIDASCREEN Gliadin from r-biopharm was chosen, because it is certified by AOAC-OMA, AOAC-RI, Codex Alimentarius and AACCI. The samples analyzed were all raw materials and not processed foods, in order to avoid the introduction of an additional source of variability in the comparison. Among the gluten containing cereals, different durum and common wheat varieties were tested, together with barley and triticale (a hybrid between wheat and rye). The naturally gluten-free samples tested were rice, soy, corn and oats.

Table 7.2 Samples analyzed with qualitative and quantitative PCR and with sandwich ELISA.

SAMPLE CODE	SPECIES	VARIETY/TYPE
1	*Triticum turgidum* spp *durum*	D240
2	*Triticum turgidum* spp *durum*	Svevo
3	*Triticum turgidum* spp *durum*	Levante
19	*Triticum aestivum*	C172
24	*Triticum turanicum*	Kamut
5910	*Triticum turgidum* spp *durum*	F312
5911	*Triticum turgidum* spp *durum*	F255
5914	*Triticum turgidum* spp *durum*	Aureo
5916	*Triticum turgidum* spp *durum*	Levante
5924	*Triticum turgidum* spp *durum*	F371
N°3	*Glycine max*	Raw brazil soybeans
N°4	*Hordeum vulgare*	French brewer's barley
N°5	*Avena sativa*	Oats
N°6	*Triticosecale*	Triticale
N°16	*Triticum aestivum*	Wheat meal fine processed
N°18	*Zea mays*	Maize gluten pellets
N°39	*Triticum aestivum*	Pure wheat gluten
N°48	*Oryza sativa*	Rice

7.3.1 DNA extraction

To have comparable results among the samples, it is fundamental to start from the same amount of DNA in every PCR reaction. Thus, the DNA concentration was quantified with the NanoDrop spectrophotometer, using the AE elution buffer as blank. The instrument measures the absorbance at 260 nm, that is the maximum absorption wavelength of nucleotides. The purity of the extracted DNA is assessed making the ratio between the absorption of nucleotides at 260 nm and that of proteins at 280 nm. Results are reported in Figure 7.4. The yields of extracted DNA were very good for all the wheat varieties, while DNA yields were lower for the gluten-free samples, probably due to a different matrix effect during the extraction, combined with a different DNA concentration in the sample. Pure wheat gluten was extracted as positive control. The purity of the extracted DNA was good, with a ratio A_{260}/A_{280} around 1.8 for all samples, except for barley and flour (1.9-2.0), corn gluten and rice (1.4-1.7, protein contamination). The ratio A260/A230 was generally good (2.0-2.2) except for soy, barley and corn gluten (0.8-2.0), where probably a contamination by sugars/phenols was present. Samples were diluted with HPLC water in order to obtain a final concentration of 10 ng/µl of DNA.

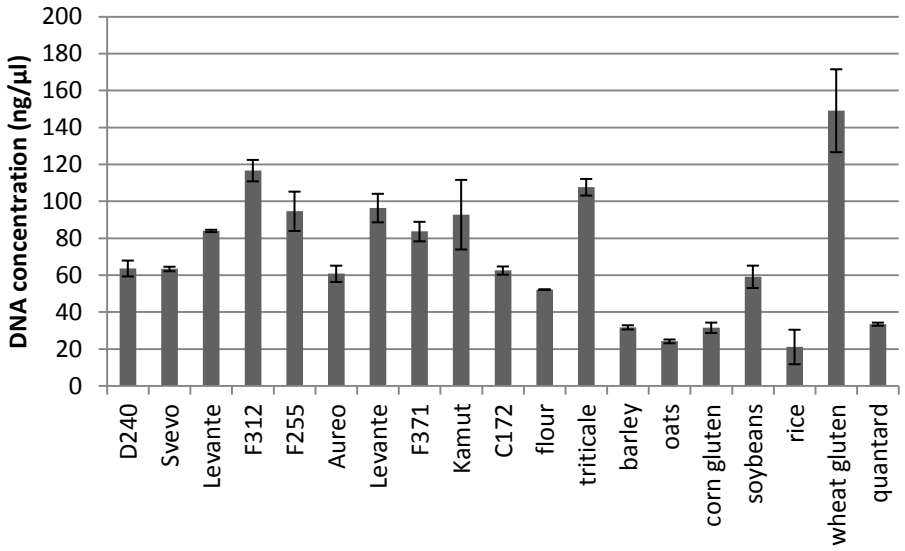

Figure 7.4 Amount of DNA extracted from the samples using the QIAgen DNeasy Mini Plant Kit.

7.3.2 Polymerase Chain Reaction – Piknova et al.

For a first screening, a qualitative Real Time PCR was performed using the method developed by Piknova et al. in 2008. The detection is achieved using a TaqMan probe and the target gene is Puroindoline B (PinB), a gene previously used in literature (Alary et al., 2002) to assess common wheat contamination/adulteration in durum wheat flour (thus not present in *Triticum turgidum* spp *durum* species). Gluten coding genes cannot be used as target (as it would be logic to do) because they are not conserved, showing a too high variability. Instead the more conservative puroindoline b gene present in wheat, but also in barley and rye (toxic for celiac patients too) is used. The limit of detection of the method is 200 ppm.

In Table 7.3 the results obtained are reported and it is clear that all the durum wheat varieties are not detected. This because the target gene (puroindoline b) is located on chromosome D, only present in hexaploid species (wheat and triticale). So, with this method, all the contaminations from durum wheat (largely used for pasta production) are not detected.

Table 7.3 Results of the polymerase reaction performed on different gluten-containing and naturally gluten free foods, performed with the method of Piknova et al.

SPECIE	VARIETY	**Ct**	RESPONSE
Triticum turgidum spp *durum*	D240	**0.00**	NEGATIVE
Triticum turgidum spp *durum*	Svevo	**40.31**	
Triticum turgidum spp *durum*	Levante	**0.00**	NEGATIVE
Triticum turanicum	Kamut	**37.10**	
Triticum turgidum spp *durum*	F312	**0.00**	NEGATIVE
Triticum turgidum spp *durum*	F255	**0.00**	NEGATIVE
Triticum turgidum spp *durum*	Aureo	**0.00**	NEGATIVE
Triticum turgidum spp *durum*	Levante	**0.00**	NEGATIVE
Triticum turgidum spp *durum*	F371	**0.00**	NEGATIVE
Hordeum vulgare	French brewer's barley	**36.74**	POSITIVE
Triticosecale	Triticale	**34.72**	POSITIVE
Triticum aestivum	C172	**30.09**	POSITIVE
Triticum aestivum	Wheat meal fine processed	**29.45**	POSITIVE
Glycine max	Brazil soybean	**0.00**	NEGATIVE
Oryza sativa	Rice	**0.00**	NEGATIVE
Avena sativa	Oats	**39.19**	POSITIVE
Zea mays	Corn gluten pellets	**42.62**	
Triticum aestivum	**Pure wheat gluten**	**28.64**	**POSITIVE**

While common wheat is detected with good Ct values (comparable with those of the pure gluten standard), barley and rye are positive but with higher Ct values. The reason is that primers and probes are designed for wheat puroindoline b gene: so, even small differences in the gene sequence can decrease primers and probe annealing efficiency to the template. This problem could be partially solved using SYBR green method instead of the TaqMan probe. This dye in fact bind aspecifically double stranded DNA, so it can overcome the problem of a poor annealing of the probe to the amplicon due to the variability in the nucleotidic sequence of the gene in different species. Soy and rice resulted negative as expected, while corn and oats showed a positive result, even if the Ct value was high, thus indicating a possible contamination.

7.3.3 Polymerase Chain Reaction – SureFood Quant Gluten

For the quantitative Real Time PCR a commercial kit from r-biopharm was used; it makes use of a TaqMan probe as is the case in PCR developed by Piknova et al. The target in this PCR method is a multicopy gene present in all cereals being toxic for celiac patients (wheat, rye, barley, oat, spelt and kamut). This method is more sensitive, showing a Limit Of Quantification under 1 ppm and a linear range for quantification up to 400 ppm. For the quantification, a standard matrix spiked with 40 ppm of gluten is provided (SureFood Quantard

Allergens 40). The calibration curve (Figure 7.5) is obtained with several dilutions of a gluten standard from 50 to 50000 copies of DNA per reaction. Thus, the number of copies is plotted against the Ct value obtained. Once the Ct of the sample is known, interpolation with the calibration curve allows to determine the number of DNA copies in the tested sample. To obtain the concentration, a simple proportion with the known 40 ppm standard is performed.

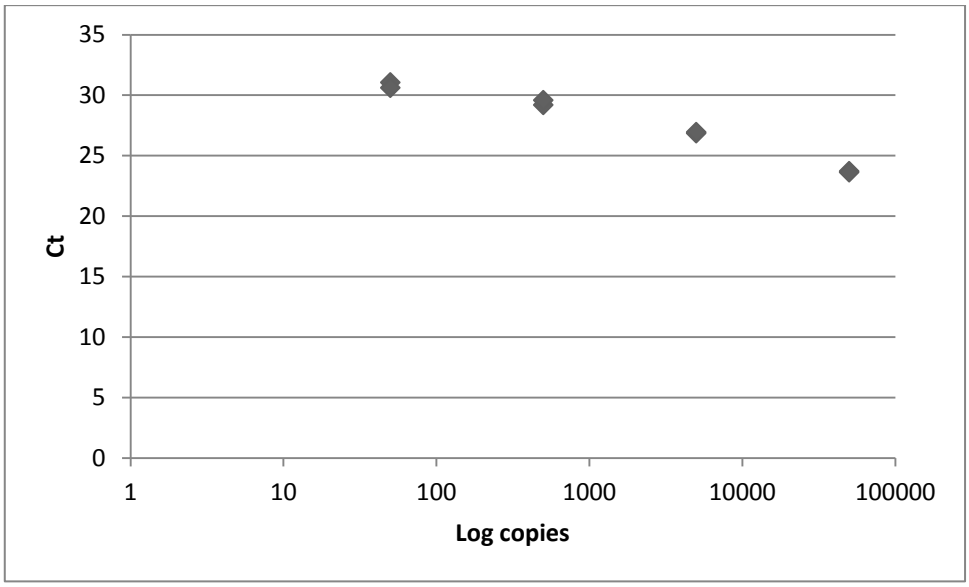

Figure 7.5 Calibration curve obtained for the gluten standard provided.

To avoid false negative results due to PCR inhibition by interfering compounds present in the sample, an inhibition control reaction is performed for each sample. In these "parallel reactions", a known template is amplified in presence of the sample and the reaction must be positive. A negative result means that the sample contains compounds that inhibit the polymerase chain reaction.

As shown in Figure 7.6, all tested gluten containing cereals resulted positive (as expected) but the strong differences among varieties exceed the natural variability in gluten content. This means that different genotypes show a different number of copies of the target gene, so the response factor is different. For what concerns naturally gluten free samples (Figure 7.7), corn, soy and rice were under 20 ppm, that is the legal limit to declare a food as "gluten free",[12] but anyway a signal was detected. Oat sample was above this limit: the consumption of oats by celiac patients is still debated among physicians, both for the sequence similarity between wheat gliadins and oat avenins, and also because of the frequent contamination of oats with wheat that can potentially occur during harvesting, milling and storage.

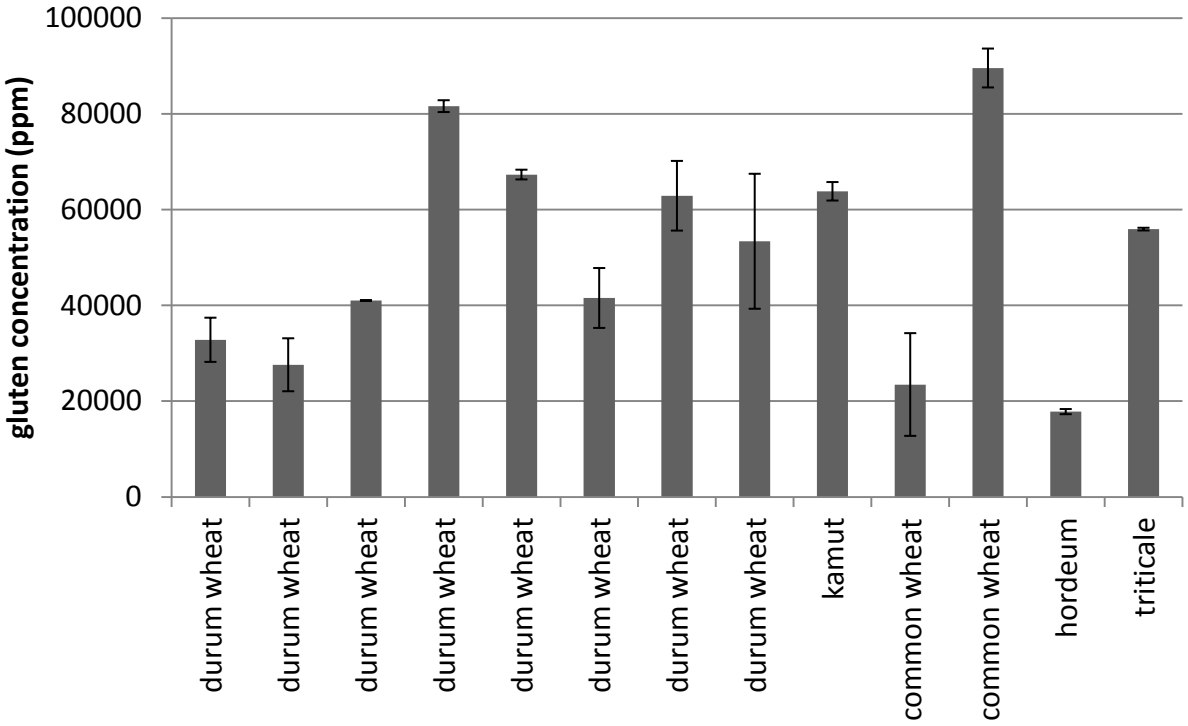

Figure 7.6 Results obtained from the quantitative Real Time PCR performed on gluten containing cereals.

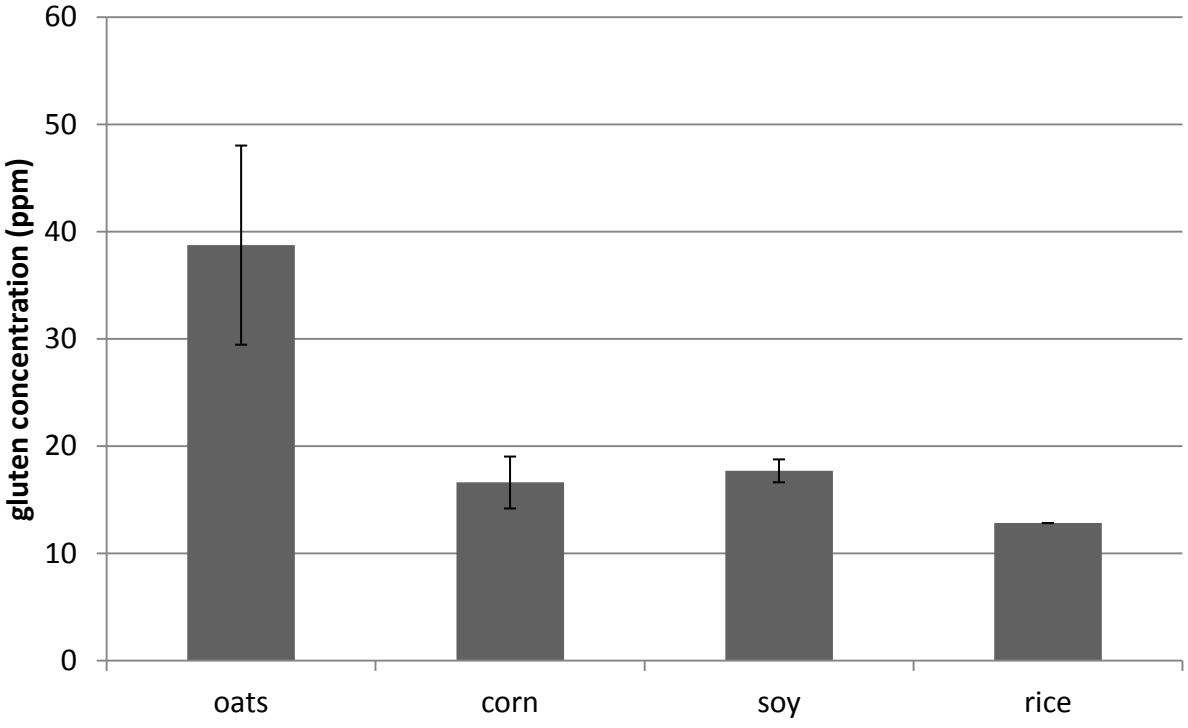

Figure 7.7 Results obtained from the quantitative Real Time PCR performed on naturally gluten free samples.

7.3.4 Sandwich R5 ELISA

For the immune-enzymatic assay the kit from R-Biopharm (Ridascreen Gliadin) was used. The capture antibody used is the R5 developed by Mendez et al.[13] in 2006: R5 is a monoclonal antibody that recognizes the specific

epitope QQPFP. The method is suitable both for native and heated gluten, but not for hydrolyzed samples. In fact, a requirement of the sandwich ELISA is the presence of two epitopes on the same polypeptide, because two antibodies must bind to the same molecule. Hydrolysis reduces the probability to have two epitopes on the same polypeptide, because proteins are cleaved into short peptides. The quantification range is 5-80 ppm of gliadin that, considering a ratio gliadins : glutenins = 1 : 1, correspond to 10-160 ppm of gluten. The standards provided are calibrated on the standard of the prolamin working group.[14] Samples were extracted with the RIDA extraction solution that extracts only the monomeric gliadins but not the polymeric glutenins, so the conversion factor to calculate the gluten amount starting from the determined gliadin amount is ×2. The composition of the solution is confidential, but anyway it contains 2-mercaptoethanol to reduce the disulphide bridges of the gluten network. Once extracted, the samples, must be diluted at least 1:500 prior to the immunoassay, otherwise 2-mercaptoethanol will denature the capture antibodies on the well surface. So, different sample dilutions were tested to check the correct dilution that allows to be in the calibration range: 1:500 (advised in the kit for samples containing around 20 ppm of gluten), 1:2500 (advised in the kit for samples containing around 100 ppm of gluten), 1:25000 and 1:250000 (since also pure wheat samples were analyzed, the gluten concentration was very high so also these last two very high dilutions were tested). In Table 7.4 the optical densities of the highest and the lowest point of the calibration curve (that thus define the calibration range) and of the different dilutions tested are reported. It is clear that the OD values of the first three dilutions (1:500, 1:2500 and 1:25000) are not consistent with the dilution factor, remaining around a value of 2 also for a 50 fold dilution. The increasing OD with the increasing dilution factor is probably due to 2-mercaptoethanol (and other interfering compounds) dilution that leads to a better performance of the capture antibodies. Since an OD value of 2 seems to correspond to the saturation of the capture antibodies, for the wheat samples a dilution of 1:250000 was used for the following tests.

Table 7.4 Optical densities at 450 nm of the lowest and the highest points of the calibration curve and of the dilutions tested.

WELL	*OD 450 nm*
Water	0.146
STD 1 – 5 ppb	0.155
STD 6 – 80 ppb	2.054
Wheat 1:500	1.887
Wheat 1:2500	2.094
Wheat 1:25000	2.355
Wheat 1:250000	0.581

From the results obtained from the analysis of gluten containing cereals (Figure 7.8) it can be observed that there are huge differences among the different varieties tested, with hexaploid species (common wheat and triticale) giving a higher response compared to tetraploid species (durum wheat and kamut). This is probably due to a different presence of the epitope detected by the R5 monoclonal antibody (QQPFP) in the different varieties. Thus, also in this case, contamination by durum wheat flour can be underestimated. Barley gave a high

response, but this data was already known in literature and it is due to the extraction protocol: with the extraction solution suggested, only the monomeric fraction of gluten (gliadins) is extracted and to obtain the gluten content it is necessary to multiply the result for a factor of 2. Since hordeins are more soluble than glutenins, the barley contamination is overestimated with this method.

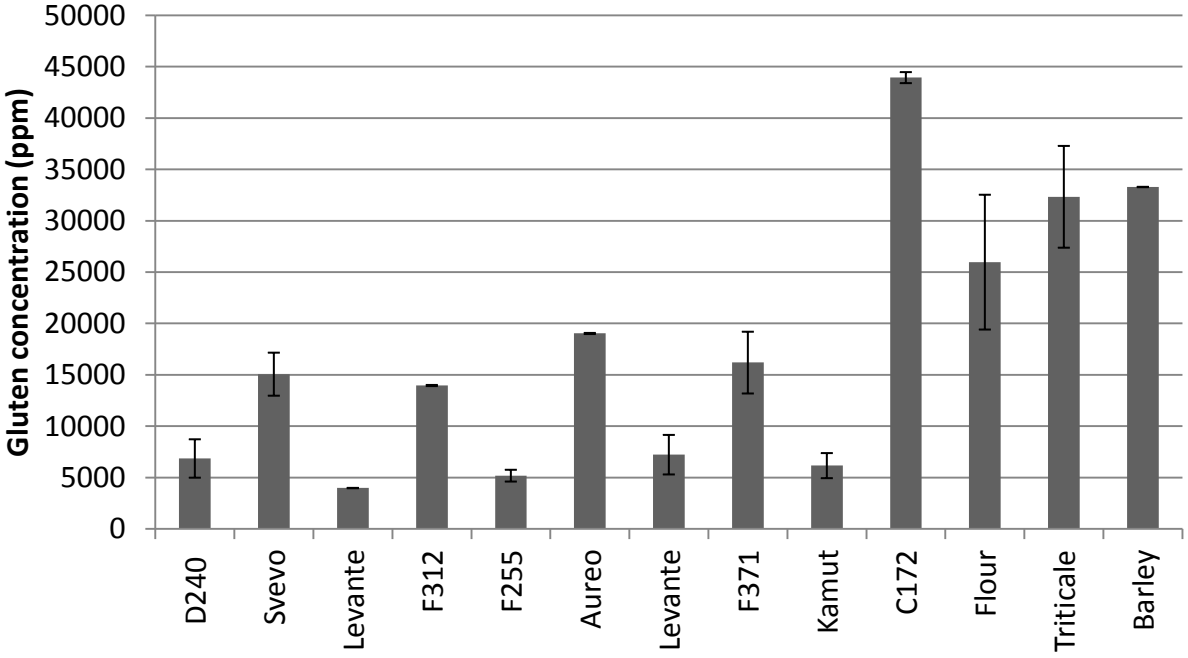

Figure 7.8 Results obtained from sandwich R5 ELISA performed on gluten containing samples.

For what concerns naturally gluten free samples (Figure 7.9), rice resulted under the Limit of Quantification, while soy was above the LOQ but anyway under the legal limit of 20 ppm. Also in this case oats resulted positive, thus confirming the doubts about its use by celiac patients. Corn gluten, which was below 20 ppm with the quantitative RT-PCR, is strongly positive now, so this could be a false positive result. Corn gluten feed is a by-product from the manufacture of cornstarch and corn syrup, obtained by adding a suitable binder (e.g. 1 - 3% of molasses, fat or colloidal clays) and then pressing the composition under high pressure in pelletizing machines or extruders to form cylindrically shaped pellets. These processes lead to chemical changes that, beside the brown colour clearly visible also in the protein extracts, probably led also to the production of interfering compounds that negatively affect ELISA assay.

So, in this first part of the work, we observed that both RT-PCR and ELISA methods are affected by genotypic differences, but RT-PCR methods have a broader range for quantification and give higher responses. Gluten is a complex network of proteins bound by disulphide bridges, so its extraction usually result more difficult than DNA extraction. As a counterpart, DNA is not directly related to the protein presence, because gene expression can be different and, anyway, DNA and gluten have different stability to food processing and show different extraction yield.

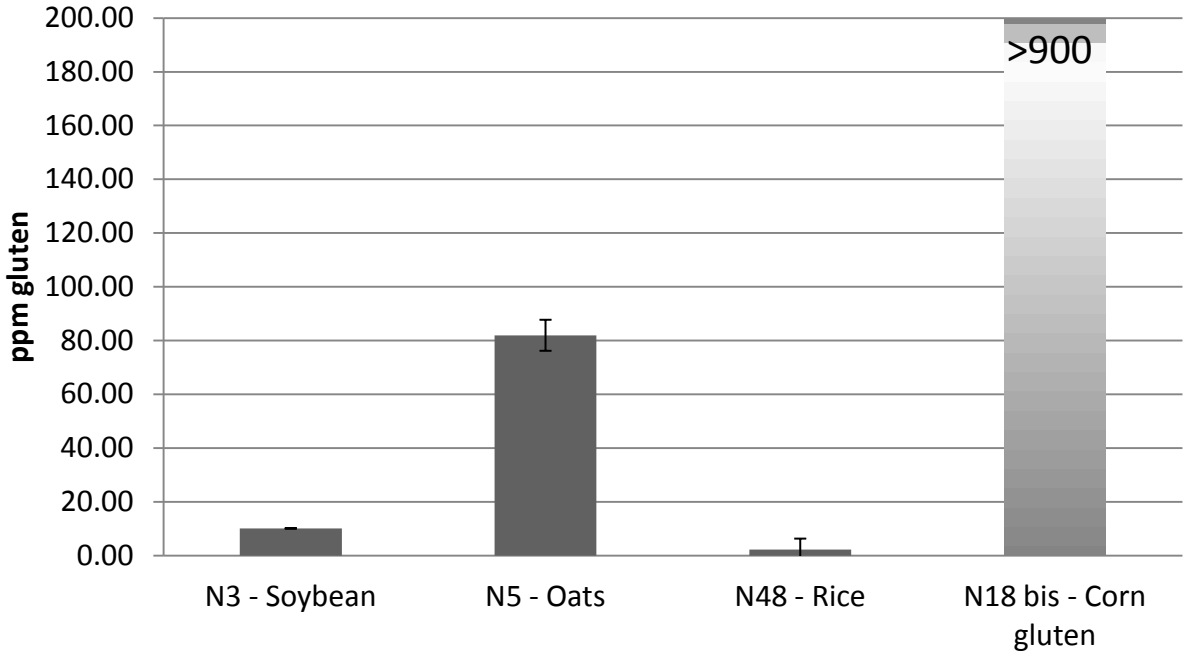

Figure 7.9 Results obtained from sandwich R5 ELISA performed on naturally gluten free samples.

7.3.5 Validation study on commercially available ELISA kits for gluten detection

For this experiment three different matrices (bread mix, pasta and muesli) were bought at a gluten free shop (Iekkerglutenvrij, Belgium) and spiked with a gluten standard (Gluten from wheat, Sigma, protein content > 80%) in order to obtain the following concentrations: 0 ppm, 0.2 ppm, 2 ppm, 20 ppm and 200 ppm. After accurate homogenization, samples were first screened with qualitative PCR using the method of Piknova et al. So, DNA was extracted from the samples using the Qiagen DNeasy Mini Plant Kit. The DNA yield of these processed food was much lower than the DNA yield from the raw materials (Figure 7.10). The high starch content of these matrices is probably the main cause of the low DNA content, because starch granules are very poor in DNA. Moreover, starch gelatinization during cell lysis (carried out at 65°C) decreases DNA extractability. In this case, maybe a pre-treatment with α-amylase of the sample could be taken into account. Other interfering compounds can be the thickeners and the emulgators used in gluten free products to mimic the gluten network. Giving the very low amount of DNA extracted the qualitative RT-PCR was performed with undiluted samples.

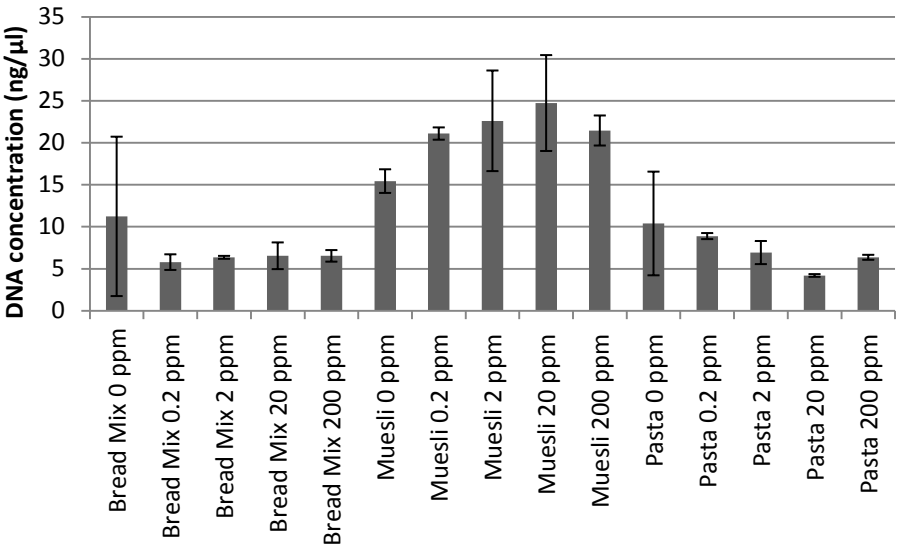

Figure 7.10 DNA concentrations obtained from the DNA extraction, determined using the absorbance at 260 nm.

As shown in Table 7.5, the limit of detection of the method is 200 ppm, so all the lower spiking levels resulted negative confirming the lack of contamination of the blank matrix employed, with a few exceptions that were probably false positive. For the 200 ppm level, only muesli showed four positive results, probably because it was the only matrix to have a good DNA concentration (Figure 7.10). The DNA yields from the pasta and bread mix were lower, so probably not sufficient to achieve a correct amplification.

Table 7.5 Results obtained from the qualitative PCR. Each sample was extracted in duplicate and each replicate was amplified in double. So, for every samples four PCR were run.

SAMPLE	RESPONSE
Bread mix 0 ppm	- - - -
Bread mix 0.2 ppm	- + - +
Bread mix 2 ppm	- - - -
Bread mix 20 ppm	- + - -
Bread mix 200 ppm	+ - - -
Pasta 0 ppm	- - - -
Pasta 0.2ppm	- - - -
Pasta 2 ppm	- - - -
Pasta 20 ppm	- - - -
Pasta 200 ppm	? - + -
Muesli 0 ppm	- - - -
Muesli 0.2 ppm	- - - -
Muesli 2 ppm	- - - +
Muesli 20 ppm	- - - -
Muesli 200 ppm	+ + + +

So, in conclusion of this first part on gluten detection using RT-PCR based methods, it can be said that these methods work very well on the raw matrices, as seen in paragraphs 7.3.2 and 7.3.3. In fact, the critical step in these methods seems to be the extraction of a good amount of DNA, both in terms of number of copies extracted and of quality of the DNA (not degraded). So, in the raw matrices (whole cereals) the amount of DNA extracted is very high and of good quality because not undergone to any food processing and thus the amplification achieved during the PCR reaction is very accurate. With more processed matrices instead, as seen in this chapter, the DNA yield can be really low and not sufficient for a good amplification reaction. In these cases, new extraction methods should be tried, maybe with the aid of an α-amylase treatment before extraction in the case of highly starchy matrices.

After the PCR, samples were analyzed with seven different ELISA kit commercially available for gluten detection (Table 7.6). So the five different spiking levels of the three different matrices were analyzed with the kits following the instruction of the manufacturer.

Table 7.6 Commercial kit for gluten detection used for the validation study. LOD, LOQ and quantification range were all converted to ppm of gluten, since the measure units were different (gliadin, gluten, peptide, protein,...).

CODE	ASSAY TYPE	LOD	LOQ	RANGE	CALIBRATION CURVE EQUATION
Kit 0	Sandwich	4 ppm	10 ppm	10-160 ppm	Cubic spline
Kit 1	Sandwich	5 ppm	5 ppm	2.5-40 ppm	Cubic spline
Kit 2	Competitive	5 ppm	5 ppm	5-160 ppm	Logarithmic curve
Kit 3	Sandwich	2 ppm	4 ppm	4-200 ppm	Cubic spline
Kit 4	Sandwich	0.3 ppm	0.3 ppm	0.3-20 ppm	Cubic spline
Kit 5	Sandwich	0.3 ppm	0.3 ppm	0.3-20 ppm	Linear curve
Kit 6	Sandwich	5 ppm	5 ppm	5-50 ppm	Polinomial 2^{nd} order

Results are reported in Table 7.7.

For the 0 ppm (blank matrix), negative results (<LOD) are obtained for all the three matrices tested using kit 0, kit 2, kit 3 and kit 5. For the other kits some cases of false positive are verified:

- Kit 1 gives results>LOQ for Bread Mix and Muesli
- Kit 4 shows false positives for Bread Mix and for Muesli (probably it is due to the very low LOQ declared for the kit, because the amount of gluten detected is really low: 0.4-0.5 ppm)
- Kit 6 gives false positives for all the matrices, but in this case the kit performance was very bad for all the spiking levels due to poor coloration of the wells also in the calibration curve.

For the 0.2 ppm level results should have been negative for all the kit tested, because the LOD of all the kits is below this value. Kit 1, kit 2 and kit 3 give negative results (according to their sensitivity). For the other kits some cases of false positive are observed: notwithstanding the amount of gluten spiked was below the LOD, a positive signal is observed. More specifically:

- Kit 4 gives false positives (detected gluten>LOQ) for Bread Mix. Like the previous level, this can be probably explained by the very low LOQ declared for the kit.
- For kit 6, false positives are observed for Pasta, while for the other two matrices (BreadMix and Muesli) results were above LOD (so detectable) but below the LOQ (not quantifiable). Like in the previous level, the kit performance was very bad for all the spiking levels due to poor coloration of the wells also in the calibration curve.

For the level 2 ppm, only the kits 4 and 5 declared a LOQ below that level. The quantification is very good in the Bread Mix matrix, while it shows lower values for Pasta and Muesli, probably due to matrix effect.

From the other kits a negative result would be expected, because they declared a LOQ>2 ppm. Indeed, only kit 2 give negative results for all the three matrices in both laboratories. Some cases of false positive are observed also in this case:

- Kit 1 shows values>LOQ for BreadMix
- Kit 3 shows values between the LOD and the LOQ (so: detectable but not quantifiable) for BreadMix and Muesli
- Kit 6 gives false positives for all the matrices, but in this case the kit performance was very bad for all the spiking levels due to poor coloration of the wells also in the calibration curve.

For the 20 ppm level, all the kits are above the LOQ. Results for pasta 20 ppm and breadmix 200 ppm (kit 1) are in single data, because one replicate was clearly an outlier, probably due to insufficient washing. Kit 2 (the only competitive assay tested) gives very underestimated results for all the three matrices tested compared to all the other kits, that are sandwich assays. Also results of kit 3 are quite underestimated, anyway the 20 ppm is always clearly detected. Concentration determined with the kits 4 and 5 are good for pasta and breadmix, while they are very low for muesli: this can be due to the fat layer in the sample extract, in fact even using the supernatant under the fat layer, the solution was not clear, thus interfering with the assay. Finally, for the 20 ppm, really low results were obtained due to a poor kit efficiency (only more concentrated samples and standard developed a coloration).

The highest spiking level (200 ppm) is clearly detected by all the kits. This gluten concentration is above the upper quantification limit for kit 1. Also in this case the competitive assay (kit 2) gives underestimated values compared to the sandwich type ELISA kits. Results lower than the actual 200 ppm are obtained with kit 3. Results from kits 4 and 5 are very good for breadmix and pasta, but very low for muesli, due to a matrix effect. Finally, kit 6 gives not reliable data (insufficient colour development).

Table 7.7 Results obtained with the seven different ELISA kit for gluten detection. All results below LOD were considered negative, "detected" means that the result is between LOD and LOQ (detectable but not quantifiable).

		KIT 0		KIT 1		KIT 2		KIT 3		KIT 4		KIT 5		KIT 6	
matrix	***level gluten spiked***	*AVG*	*STD DEV*	*AVG*	*STD DEV*	*AVG*	*STD DEV*	*AVG*	*STD DEV*	*AVG*	*STD DEV*	*AVG*	*STD DEV*	*AVG*	*STD DEV*
bread mix	0 ppm	<LOD	NA	5.58	1.40	<LOD	NA	<LOD	NA	0.46	0.02	<LOD	NA	5.53	0.07
pasta		<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	5.34	0.07
muesli		<LOD	NA	5.78	1.31	<LOD	NA	<LOD	NA	0.39	0.10	<LOD	NA	5.24	0.07
bread mix	0.2 ppm	<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	1.20	1.14	<LOD	NA	DETECTED	NA
pasta		<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	5.86	1.63
muesli		<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	DETECTED	NA
bread mix	2 ppm	DETECTED	NA	5.01	0.40	<LOD	NA	DETECTED	NA	3.93	2.63	2.01	0.33	5.78	0.14
pasta		<LOD	NA	<LOD	NA	<LOD	NA	<LOD	NA	0.41	0.12	0.73	0.08	5.10	0.14
muesli		<LOD	NA	<LOD	NA	<LOD	NA	DETECTED	NA	0.60	0.19	<LOD	NA	5.58	0.27
bread mix	20 ppm	15.04	0.51	21.47	18.35	6.52	0.08	11.13	2.47	13.65	6.96	16.78	1.30	5.53	0.07
pasta		13.82	NA	12.32	NA	6.31	0.13	8.28	4.88	14.77	2.07	19.12	4.53	5.49	0.14
muesli		DETECTED	NA	14.50	2.63	6.15	0.27	10.43	1.98	7.02	1.12	3.00	0.01	5.34	0.21
bread mix	200 ppm	89.63	7.20	135.33	NA	69.83	3.64	105.65	38.17	198.86	23.62	215.15	29.20	270.83	44.16
pasta		82.27	1.62	297.37	126.29	63.61	0.55	94.23	20.60	141.55	3.25	221.52	7.09	52.92	2.75
muesli		81.85	5.05	226.17	17.92	47.44	3.04	87.90	2.41	27.98	2.73	50.56	17.46	54.86	1.37

The validation study performed highlighted some strength and weakness of the ELISA kits for gluten detection present on the market. First of all, a number of false positive was observed at all the three lower spiking levels (0, 0.2 and 2 ppm), but in all the cases the amount of gluten detected was very low (<6 ppm), so anyway largely

below the legal threshold of 20 ppm. For the higher levels (20 and 200 ppm), results showed a very high variability from kit to kit. For the 20 ppm it can be noted a general underestimation of gluten content for most of the cases, that can be a real problem in order to verify the correct labelling of food products, according to the EC No 41/2009. For the 200 ppm instead some kit were clearly underestimating while some other gave too high results, confirming the interkit variability. That means that the estimated gluten content can vary using different kits, thus lending itself to uncertainties in the field of food safety control. So, indications about the official kit to be used should be provided by the legislation authorities in order to guide food safety monitoring. A couple of kit on the market have already received the approval by the AOAC. But, even using the same kit, strongly different results can be obtained in different matrices, and it can be seen that the underestimation of gluten content is directly proportional to the complexity of the food matrix (higher content in sugars and fats). This is a serious issue for gluten detection in processed foods, in which it could be underestimated or even not detected while still harmful for celiac patients.

7.3.6 MS/MS method development for gluten detection

Since, as previously seen, PCR methods cannot be always applied to processed foods because technological treatments, together with food formulation, can lead to low DNA presence and poor extractability and ELISA methods are affected by matrix type as well, we decided to develop a tandem mass spectrometry methods for gluten detection using marker peptides generated from enzymatic digestion. We have already demonstrated in a previous work (see Chapter 3) that simplified digestion models leads to a strong underestimation of gluten content in thermally treated foods, such as cooked pasta. So, for the proteolytic cleavage of gluten proteins into shorter peptides we used the method developed by Versantvoort et al.[15] For gluten quantification the more abundant peptides were chosen, independently from their immunotoxicity, because the aim of the experiment was gluten detection and not immunotoxic peptides quantification. In this way it has been possible to use peptides with a better response factor, thus increasing the sensitivity of the method. So, on the basis of the peptides previously identified, we screened all the areas of the peptides in different wheat samples that had undergone simulated gastrointestinal digestion. The five more abundant peptides identified were QQQPL, QQPPFS and PQQPPFS among those not containing immunotoxic sequences and TQQPQQPFPQQPQQPFPQ among those containing immunotoxic sequences. An important requirement for a marker peptide is its specificity for the target protein. So, the candidate marker peptides were aligned in the protein database to check their specificity for gluten proteins. Peptide QQQPL was too short to achieve a correct alignment, so it was discarded despite being the most abundant one. The other three peptides were present in all the Triticeae tribe, that comprises all the species toxic for celiac patients, such as wheat, barley and rye, but not oats (*Avena sativa*). Once the candidate peptides were selected, Daughter Scan experiments were performed to identify the characteristic to be used for the quantification and confirmation transitions in the subsequent MRM experiment. Transitions fragments monitored are reported in Table 7.8. Collision energies were optimized in order to achieve the maximum intensity of the Daughter Ion used for the quantification.

Table 7.8 Transitions of the marker peptides monitored for gluten quantification; the first daughter ion is for quantification and the last two for confirmation.

PEPTIDE	Rt (min)	PARENT ION (m/z)	DAUGHTER IONS (m/z)	COLLISION ENERGY
QQPPFS	20.67	703.6 $[M+H^+]^+$	447.2 (Q) 226.5 (C) 350.3 (C)	27
TQQPQQPFPQQPQQPFPQ	25.63	717.6 $[M+3H^+]^{3+}$	244.0 (Q) 354.0 (C) 468.4 (C)	28
PQQPPFS	21.23	800.6 $[M+H^+]^+$	447.2 (Q) 226.5 (C) 350.3 (C)	31

In a previous study the development of MRM methods for gluten quantification was described, but there the result was expressed in term of immunogenic peptide amount. In this study, our aim was to obtain a result that expresses the amount of gluten present in food samples. To achieve this purpose, three different gluten-free matrices were used: bread mix, pasta and muesli. So, 2.25 g of blank matrix were spiked with 23 mg of pure wheat gluten in order to obtain a 10000 ppm standard solution. The standard was then digested and serial dilutions were made with the respective blank matrices undergone digestion, in order to obtain a matrix calibration curve. The calibration curves were made in the range 0-300 ppm with the following points: 0, 5, 10, 20, 50, 100, 200 and 300 ppm (Figure 7.11, Figure 7.12 and Figure 7.13). As it can be seen from the slopes of the calibration curves, the matrix effect is generally increasing from breadmix, to pasta, to muesli (that is the more complex matrix, with a lot of fats and sugars in the formulation).

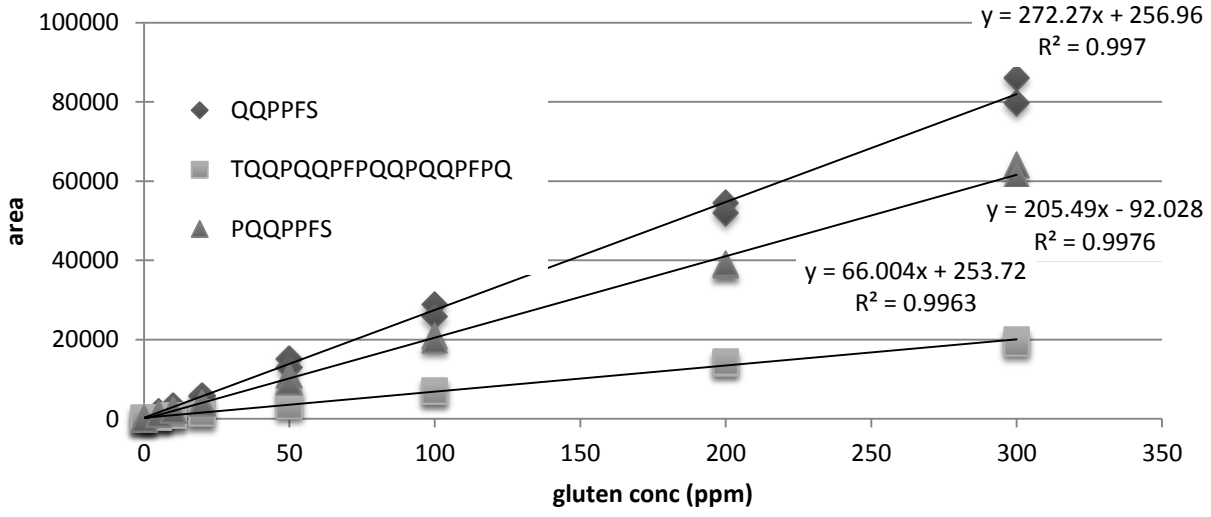

Figure 7.11 Calibration curves obtained for the three marker peptides in breadmix matrix.

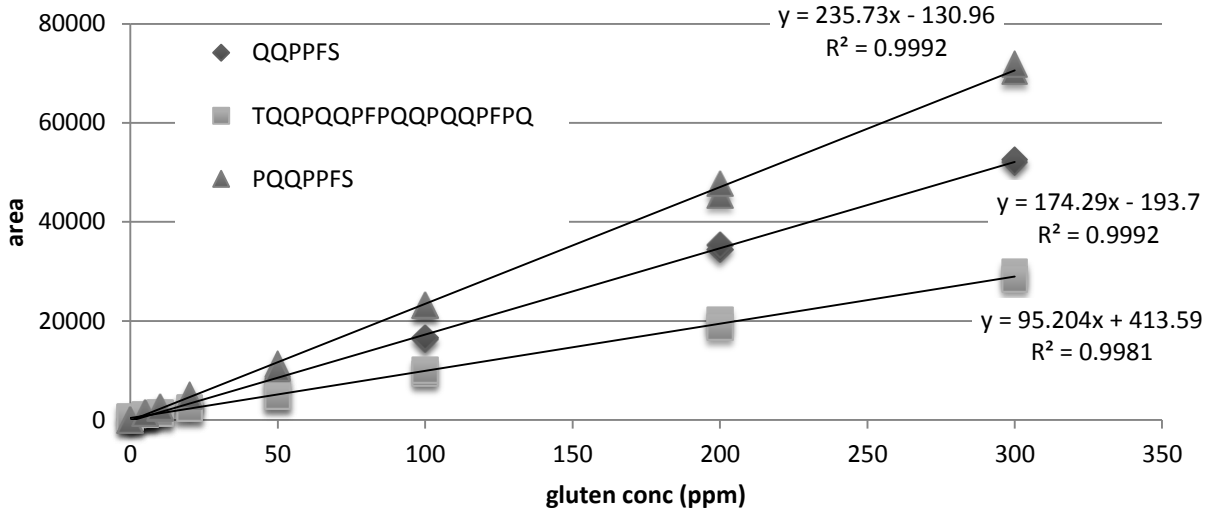

Figure 7.12 Calibration curves obtained for the three marker peptides in pasta matrix.

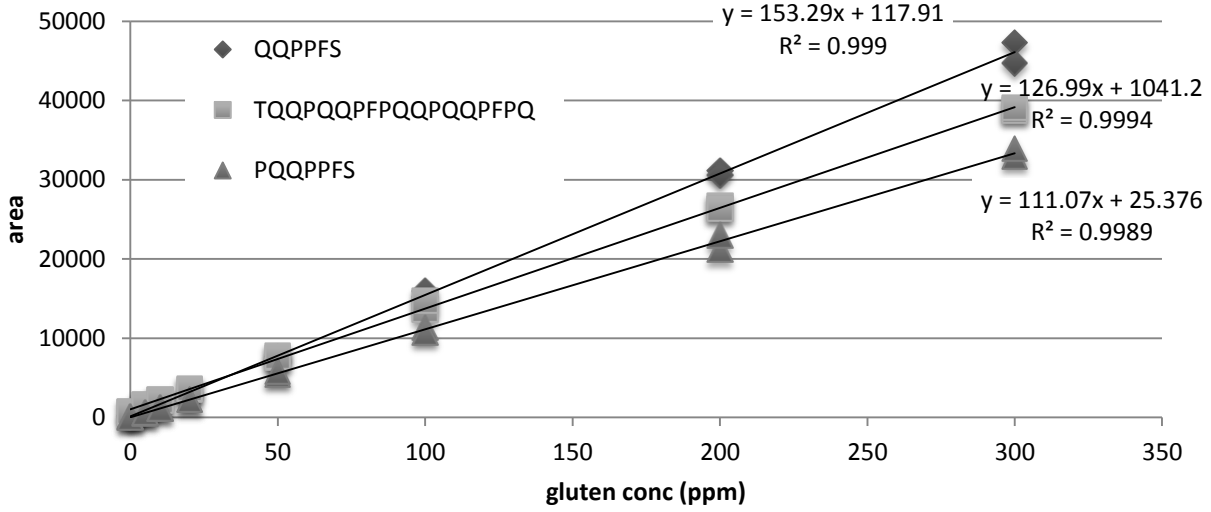

Figure 7.13 Calibration curves obtained for the three marker peptides in muesli matrix.

So, the same spiked samples used for the validation study were submitted to simulated gastrointestinal digestion and analyzed with the MRM method developed. Results are reported in Table 7.9. Generally, peptides QQPPFS and PQQPPFS seem to perform much better than TQQPQQPFPQQPQQPFPQ. The 0 ppm resulted negative for all the three matrices, except a very low amount (around 1 ppm) detected in muesli. The 0.2 ppm was detected only in muesli, but the amount is similar to that of the 0 ppm, so probably it is just a case of false positive. The 2 ppm can be detected only with peptide QQPPFS and PQQPPFS. The matrix effect is particularly evident in the bread mix and pasta, in which only one of the two replicated gave a signal. On the other side, muesli seems to be the matrix in which gluten is better detected, giving higher responses for the 20 and the 200 ppm.

Table 7.9 Gluten concentration determined with MRM method

spiked gluten concentration (ppm)	QQPPFS		TQQPQQPFPQQPQQPFPQ		PQQPPFS	
BREAD MIX	*AVG*	*STD DEV*	*AVG*	*STD DEV*	*AVG*	*STD DEV*
0	ND	ND	ND	ND	ND	ND
0.2	ND	ND	ND	ND	ND	ND
2	0.90	-	ND	ND	1.99	1.06
20	5.31	0.50	ND	ND	12.39	2.88
200	75.09	22.89	19.48	6.65	138.79	38.52
PASTA	*AVG*	*STD DEV*	*AVG*	*STD DEV*	*AVG*	*STD DEV*
0	ND	ND	ND	ND	ND	ND
0.2	ND	ND	ND	ND	ND	ND
2	3.11	-	ND	ND	3.65	-
20	9.89	2.34	6.99	1.26	9.93	2.15
200	100.49	29.30	95.68	30.69	141.88	44.39
MUESLI	*AVG*	*STD DEV*	*AVG*	*STD DEV*	*AVG*	*STD DEV*
0	0.77	0.54	ND	ND	1.05	0.10
0.2	0.86	0.76	ND	ND	1.30	0.23
2	1.26	0.31	ND	ND	4.22	0.32
20	25.45	19.64	10.22	8.55	64.28	48.63
200	243.83	32.33	85.25	13.68	599.79	91.47

The same results, compared with the ELISA kits tested, are reported in Table 7.10. Among the three peptide tested, PQQPPFS seems to be the most accurate one. It is clear that, both for MRM as well as ELISA detection methods, there is a general underestimation of gluten content. This is probably due to a poor gluten extractability in the case of ELISA and to a strong matrix effect for MS/MS detection.

Table 7.10 Comparison among different ELISA kits and three different MRM method for gluten detection in three gluten free matrices spiked with known amount of gluten.

matrix	*gluten spiked*	KIT 0	KIT 1	KIT 2	KIT 3	KIT 4	KIT 5	KIT 6	MS/MS 1	MS/MS 2	MS/MS 3
bread mix	2 ppm	POS	5.01	<LOD	POS	3.9	2.0	5.8	0.9	<LOD	2.0
pasta		<LOD	<LOD	<LOD	<LOD	0.4	0.7	5.1	3.1	ND	3.7
muesli		<LOD	<LOD	<LOD	POS	0.6	<LOD	5.6	1.3	ND	4.2
bread mix	20 ppm	15.0	21.5	6.5	11.1	13.6	16.8	5.5	5.3	ND	12.4
pasta		13.8	12.3	6.3	8.3	14.8	19.1	5.5	9.9	7.0	9.9
muesli		POS	14.5	6.2	10.4	7.0	3.0	5.3	25.5	10.2	64.3
bread mix	200 ppm	89.6	135.3	69.8	105.7	198.9	215.2	270.8	75.1	19.5	138.8
pasta		82.3	297.4	63.6	94.2	141.6	221.5	52.9	100.5	95.7	141.9
muesli		81.8	226.2	47.4	87.9	28.0	50.6	54.9	243.8	85.3	599.8

Finally, different samples were analyzed for the gluten content using the newly developed MRM method as well as a commercially available ELISA kit. All tested samples should be gluten-free, except for the barley. Actually, only the vol au vent, the soy and the corn gluten pellets resulted below the legal limit for declaring a food as gluten free (20 ppm). Cheese croquettes and oats were slightly contaminated by gluten, and the detection was comparable among the four results. Chicken nuggets were strongly contaminated and gluten was clearly detected. Corn gluten pellets resulted strongly positive with the ELISA assay and with the second peptide, but the quantitative RT-PCR performed previously (see Chapter 7.2.5) confirms the value obtained with peptide 1 and 3. Finally, from these results emerged that the better peptide for the quantification of barley contamination is PQQPPFS, that is the one giving the highest response.

Table 7.11 Gluten content (in ppm) determined using the three marker peptides and a commercially available ELISA kit.

	QQPPFS	TQQPQQPFPQQPQQPFPQ	PQQPPFS	ELISA
Vol au vent	0.7	1.0	0.2	0.9
Chicken nuggets	538.0	451.6	222.1	466.6
Cheese croquettes	35.5	110.3	29.1	68.6
Brazil soybeans	0.8	5.4	2.7	10.1
Corn gluten pellets	14.7	59.2	7.8	332.2
Oats	103.4	115.4	71.3	81.9
French brewer's barley	84.6	396.5	437.7	33292.0

7.4 Conclusions

A LC/MS method ahs been developed for detecting marker peptides for gluten content in order to measure the amount of gluten in different food materials. The method has been compared to ELISA and PCR detection. The high variability of gluten proteins has clearly emerged, as shown in Chapters 2, 3 and 4, the peptidic profile generated after in vitro digestion of several wheat varieties can be strongly different. So, depending on the gluten composition that gives rise to contamination, different results (in term of gluten concentration) can be obtained using three different peptides as markers in LC/MS. But even the same gluten composition, when used for validation experiments, can lead to different result when spiked in different matrices, as happened for bread mix, pasta and muesli. This is probably due to the influence that the matrix has on the outcome of the digestion, leading to a higher or lower digestibility of the gluten spiked: not surprisingly the matrix that gives the more different result is the more complex one (muesli). Peptide 2 (TQQPQQPFPQQPQQPFPQ) is the one that most suffers of matrix effect, as it can be also observed from the calibration curves. The use of isotopically labelled standards for peptides quantification could be useful to achieve a more accurate result in the LC-MS/MS analysis, but the problem of how to convert the peptide content in gluten content still remains. So, maybe in this case it would be a good approach to quantify different marker peptides and, in the view of consumer protection in the worst case, keep in consideration the highest value.

7.5 ACKNOWLEDGEMENT

ILVO T&V (Instituut voor Landbouw- en Visserijonderzoek, Technologie en Voeding) is gratefully acknowledged for hosting the experiments and made available materials and skills.

---

[1] Janatuinen EK, Pikkarainen PH, Kemppainen TA, Kosma VM, Järvinen RMK, Uusitupa MIJ, Julkunen RJK (1995) A Comparison of Diets with and without Oats in Adults with Celiac Disease. *N Engl J Med* 333:1033-1037

[2] Arentz-Hansen H, Fleckenstein B, Molberg Ø, Scott H, Koning F, Jung G, Roepstorff P, Lundin KEA, Sollid LM. The Molecular Basis for Oat Intolerance in Patients with Celiac Disease. *PLoS Med* 1(1): e1.

[3] Janatuinen EK, Kemppainen TA, Julkunen RJK, Kosma3 VM, Mäki M, Heikkinen M, UusitupaMIJ (2002) No harm from five year ingestion of oats in coeliac disease. Gut 50: 332-335

[4] Lundin KEA, Nilsen EM, Scott HG, Løberg EM, Gjøen A, Bratlie J, Skar V, Mendez E, Løvik A, Kett K (2003) Oats induced villous atrophy in coeliac disease. Gut 52: 1649-1652

[5] Collin P, Thorell L, Kaukinen K, Mäki M (2004) The safe threshold for gluten contamination in gluten-free products. Can trace amounts be accepted in the treatment of coeliac disease? *Aliment Pharmacol Ther* 19: 1277–1283

[6] Li X, Wang X (2000) Application of real-time polymerase chain reaction for the quantitation of interleukin-1β mRNA upregulation in brain ischemic tolerance. Brain Res Protoc 5(2):211-217.

[7] http://biosiva.50webs.org/agab.htm

[8] Hewel JA, Phanse S, Liu J, Bousette N, Gramolini A, Emili A (2013) Targeted protein identification, quantification and reporting for high-resolution nanoflow targeted peptide monitoring. *Journal of Proteomics* 81:159-172.

[9] Simonato B, Mainenete F, Tolin S, Pasini G (2011) Immunochemical and mass spectrometry detection of residual proteins in gluten fined red wine. *J Agric Food Chem* 59, 3101–3110

[10] Sealey-Voyksner JA, Khosla C, Voyksner RD, Jorgenson JW (2010) Novel aspects of quantitation of immunogenic wheat gluten peptides by liquid chromatography–mass spectrometry/mass spectrometry. *J Chromatogr A* 1217: 4167–4183

[11] Piknovà L, Brežnà B, Kuchta T (2008) Detection of gluten-containing cereals in food by 5'-nuclease real-time polymerase chain reaction. *J Food Nutr Res* 47(3): 114–119

[12] Directive 2007/68/EC.

[13] Kahlenberg F, Sanchez D, Lachmann I, Tuckova L, Tlaskalova H, Méndez E, Mothes T (2006) Monoclonal antibody R5 for detection of putatively coeliac-toxic gliadin peptides. *Eur Food Res Technol* 222:78-82

[14] Van Eckert R, Berghofer E, Ciclitira PJ, Chirdo F, Denery-Papini S, Ellis HJ, Ferranti P, Goodwin P, Immer U, Mamone G, Méndez E, Mothes T, Novalin S, Osman A, Rumbo M, Stern M, Thorell L, Whim A, Wieser H (2006) Towards a new gliadin reference material–isolation and characterisation. *J Cer Sci* 43(3), 331–341

[15] Versantvoort CHM, Oomen AG, Van de Kamp E, Rompelberg CJM, Sips AJAM (2005) Applicability of an in vitro digestion model in assessing the bioaccessibility of mycotoxins from food. *Food Chem Toxicol* 43, 31-40

8 COMMON WHEAT DETERMINATION IN DURUM WHEAT SAMPLES THROUGH LC/MS ANALYSIS OF GLUTEN PEPTIDES

Based on: Prandi B., Bencivenni M., Tedeschi T., Marchelli R., Dossena A., Galaverna G., Sforza S. (2012) Common wheat determination in durum wheat samples through LC/MS analysis of gluten peptides. Analytical and Bioanalytical Chemistry 403(10): 2909-14.

8.1 INTRODUCTION

Cereal are the most widely harvested crops in the world and their products have always been basic foods for human nutrition. Wheat and rice are the most important cereals for human consumption and in the Mediterranean area wheat consumption is particularly high. As far as wheat is concerned, the most diffuse species are durum wheat (*Triticum turgidum* spp *durum*), whose typical intended use is the pasta industry, and common wheat (*Triticum aestivum*), usually employed to make bread or other baked goods. In Italy, dried pasta must be made exclusively of durum wheat, and a maximum contamination of 3% of common wheat flour in durum wheat flour is allowed.[1] As a matter of fact, pasta dough made from durum wheat flour has rheological properties ideally suited to the pasta manufacturing process: when cooked, durum wheat pasta resists disintegration and retains a firm texture.[2] However, the Italian law also allows import-export of pasta totally or partially prepared using *T. aestivum*, using a clear indication on the label. Since durum wheat price is quite 25% higher than common wheat,[3] useful tools for the detection of *T. durum* adulteration with *T. aestivum* are required. The official analysis method is based on the extraction of soluble proteins from wheat and their electrophoretic separation by isoelectric focusing on polyacrylammide gel, since the band patterns are different between common and durum wheat. In this way it is possible to recognize the two species either alone or mixed together.[4] However the quantitative evaluation is based on the comparison of band intensities between the sample and the standard and thus is not very accurate. Moreover, the method does not allow a correct evaluation of the common wheat content for pasta dried at high temperature.

Many studies are present in literature to overcome these problems, most of them taking advantage of the different ploidy level of common (AABBDD) and durum wheat (AABB). The amplification with end-point PCR of DNA sequences belonging to DD genome, using appropriate primers, is indication of common wheat presence;[5] this presence can also be quantify using real-time PCR.[6] The problem of DNA degradation during technological processing was overcome amplifying repeated DNA sequences giving short amplicons, called microsatellites.[7]

This different genomic structure also affects protein expression. Differences in protein composition between common and durum wheat have been known for a long time: the different electrophoretic pattern of chloroform/methanol soluble proteins was used to detect and define the range of adulteration of *T. durum* with *T. aestivum*.[8] A similar electrophoretic method can be used also for pasta, extracting and separating the prolamine[9] or the albumin fraction.[10] With the development of high performance liquid chromatography and free zone capillary electrophoresis, new methods for the detection of common wheat were generated. The extraction and subsequent HPLC analysis of gliadins from wheat can be used for the individuation of adulteration with common wheat,[11] also when pasta is dried at high temperature.[12] Both in the chromatogram and in the electropherogram

of proteic extracts from common wheat are present some peaks not detectable in durum wheat; in this way it was possible to make a calibration curve using samples of known composition.[13]

Not only protein analysis were study to detect *T. durum* adulterations, but also analysis of lipid fraction. In fact, the analysis of sterol-palmitate content can be used to detect adulteration in wheat and pasta dried at different temperature.[14] Alkylresorcinols composition of common wheat is different from durum wheat, so the C17:C21 ratio can be used to estimate such adulterations; unfortunately alkylresorcinols are present only in the hyaline layer, inner pericarp and testa of cereal grains, limiting the analysis to whole grain products.[15] Unfortunately, many of the method based on metabolites analysis are affected from the variability due to different cultivar.

In the present work a method for the identification and quantification of durum wheat adulteration with common wheat was developed, using a marker peptide generated from the enzymatic treatment of grinded kernels or flours.

8.2 MATERIALS AND METHODS

8.2.1 Reagents and solvents

Deionised water was obtained from a Millipore Alpha Q-Waters purification system. Pepsin from porcine gastric mucosa, α-Chymotrypsin from bovine pancreas, formic acid (≥99,9%), sodium dihydrogen phosphate and acetonitrile were purchased from Sigma-Aldrich (Stockholm, Sweden). Hydrochloric acid (37% V/V) was purchased from Carlo Erba (Milan, Italy).

8.2.2 Samples

All wheat varieties were provided by Società Produttori Sementi SpA (Bologna, Italy). Common and durum wheat used for the calibration curve were respectively of Cerere and Levante varieties. The calibration curve was done in triplicate for eight points, corresponding to the following percentage of common wheat: 0, 5, 10, 20, 40, 60, 80 and 100. Some mixed samples of known composition were prepared by using different durum (Levante and Svevo) and common (Cerere, C172, C173, C181) wheat varieties, and used for the determination of common wheat in blind experiments. Commercial samples of durum wheat flour were purchased in the market.

8.2.3 Enzymatic cleavage

Wheat kernel were grinded to obtain a fine granulometry; common and durum flours were blended at predefined percentages and accurately mixed. 100 mg of each sample were weighted in a 15 ml tube and added of 4 ml of 10 mM HCl (pH 2). After the addition of 200 µl of a 1 mg/ml pepsin solution, the tube was shaken and left stirring for 3 hours at 37°C. Then, the sample was added with 4 ml of phosphate buffer (100 mM, pH 7.2) and 200 µl of a 1 mg/ml chymotrypsin solution; the tube was shaken and left stirring for 4 hours at 37°C. After the enzymatic cleavage, 1 ml of acetonitrile with 0.1% formic acid was added to stop the reaction and the tubes were centrifuged at 6226 g (4°C) for 20 minutes. The supernatant was collected in a flask and dried under vacuum, then the samples were reconstituted with 1 ml formic acid 0.1% and transferred into a 1 ml tube. After drying under nitrogen flux, the samples were reconstituted with 500 µl of 0.1% formic acid and centrifuged at 16602 g (at 4°C) for 10 minutes to obtain a clear solution. The supernatant was then analyzed by UPLC/ESI-MS.

8.2.4 UPLC/ESI-MS analysis

The complex mixtures obtained from enzymatic cleavage were separated by a RP column (ACQUITY UPLC BEH 300 C18 1,7 µm 2,1*150 mm) in an UPLC/ESI-MS system (UPLC Acquity Waters with a single quadrupole mass spectrometer Waters Acquity Ultraperformance) using a gradient elution. Eluent A was water with 0.1% formic acid and 0.2% acetonitrile, eluent B was acetonitrile with 0.1% formic acid; gradient: 0÷7 min 100% A, 7÷50 min from 100% A to 50% A, 50÷52.6 min 50% A, 52.6÷53 min from 50% A to 0% A, 53÷58.2 min 0% A, 58.2÷59 min from 0% A to 100% A, 59÷72 min 100% A. The samples were analyzed by the mass analyzer the Full Scan mode. Flow 0.2 ml/min; analysis time 72 min; column temperature 35°C; sample temperature 6°C; injection volume 5 µl; acquisition time 7÷58.2 min; ionization type positive ions; scan range 100÷2000 m/z; capillary voltage 3.2 kV; cone voltage 30 V; source temperature 100°C; desolvation temperature 200°C; cone gas flow 100 l/h; desolvation gas flow 650 l/h.

To calculate the common wheat percentage, the characteristic ions of wheat (durum + common) marker peptide ($[M+4H^+]4^+$=684.9; $[M+3H^+]3^+$=913 and $[M+2H^+]2^+$=1368.9) and of common wheat marker peptide ($[M+4H^+]^{4+}$=978.5 and $[M+3H^+]^{3+}$=1304.3) were extracted (obtaining eXtract Ion Chromatograms, XICs) and integrated with the MassLynx software.

8.3 RESULTS AND DISCUSSION

8.3.1 Identification of marker peptides

The lacking of DD genome in durum wheat implies a different protein expression between *Triticum aestivum* and *Triticum turgidum* spp *durum*, which also affects the peptides generated by the enzymatic cleavage. As a matter of fact, in common wheat samples treated with pepsin and chymotrypsin, a peptide having a molecular mass of 3909 Da was identified and the same peptide was absent in durum wheat samples equally treated (Figure 8.1).

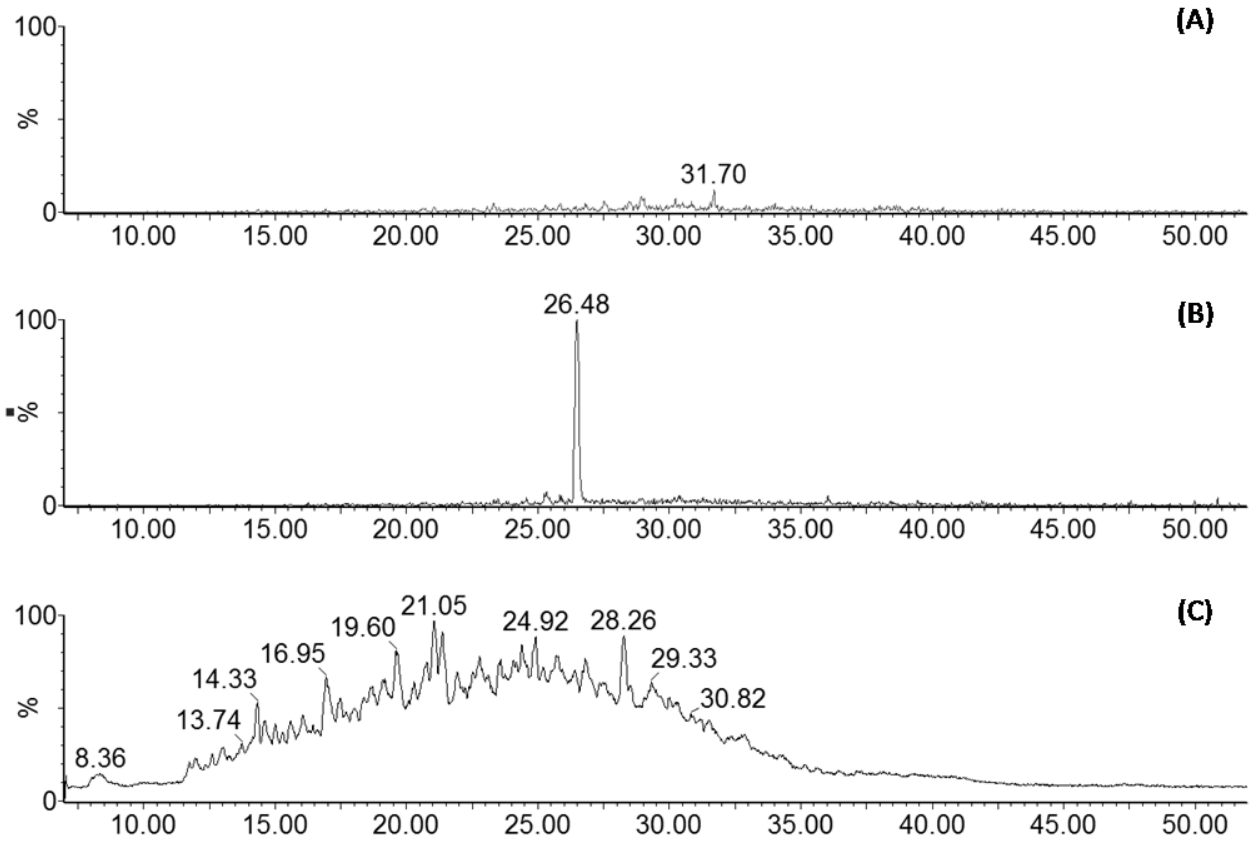

Figure 8.1 EXtract ions chromatograms of m/z 978.5 $[M+4H^+]^{4+}$ and 1304.3 $[M+3H^+]^{3+}$, belonging to the peptide marker for common wheat presence LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF, are reported for a durum wheat sample (A) and a common wheat sample (B); chromatograms are scaled to the same intensity. In (C) is reported a Full Scan chromatogram of a wheat sample.

The mass spectrum obtained from the fragmentation of the ion with m/z=1303 ($[M+3H^+]^{3+}$ of this marker peptide for common wheat presence) agreed with the aminoacidic sequence LQLQPFPQPQLPYPQPQLPYPQPQPF. The annotated spectrum is shown in Figure 8.2.

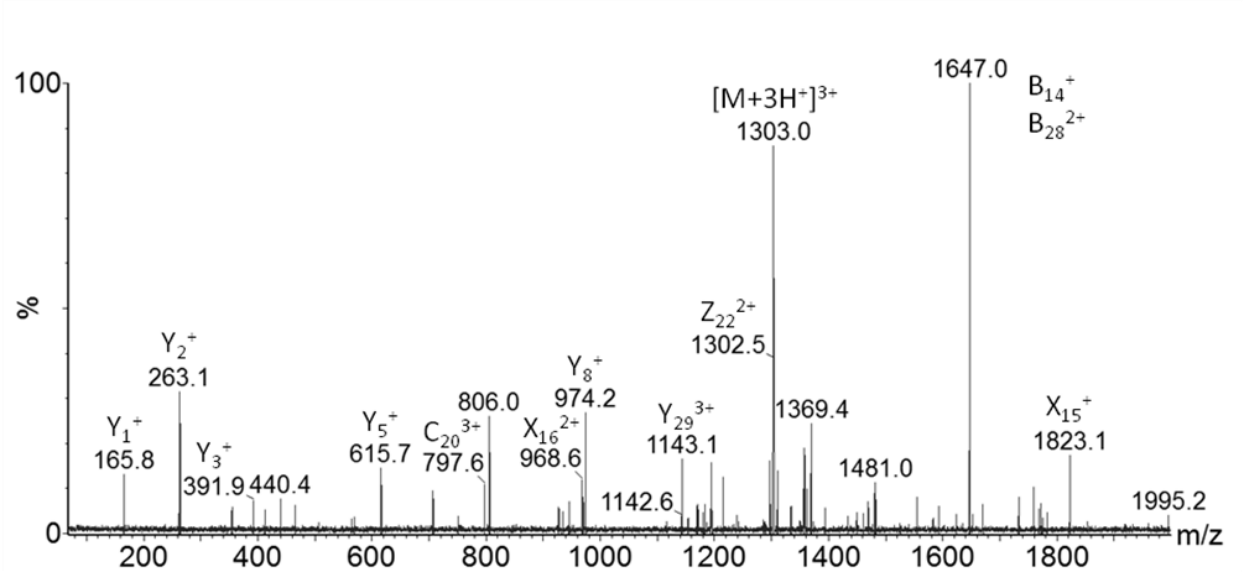

Figure 8.2 Mass spectrum obtained from the fragmentation of the ion with m/z=1303, used to identify the peptide marker for common wheat presence.

A comparison in Uniprot Data Bank outlined that this sequence was present only in six α-gliadin sequences, all belonging to the species *Triticum aestivum* or to the hybrid *Thinopyrum ponticum* x *Triticum aestivum*. Coding genes for some of the protein containing this sequence could also be identified: the gene Gli-Z1 (encoding for the protein Q1WA39) and the gene Gli-G3 (encoding for the protein A5JSA6) are located on chromosome 6D,[16] absent in durum wheat genome. Also the gene Gli-H59 (encoding for the protein A7LHC8) is located on genome D.[17, 18] In general, sequences equivalent to LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF were not found to be present in the α-gliadin loci of the A or B chromosomes (*Gli-2A* or *Gli-2B*) but only in *Gli-2D* on the short arm of chromosome 6D.[19] The absence of this peptide in the enzymatically treated durum wheat extracts is in agreement with these data.

In order to use this peptide as a marker for common wheat, another marker peptide accounting for the total (durum + common) wheat content was chosen: the peptide having molecular mass of 2735 Da was found to be one of the most abundant in all wheat species, so it was taken as a reference (Figure 8.3).

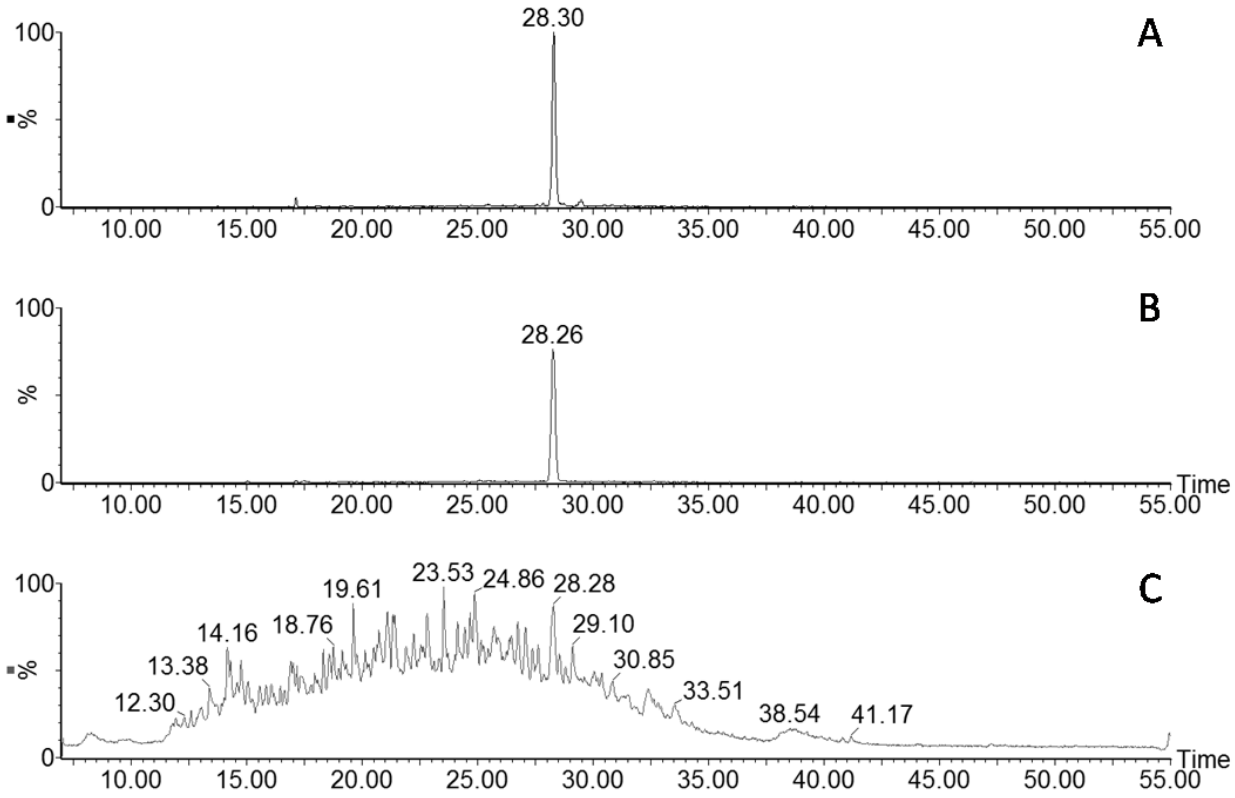

Figure 8.3 EXtract ions chromatograms of m/z 684.9 $[M+4H^+]^{4+}$, 913 $[M+3H^+]^{3+}$ and 1368.9 $[M+2H^+]^{2+}$, belonging to the peptide marker for total wheat presence VRVPVPQLQPQNPSQQQPQEQVPL, are reported for a durum wheat sample (A) and a common wheat sample (B); chromatograms are scaled to the same intensity. In (C) is reported a Full Scan chromatogram of a wheat sample.

Figure 8.4 shows the mass spectrum obtained from the fragmentation of the ion with m/z=1368 ($[M+2H^+]^{2+}$ of the peptide used as a marker for the total wheat content), which was in agreement with the theoretical fragmentation of the peptide VRVPVPQLQPQNPSQQQPQEQVPL. This peptide is the N-terminal sequence of all α-gliadins, and was very abundant in all our samples, so it was taken as an internal reference.

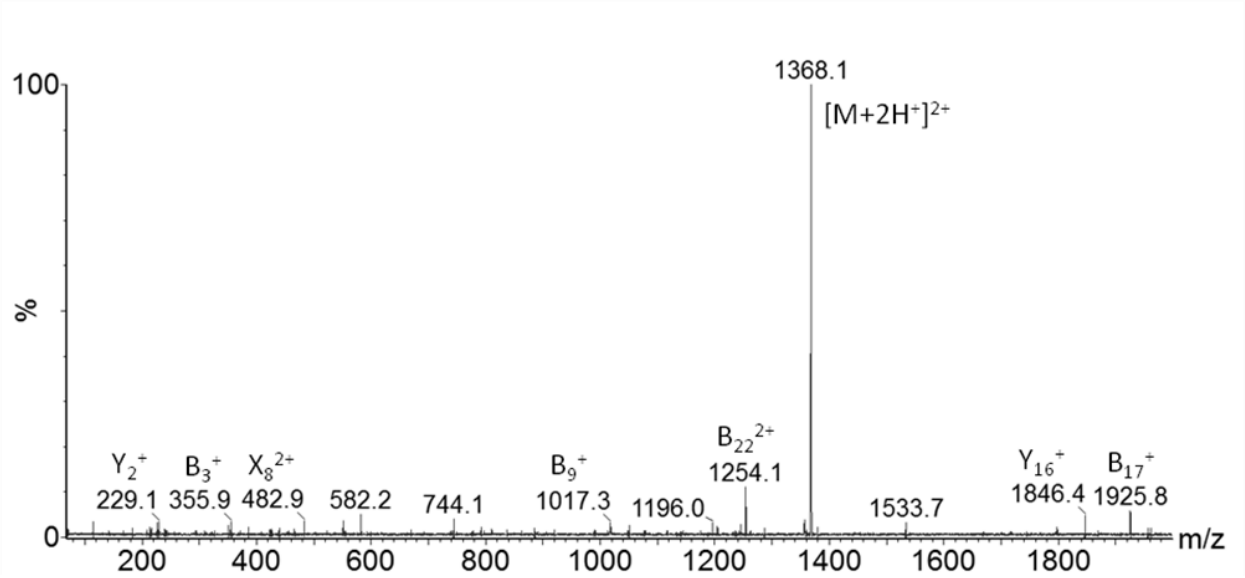

Figure 8.4 Mass spectrum obtained from the fragmentation of the ion with m/z=1368, used to identify the peptide marker for total wheat amount.

8.3.2 Calibration curve

The calibration curve was made mixing common and durum wheat grinded kernels in eight different percentages: 0, 5, 10, 20, 40, 60, 80 and 100% common wheat. 0.1 g of each sample was processed in triplicate as described in the experimental session and analyzed by UPLC/ESI-MS. The peak due to marker peptides were integrated in the Extract Ion Chromatogram, and the ratio between the marker peptide for common wheat and the marker peptide for total wheat was plotted against common wheat proportion.

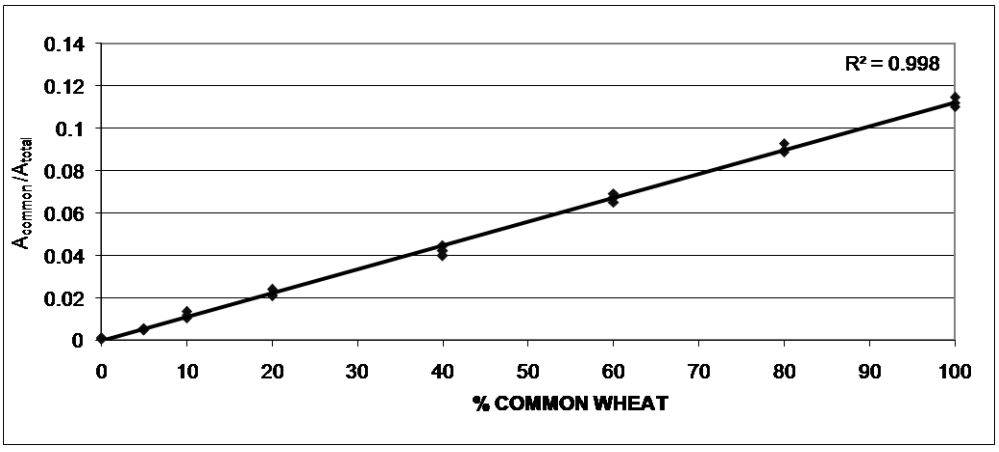

Figure 8.5 Calibration curve for the determination of common wheat proportion: on the Y-axis is reported the ratio between the area of the peptide marker for common wheat presence and the area of the peptide marker for total wheat presence, plotted against the common wheat percentage.

The calibration curve obtained (Figure 8.5) can be used to get an assessment of the proportion of common wheat on total wheat content, using the equation obtained with a linear regression. When the common wheat percentage calculated was greater than 30%, for a more accurate determination only the last four point of the

calibration curve were to be used. The variation of the linear regression equation at high common wheat percentages is probably due to a loss of linearity caused by a too wide range of concentrations.

8.3.3 Determination of common and durum wheat composition in mixed samples from different cultivars

Since protein content of wheat is dependent upon the cultivar and the agronomic conditions, some samples were prepared mixing different common and durum wheat varieties, in order to test the method against interspecific variability (Table 8.1).

Table 8.1 Composition of the samples used to test the interspecific variability and results obtained from the quantitation.

SAMPLE	DURUM WHEAT VARIETY	COMMON WHEAT VARIETY	ACTUAL COMMON WHEAT CONTENT (%)	DETECTED COMMON WHEAT CONTENT (%)
1	Svevo		0	0.2±0.1
2	Levante	C173	5	6.4±1.6
3	Levante	Cerere	20	15.5±0.8
4	Levante	Cerere	56	55.2±8.7
5	Svevo	C181	70	63.9±3.1
6	Levante	C172	75	82.2±1.7
7	Levante	Cerere	90	86.8±3.7
8		C172	100	97.7±1.9

Svevo, Levante and Cerere are wheat varieties widely harvested, instead C172, C173 and C181 are experimental varieties cross-breeded to have a lower α-gliadin content.

Eight different mixtures were prepared, extracted and enzimatically treated as reported before, analyzed by UPLC/ESI-MS and finally blindly assessed using the calibration curve reported in Figure 8.5.

The results shown in Table 8.1 demonstrate the independence of the method from the varieties used for sample preparation, and the good accuracy of the calculated common wheat proportions.

8.3.4 Testing the method with commercial samples

Using the developed method, a screening of the commercial durum wheat flour brands was done. As shown in Figure 8.6, only two samples showed a common wheat content higher than the Italian legal limit (3%), but in all samples the presence of common wheat-derived peptides was identified, indicating a diffuse contamination from common wheat of the samples of durum wheat found in the Italian market.

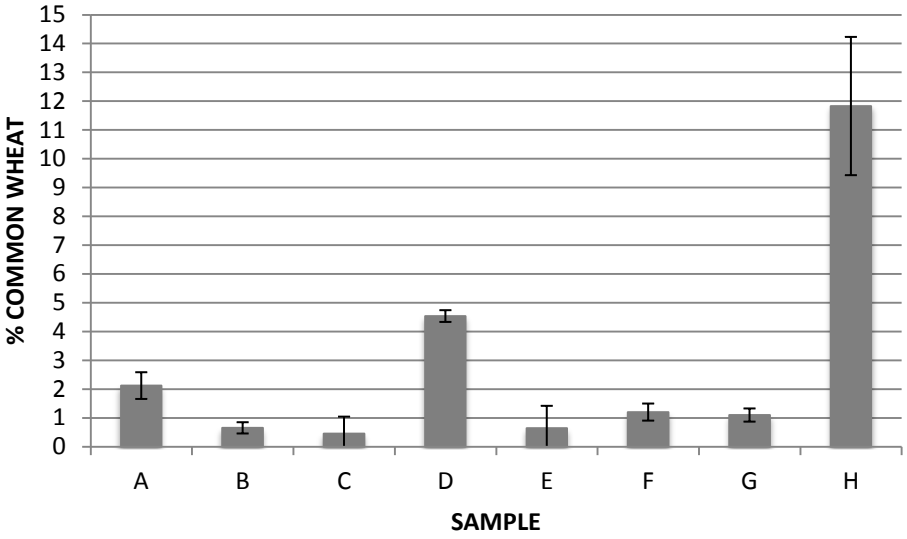

Figure 8.6 Results of the screening of the commercial samples of durum wheat flour.

8.4 CONCLUSIONS

In the present work, a peptide marker for common wheat presence was identified. The ratio of this peptide content against a marker peptide present in all wheat species was used, analyzing mixtures of pure common (Cerere) and durum (Levante) wheat varieties, for obtaining a calibration curve related to the common wheat percentage. This determination was showed not to be affected from the different varieties used for making mix samples, indicating that it was generally suitable for determining the common wheat content, showing results quite accurate for all the common wheat proportions tested. Finally, the method was applied for a market survey, showing a diffuse, albeit not very high, common wheat contamination in the commercially available durum wheat flour brands.

8.5 ACKNOWLEDGEMENTS

The AGER Foundation is gratefully acknowledged for its financial support to the project “From seed to pasta.”

---

[1] DPR 187/01 of 9th February 2001: Regulations for the review of legislation on production and marketing of flour and pasta.

[2] Dexter JE, Matsuo RR (1980) Relationship between durum wheat protein properties and pasta dough rheology and spaghetti cooking quality. *J Agric Food Chem* 28, 899-902

[3] Institute of Food Services for the Agricultural Market (ISMEA)

[4] Approval of official methods of analysis of grain. Recognition and measurement of common wheat flours in the durum wheat flour and pasta with isoelectric focusing. G. U. supplement n°2 to n°4 of 5/1/1980

[5] Arlorio M, Coïsson JD, Cereti E, Travaglia F, Capasso M, Martelli A (2003) Polymerase chain reaction (PCR) of puroindoline b and ribosomal/puroindoline b multiplex PCR for the detection of common wheat (Triticum aestivum) in italian pasta. *Eur Food Res Technol* 216, 253-258

[6] Alary R, Serin A, Duviau MP, Joudrier P, Gautier MF (2002) Quantification of common wheat adulteration using real-time quantitative polymerase chain reaction (PCR). *Cereal Chem* 79, 553-558

[7] Sonnante G, Montemurro C, Morgese A, Sabetta W, Blanco A, Pasqualone A (2009) DNA micro satellite region for a reliable quantification of soft wheat adulteration in durum wheat-based foodstuffs by real-time PCR. *J Agric Food Chem* 57, 10199-10204

[8] Garcìa-Olmedo F, Garcia-Faure R (1969) A new method for the estimation of common wheat (Triticum aestivum L.) in pasta products. *Lebensm Wiss uTechnol* 2, 94-96

[9] Burgoon AC, Ikeda HS, Tanner SN (1985) A method for detecting adulteration in durum wheat pasta by polyacrylamide gel electrophoresis. *Cereal Chem* 62, 72-74

[10] Feillet P, Kobrehel K (1974) Determination of common wheat content in pasta products. *Cereal Chem* 51, 203-208

[11] McCarthy PK, Scanlon BF, Lumley ID, Griffin M (1990) Detection and quantification of adulteration of durum wheat flour by flour from common wheat using reverse phase HPLC. *J Sci Food Agric* 50, 211-226

[12] Barnwell P, McCarthy PK, Lumley ID, Griffin M (1994) The use of high performance liquid chromatoghraphy to detect common wheat (Triticum aestivum) adulteration of durum wheat (Triticum durum) pasta products dried at low and high temperatures. *J Cereal Sci* 20, 245-252

[13] Bonetti A, Marotti I, Catizone P, Dinelli G, Maietti A, Tedeschi P, Brandolini, V (2004) Compared use of HPLC and FZCE for cluster analysis of Triticum spp and for the identification of T. durum adulterations. *J Agric Food Chem* 52, 4080-4089

[14] Sarwar M, McDonald CE (1993) Detection of bread wheat farina adulterant in durum wheat semolina and pasta dried at low, high, and ultra-high temperatures. *Cereal Chem* 70, 405-411

[15] Knödler M, Most M, Schieber A, Carle R (2010) A novel approach to authenticity control of whole grain durum wheat (Triticum durum Desf.) flour and pasta, based on analysis of alkylresorcinol composition. *Food Chem* 118, 177-181

[16] Xie Z, Wang C, Wang K, Wang S, Li X, Zhang Z, Ma W, Yan Y (2010) Molecular characterization of the celiac disease epitope domains in α-gliadin genes in Aegilops tauschii and hexaploid wheats (Triticum aestivum L.). *Theor Appl Genet* 121, 1239-1251

[17] Chen F, Xu C, Chen M, Wang Y, Xia G (2008) A new α-gliadin gene family for wheat breedings: somatic introgression line II-12 derived from Triticum aestivum and Agropyron elongatum. *Mol Breed* 22, 675-685

[18] Altenbach SB, Vensel WH, DuPont FM (2010) 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

[19] Molberg Ø, Uhlen AK, Jensen T, Flæte NS, Fleckenstein B, Arentz-Hansen H, Raki M, Lundin KEA, Sollid LM (2005) Mapping of gluten T-cell epitopes in the bread wheat ancestors: implications for celiac disease. *Gastroenterol* 128, 393-401

9 GENERAL DISCUSSION AND PERSPECTIVES

Food allergy is one of the most important public health issue worldwide. In fact, approximately 8% of children and 1-2% of adults have some type of food allergy,[1] but the prevalence of perceived food allergy seems to be much higher than verified food allergy, up to 22% of the adult population.[2] So, the foods eliciting these adverse reactions are safe to consume for the vast majority of consumers, but the resultant adverse reactions in allergic individuals can be quite serious on occasion.[3] About 90% of food allergies are due to proteins from eight main foods: milk, eggs, fish, crustaceans (crab, shrimp, and lobster), peanuts, tree nuts, soybeans, and wheat.[4] The prevalence of food allergies is continuously increasing, especially in developed countries. Since no cure is available for allergic patients, disease management is achieved by strict avoidance of the offending food. This means that allergic consumers must absolutely avoid eating foods that could provoke potentially life-threatening reactions, and successful avoidance depends on having complete and accurate information on food labels. Thus, huge efforts are made by regulatory agencies, with the collaboration of food industry, to protect allergic consumers, to ensure that all food allergens present in a food are declared on the label and that effective controls are used to prevent the presence of unintended allergens.[5] In the case of children, dietary elimination of nutrient-dense foods may result in inadequate nutrient intake and impaired growth: children with multiple food allergies have a higher risk of impaired growth and may have a higher risk of inadequate nutrient intake than children without food allergies. In addition to this, the social lifestyle of individuals with a food allergy, or of the families with an allergic child or family member, can be seriously worsened by the need for continuous vigilance to avoid foods to which they are allergic.[6] Moreover, the repercussions of food allergy are not only limited to individuals or households: the food industry must also sustain a lot of extra costs due to food allergy.[7] In primis, legislative changes, such as the new EU-legislation on food labelling (EU Directive 2003/89/EC amending Directive 2000/13/EC), force the industry to adapt productive processes, food labelling and monitoring to improve allergic consumer protection. The onus of responsibility falls to the food manufacturer, who is required to manage production processes to ensure allergenic ingredients are labelled.[8] Up to now, the potential social impact and economic costs of food allergy on the individual, families, health-related services and food industry is really relevant.

Wheat is in the list of the eight main allergenic foods, because the gluten contained in it is the main external trigger of celiac disease. Celiac patients eat several types of gluten-free products, some of them are naturally gluten-free foods (fruits, vegetables, and unprocessed meat, fish and poultry) but some others are gluten-free substitute foods (pasta, bread, cereals, crackers and snack foods) in which wheat flour is replaced by gluten-free flours. Gluten free products can be purchased at general and specialty food stores as well as via internet. Several studies demonstrated that gluten-free food is not always readily available and it is considerably more expensive than regular, gluten-containing foods.[9] The increasing incidence of celiac disease in the population has negative effects not only on people quality of life, but also on the health care system: it has been estimated that the average per-patient annual healthcare costs in primary care significantly increased by 91% for CD patients after they had been diagnosed with the disease.[10] The impact is also evident for the agricultural and food sectors:

wheat is one of the first three cereals for diffusion and cultivation for human nutrition. Gluten, the main trigger of celiac disease, is at the basis of rheological properties of wheat based products. In fact the formation of a gluten network in the dough is of outmost importance for air bubbles and starch retention (respectively for leavened products and pasta). A low gluten content of the flour leads to loss of product shape in the case of leavened products and to soft and mushy pasta. The consequence is that wheat breeding has been, during the last decades, oriented toward increasing yield and the amounts of amylopectin, gluten and protein.[11]

At the moment no therapies are available for people that are already celiac, so the only treatment is the gluten free diet. But, on the other hand, efforts can be made in the direction of decreasing celiac disease incidence. Different hypothesis have been made on the reasons of the increased incidence of celiac disease. Since celiac disease affect the gastrointestinal tract, the gut microflora can play a key role in the loss of the immunological tolerance. For example, rodshaped bacteria in the upper small bowel are present in one-third of the children with CD but in less than 2% of the controls;[12] another study showed that the species *Bacteroides fragilis* is more abundant in the intestinal microbiota of CD patients, whereas *Bacteroides ovatus* is less abundant in comparison to healthy controls.[13] Beside usual microflora, also viral and bacterial gastroenteritis may have a role in celiac disease pathogenesis: in fact it has been previously demonstrated that a high frequency of rotavirus infections may increase the risk of celiac disease autoimmunity in childhood in genetically predisposed individuals.[14] For what concerns gluten, timing of gluten introduction into the infant diet is associated with risk of celiac disease autoimmunity.[15] Recent studies demonstrated that the oral tolerance to gluten can be lost also in the elderly:[16] the study was conducted following a cohort from 1974 up to now. Parallely, it appears that vital gluten consumption has tripled since 1977. This increase is of interest because it is in the time frame that fits with the predictions of an increase in celiac disease.[17] So, it seems that a massive and early exposure to gluten can be one of the causes of the switch from oral tolerance to celiac disease. Another cause that it has been hypothesized is the transition from sourdough fermentation of break and baked products to yeast fermentation. So, the bacterial proteolytic activity is rather promising not only as currently demonstrated for eliminating traces of contaminant gluten but probably also in perspective for the manufacture of tolerated baked goods.[18]

Thus, trying to decrease these risk factors could help to stop the rising of celiac disease incidence. It is known since a long time that breast feeding has a protective effect against the development of celiac disease, especially when it is still ongoing during gluten introduction in the diet. Also the improvement of infant milk formula, decreasing protein content and osmolarity, has helped to reduce celiac disease incidence.[19] Obviously, the easiest way to reduce the amount of gluten ingestion is the reduction of wheat-derived products consumption, but this would mean a kind of “preventive gluten free diet”, with all the problems and limitations previously described (first of all the decrease in life quality). An alternative way could be the reduction of gluten content in wheat (in contrast with what done in the last decades), but this would mean a dramatic decrease in the texture quality of baked products and pasta. Since gluten proteins have reserve role (nitrogen stock), they underwent to a limited evolutionary pressure, thus showing a high sequence variability with a lot of different isoforms. This lays the groundwork for a possible varietal selection aimed to have the same total gluten amount (so maintaining the same rheological properties) but expressing protein isoforms with a reduced content of sequences involved in

celiac disease. In this way the exposure of the population to immunotoxic sequences will be reduced and, hopefully, also the incidence of the disease.

Differences Gluten coding genes have been extensively characterized, beside for their technological and functional implication in baked products, also for evaluating how much the wheat genetic characteristics can impact on the final immunotoxicity of gluten. One of the most studied immunogenic peptides, the 33-mer LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF, has been demonstrated to be encoded by the 6D chromosome, thus being absent in diploid and tetraploid *Triticum* species. The lacking of genes encoding for some immunogenic sequences in diploid/tetraploid varieties has also found evidences in T-cell assays performed on gluten chymotryptic digests, that gave different results among the species tested.[20] A lot of efforts have been made in the last years in the direction of decreasing wheat toxicity for celiac patients, for example, making use of ex vivo organ cultures and immunohistochemistry assays.[21] The K562 cell agglutinating activity of wheat has been known for many years[22] to screen cereal toxicity for celiac patients.[23] In the great majority of the studies performed in the past to assess gluten toxicity with immunochemical methods, there was a lacking of a complete characterization of the digestion mixture obtained. Except for some rare cases[24], a peptidic characterization of the digested mixture was not carried out and, if performed, only the main immunotoxic peptides were identified. In the recent years the advances in the field of mass spectrometry allowed a more accurate identification of the peptides generated upon digestion, but the majority of these studies have been carried out mainly on recombinant gliadin and synthetic peptides.[25, 26, 27, 28] Other studies took in consideration purified gliadins, to identify the target peptides of tissue transglutaminase.[29] Gluten has been so infrequently used in such studies because the complexity of the chemistry of the cereal proteins has made it extremely difficult to find their toxic components. Previous studies have demonstrated that gliadin and glutenin are each composed of at least 50 closely related species.[30] Thus, this thesis (Chapter 4) deeply investigated the composition of peptide mixtures derived from simulated gastrointestinal digestion of the whole gliadin fraction of wheat. The determination of the amino acid sequence of the gluten derived peptides was achieved using reverse phase chromatography (HPLC, UPLC and μHPLC) coupled with different types of mass spectrometer (both high and low resolution, e.g. LTQ OrbiTrap, single and triple quadrupole). This extensive characterization can give useful information for a better understanding of the peptides that presumably come in contact with the intestinal mucosa, triggering the immunological response in celiac patients. Moreover, it allowed to quantify (using the isotopically labelled internal standard method) peptides containing sequences involved in celiac disease: these data can be very helpful for interpretation of the results of immunological assays, since the different response can be due both to different epitopes generation in terms of amino acid sequence and to a different relative amount of pathogenic peptides. The accurate molecular characterization of the digested mixtures is, as demonstrated in Chapter 5, an interesting tool for the screening of different wheat lines aimed to identify those producing a smaller amount of pathogenic peptides. The results obtained confirm, from a molecular point of view, what has been previously assessed using genetic and immunologic approaches, that there is a strong influence of the genotype in the final amount of peptides containing sequences involved in celiac disease. More specifically, the peptides that are more affected by genetic features are those eliciting the adaptive immune system (immunogenic peptides). Recent studies demonstrated that number of subjects that lost the immunological tolerance to gluten in their adulthood is increasing and among the possible causes there is also the amount and the quality of ingested gluten.[31] This

means that the use of less immunogenic wheat varieties (especially in the preparation of baby foods) can reduce the exposure to gluten, possibly decreasing the incidence of the disease. Of course, these data take in consideration the molecular point of view, so it would be really interesting to cross the data with immunological tests (such as T cell proliferation assays or K562 cells agglutination) on the samples in order to verify the quality of the correlation between pathogenic peptides content and immune response.

In order to perform immunological assays on gluten peptides, it is necessary to simulate the human gastrointestinal digestion on gliadin/gluten/wheat samples. In literature a high number of articles in which *in vitro* digestion model were applied to study gluten peptides is present, but the weakness of these approaches is that they are all different. Previously used models are not consistent with each other for the type of enzyme used (peptic/tryptic digests, peptic/chymotryptic, pancreatin, eventual exoproteases, ...), for the digestion times of the gastric and intestinal phase (from 20 min to several hours), for the buffering agents used (HCl, formic acid, bicarbonate or phosphate buffer, ...), for the protein:enzyme ration used and so on. All these factor could have a strong influence on the outcome of the digestion, in terms of peptide sequence and amount, that can be reflected also on subsequent analysis on the gluten digest. So, in this thesis work (Chapter 3) a qualitative and quantitative comparison of the peptides generated was performed applying two extremely different digestion model: a very simple peptic/tryptic-chymotryptic digestion of a gliadin extract and a more complex and more physiological method involving the use of artificial digestive juices. Results clearly showed that the peptide composition obtained is completely different: the method are quite well related in term of total amount of immunotoxic sequences produced but the peptides generated are different. This means that simplified methods are suitable mainly for varietal screening or comparison purposes, but for biological experiment it should be better to use more physiological systems. Of course, the limit of these models is the lacking of brush border membrane enzymes (that can further proteolyzed the peptides) and of the intestinal microflora. So, a possible interesting continuation of the work could be the use of a more physiological model taking into account also these latter variables and studying the effects in terms of peptides produced.

During the course of this PhD, it has been discovered that α-amylase/trypsin inhibitors (in particular CM3 and 0.19) can activate a strong innate immune response via activation of the Toll Like Receptor 4. This means that they are involved in a wide variety of diseases, like celiac disease but also irritable bowel syndrome, inflammatory bowel disease and even non intestinal inflammations.[32] In addition to this, CM3 was already known in literature as an allergen triggering in particular baker's asthma.[33] The development of a LC-MS method for the detection of this allergen could be useful method for routine analysis and it has been used also in this thesis to demonstrate the strong influence of the cultivation area on CM3 content. The further identification of agronomical practices aimed to reduce the content of this allergen could be a good tool for reducing people exposure, thus the incidence of baker's asthma and, possibly, also celiac disease. It must be underlined that this is one of the first works aimed at the quantification of this new non-gluten trigger for celiac patients (Chapter 6).

So, according to the results obtained, a switch in the selection criteria for wheat varieties is advisable and breeders should take into account, beside gluten and total protein content, also the amount of immunogenic peptides generated upon digestion. This new approach to wheat genotypic selection should allow to reduce the exposure of predisposed subject to immunogenic sequences, reducing the risk of a further disease development

and constituting a very efficient approach in celiac disease prevention. Moreover, it has also been demonstrated that also cultivation practices should be investigated and adapted in order to reduce allergenic potential of wheat crops, with particular regards to the allergen CM3. This should be also useful on two different aspects. First, the reduction of the exposition to the sensitizing allergen for people employed in the production chain of pasta and bakery products leading to a possible decrease of subjects with baker's asthma or wheat allergy. Second, the reduction of this protein will reduce the innate immune response and, since the activation of innate immunity mechanisms seems to precede the adaptive immune response in the mechanism of celiac disease,[34] it could significantly contribute to celiac disease prevention. Patents for pharmaceutical antibodies formulations aimed to neutralize α-amylase/trypsin inhibitors already exist,[35] but agronomical practices can be an alternative or complementary way to reduce innate response to this class of proteins. A synergistic action of varietal breeding aimed to reduce immunogenic peptides (adaptive immune response) and of agronomical practices aimed to reduce CM3 content (innate immune response) is therefore desirable.

It must be underlined that all these studies are aimed to celiac disease prevention, and wheat varieties (even with a reduce content of celiac disease related peptides) are not suitable for the consumption by celiac patient. For celiac patients, the only treatment is the strict avoidance of gluten, so the only defence are good detection methods in order to allow a correct verification of labelling. Nowadays two main "schools" of gluten detection method are available, also in form of commercial kits. The first methods are PCR-based, that act at genetic level: it is important to underline that these are indirect methods, since they do not quantify the specific allergenic protein but the DNA of the ingredient that contains the allergen. The second method is based on immunoenzymatic assays, that determine the proteic fraction of the allergenic ingredient, thus being very sensitive to protein denaturation or hydrolysis that can occur in processed foods. In the last years also MS detection of food allegen is in rapid development, so in this PhD thesis a broad evaluation of detection methods for gluten quantification was performed. Two PCR-based methods (one commercial kit and one previously developed in literature) were compared with seven immunoenzymatic assay (covering almost all the kits on the market) and a mass spectrometry methods here developed. The high gluten genetic variability affect all the three methods, but PCR-methods seemed to be much more affected by the matrix effect, since in different matrices it resulted impossible to extract a sufficient DNA amount to achieve good amplification in the subsequent RT-PCR. This problem was not present with ELISA kits, that anyway gave very different results from one kit to another: that means that the determined gluten content can be different depending on the kit used, and this is a problem especially for "border line" samples, with a gluten content around the legal limit of 20 ppm. MS methods allow to avoid the extraction step, since the proteolytic cleavage can be performed on the whole matrix, thus avoiding matrix effect during extraction, especially in the case of processed foods. The use of calibration curves in matrix allows to obtain quite accurate quantification, especially at trace levels (2 ppm), at which most of the immunoenzymatic assays are under the limit of quantification or give very unreliable results (Chapter 6). Anyway, further improvement can be made, for example with the use of isotopically labelled peptides; moreover, MS could allow, with the development of opportune methods, the detection of multiple allergens in a single digestion/run that could be very useful for routine screening of food products. Marker peptides approach can be used also, beside safety assessment, for quality evaluation of wheat derived products. As demonstrated here, marker peptides present only in common wheat can be used to detect the adulteration/contamination of durum wheat

flours with the less expensive common wheat (Chapter 7). Also in this case the direct proteolytic approach on the food matrix can overcome problems of DNA extraction especially with high starchy matrices; moreover the MS can complement the routine electrophoretic technique, giving a more accurate quantification.

---

[1] Maleki SJ, Hurlburt BK (2002) Food allergy: recent advances in food allergy research. *Crop biotechnology*, ACS Symposium Series, pp 192-204.

[2] Woods RK, Stoney RM, Raven J, Walters EH, Abramson M, Thien FC (2002) Reported adverse food reactions overestimate true food allergy in the community. *Eur J Clin Nutr* 56,31–36

[3] Taylor SL, Nordlee JA, Bush, RK (1992) Food allergies. *Food Safety Assessment*, ACS Symposium Series, pp 316-329

[4] Wilkinson SL (1998) Deconstructing Food Allergies. *Chem. Eng. News* 76(36), 38–40

[5] Gendel SM (2013) The regulatory challenge of food allergens. *J Agric Food Chem* 61(24), 5634–5637

[6] Munoz-Furlong A (2001) Living with food allergies: not as easy as you might think. *FDA Consum* 35-40

[7] Wilson E, Fordham R, Mugford M (2006) A review of allergy-related factors affecting food industry efficiency. University of East Anglia: Report from the Health Economics Group,

[8] de Blok BM, Vlieg-Boerstra BJ, Oude Elberink JN, Duiverman EJ, DunnGalvin A, Hourihane JO, Cornelisse-Vermaat JR, Frewer L, Mills C, Dubois AE (2007) A framework for measuring the social impact of food allergy across Europe: a EuroPrevall state of the art paper. Allergy 62,733–737

[9] Lee AR, Ng DL, Zivin J, Green PHR (2007) Economic burden of a gluten-free diet. J Hum Nutr Diet 20(5), 423-430

[10] Violato M, Gray A, Papanicolas I, Ouellet M (2012) Resource use and costs associated with coeliac disease before and after diagnosis in 3,646 cases: results of a uk primary care database analysis. PLoS ONE 7(7): e41308. doi:10.1371/journal.pone.0041308

[11] Morris CE, Sands DC (2006) The breeder's dilemma—yield or nutrition? *Nat Biotechnol* 24(9), 1078-1080

[12] Forsberg G, Fahlgren A, Hörstedt P, Hammarström S, Hernell O, Hammarström ML (2004) Presence of bacteria and innate immunity of intestinal epithelium in childhood celiac disease. *Am J Gastroentero* 99,894–904

[13] Sànchez E, Laparra JM, Sanz Y (2012) Discerning the role of bacteroides fragilis in celiac disease pathogenesis. Appl Env Microbiol 78(18), 6507-6515

[14] Stene LC, Honeyman MC, Hoffenberg EJ, Haas JE, Sokol RJ, Emery L, Taki I, Norris JM, Erlich HA, Eisenbarth GS, Rewers M (2006) Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. *Am J Gastroenterol* 101, 2333-2340

[15] Norris JM, Barriga K, Hoffenber EJ, Taki I, Miao D, Haas JE, Emery LM, Sokol RJ, Erlich HA, Eisenbarth GS, Rewers M (2005) Risk of celiac disease autoimmunity and timing of gluten introduction in the diet of infants at increased risk of disease. *JAMA* 293(19),2343-2351

[16] Catassi C, Kryszak D, Bhatti B, Sturgeon C, Helzlsouer K, Clipp SL, Gelfond D, Puppa E, Sferruzza A, Fasano A (2010) Natural history of celiac disease autoimmunity in a USA cohort followed since 1974. *Ann Med* 42, 530–538

[17] Kasarda DD (2013) Can an increase in celiac disease be attributed to an increase in the gluten content of wheat as a consequence of wheat breeding? *J Agr Food Chem* 61, 1155−1159

[18] Gobetti M, Rizzello CG, Di Cagno R, De Angelic M (2007) Sourdough lactobacilli and celiac disease. Food Microbiol 24(2), 187-196

[19] Stevens FM, Egan-Mitchell B, Cryan E, McCarthy CF, McNicholl B (1987) Decreasing incidence of coeliac disease. *Arch Dis Child* 62, 465-468

[20] Molberg Ø, Uhlen AK, Jensen T, Flæte NS, Fleckenstein B, Arentz-Hansen H, Raki M, Lundin KEA, Sollid LM (2005) Mapping of gluten T-cell epitopes in the bread wheat ancestors: implications for celiac disease. *Gastroenterol* 128, 393-401

[21] Carroccio A, Di Prima L, Noto D, Fayer F, Ambrosiano G, Villanacci V, Lammers K, Lafiandra D, De Ambrogio E, Di Fede G, Iacono G, Pogna N (2011) Searching for wheat plants with low toxicity in celiac disease: Between direct toxicity and immunologic activation. *Digest Liver Dis* 43, 34-39

[22] De Vincenzi M, Dessi MR, Giovannini C, Maialetti F, Mancini E (1995) Agglutinating activity of wheat gliadin peptide fractions in coeliac disease. *Toxicology* 96, 29-35

[23] Silano M, De Vincenzi M (1999) In vitro screening of food peptides toxic for coeliac and other gluten-sensitive patients: a review. *Toxicology* 132,99–110

[24] Van Dewal Y, Kooy YMC, Van Veelen PA, Peña SA, Mearin LM, Molberg Ø, Lundin KEA, Sollid LM, Mutis T, Benckhuijsen WE, Drijfhout J, Koning F (1998) Small intestinal T cells of celiac disease patients recognize a natural pepsin fragment of gliadin. *Proc Natl Acad Sci USA* 95, 10050–10054

[25] Molberg O, Mcadam SN, Korner R, Quarsten H, Kristiansen C, Madsen L, Fugger L, Scott H, Noren O, Roepstorff P, Lundin KE, Sjostrom H, Sollid LM (1998) Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nat Med* 4, 713-717

[26] van de Wal Y, Kooy Y, van Veelen P, Pena S, Mearin L, Papadopoulos G, Koning F (1998) Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *Eur J Immunol* 161, 1585-1588

[27] Quarsten H, Molberg Ø, Fugger L, McAdam SN, Sollid LM (1999) HLA binding and T cell recognition of a tissue transglutaminase-modified gliadin epitope. *Eur J Immunol* 29, 2506-2514

[28] Arentz-Hansen H, Korner R, Molberg O, Quarsten H, Vader W, Kooy YM, Lundin KE, Koning F, Roepstorff P, Sollid LM, McAdam SN (2000) The intestinal T cell response to α-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *J Exp Med* 191, 603-612

[29] 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 Letters* 562, 177-182

[30] 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

[31] Catassi C, Kryszak D, Bhatti B, Sturgeon C, Helzlsouer K, Clipp SL, Gelfond D, Puppa E, Sferruzza A, Fasano A (2010) Natural history of celiac disease autoimmunity in a USA cohort followed since 1974. *Annals of Medicin*, 42, 530–538

[32] Junker Y, Zeissig S, Kim SJ, Barisani D, Wieser H, Leffler DA, Zevallos V, Libermann TA, Dillon S, Freitag TL, Kelly CP, Schuppan D (2013) Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. J Exp Med 209(13), 2395-2408

[33] Kusaba-Nakayama M, Ki M, Iwamoto M, Shibata R, Sato M, Imaizumi K (2000) CM3, one of the wheat α-amylase inhibitor subunits, and binding of IgE in sera from Japanese with atopic dermatitis related to wheat. *Food Chem Toxicol* 38, 179-185

[34] Maiuri L, Ciacci C, Ricciardelli I, Vacca L, Raia V, Auricchio S, Picard J, Osman M, Quaratino S, Londei M (2003) Association between innate response to gliadin and activation of pathogenic T cell in coeliac disease. *Lancet* 362, 30–7.

[35] Schuppan D, Junker Y, Libermann T, Dillon ST (2013) Methods and compositions for treating celiac disease. Patent n° US20130266584 A1.

10 SUMMARY

In this PhD thesis, peptides deriving from simulated gastrointestinal digestion of wheat proteins were extensively studied, with particular focus on those containing aminoacidic sequences known in literature to be involved in celiac disease. Celiac disease is an autoimmune entheropathy that develops in some genetically predisposed subject after consumption of gluten or related proteins of barley and rye. Peptides derived from gastrointestinal gluten digestion can trigger the immune response following two different pathways: some peptides, usually called immunogenic peptides in literature, contain sequences able to elicit the adaptive immune response, while some other peptides, usually referred to as toxic in literature, are able to stimulate the innate immune response. It has been previously demonstrated in literature that a massive exposure to gluten during weaning leads to a higher risk of celiac disease in the following years, but the loss of the immunological tolerance to gluten can occur even in adults due to a complex interplay of several factors: some examples are changes in the intestinal microflora, the type of fermentation in baked products, stress factors such as intestinal infections and, last but not least, the amount and quality of gluten ingested. This could be a key factor on which we can act because, since gluten proteins have reserve role (nitrogen stock) they did not undergo to evolutionary pressure, thus showing a high sequence variability with a lot of different isoforms. This lays the groundwork for a possible varietal selection aimed to have the same total gluten amount (so maintaining the same rheological properties) but expressing protein isoforms with a reduced content of sequences involved in celiac disease.

So, the first step of the PhD project was the comparison of different simulated gastrointestinal digestion methods applied to wheat proteins and the subsequent characterization of the peptides generated using liquid chromatography-mass spectrometry techniques (Chapter 3). Once decided the in vitro digestion model to be adopted, the subsequent step was a varietal screening aimed to evaluate differences in the total amount of peptides containing sequences known in literature to be involved in celiac disease (Chapters 4 and 5). Finally, a recently discovered non-gluten trigger for celiac disease patients was studied, developing a liquid chromatography-mass spectrometry method for its detection (Chapter 6). Also in this the impact of genetic and environmental factors was evaluated.

In literature are present several digestion methods applied to gluten proteins: generally these models are very simple and they involve the use of only the main gastric and pancreatic proteases (pepsin, trypsin and chymotrypsin) and a buffering agent to keep the correct pH value for every phase. These methods were compared to a more complex digestion procedure, previously used in literature to assess the release of mycotoxins and heavy metals from food matrices. This method involves the use of digestive juices whose chemical composition strictly reflects the physiological one. In both cases (simple and complex model) the peptides generated from the digestion were characterized using liquid chromatography coupled with mass spectrometry. In this *in vitro* experiments, the processes occurring in the human gastrointestinal tract during food digestion were simulated, and the outcome of the digestion was studied by LC-MS techniques. With the use of tandem mass spectrometry the exact aminoacidic sequence of the peptides generated by the digestion was determined and, among all the peptides, the ones containing sequences known to be implied in celiac disease were identified.

Strong differences were present between the two digestion models. First, with the simplified model almost all the peptides derive from α-gliadin, while with the physiological method they are equally distributed among α- and γ-gliadins and LMW-glutenins. This can be explained by the observation that α-gliadin derived peptides of the simplified method are further proteolyzed into shorter peptides in the physiological model and often these shorter peptides did not contain immunotoxic sequences anymore. Moreover, in the physiological model are present enzymes other than proteases (like amylase and lipase) that, even if not directly implied in protein cleavage, can contribute (together with bile salts) to matrix degradation, thus improving the extractability and digestibility of higher molecular weight proteins such as γ-gliadins and glutenins.

Thus, in the case a subsequent immunological experiments or biological trials have to be performed, the more physiological method is more suitable than the simplified one, because the peptides generated are really different and the more complex method is more similar to what really happens in the human gastrointestinal tract.

The peptides containing immunotoxic sequences were quantified for both the *in vitro* digestion models along the pasta production chain, in order to evaluate also the suitability of the two methods for processed foods. The samples (kernels, semolina, dough, extruded pasta, dried pasta and cooked pasta) were obtained from three different durum wheat varieties (Svevo, Meridiano and Saragolla). The physiological digestion method produced lesser amount of toxic and a higher amount of immunogenic peptides compared to the simplified one, probably due to the different molecular weight of the peptides generated. Anyway a noticeable result is that the difference among the varieties tested remains unchanged, with Saragolla showing a lower content of peptides involved in celiac disease compared to Svevo and Meridiano. Another remarkable result is that the simplified method cannot be applied to thermally treated foods, because heating induces gluten polymerization leading to a poor proteins extractability. Anyway, the two different model are very well correlated in terms of total amount of immunotoxic peptides generated so, since the aim of the thesis was to perform a varietal screening, the simplified method was more suitable for large amount of samples to be analyzed.

The simplified in vitro digestion method was then applied to 45 durum wheat samples belonging to five different varieties and harvested in three different Italian areas (Argelato in the North of Italy, Falconara in the Centre and Lucera in the South). From the results no major differences due to the different cultivation place (consistently with the reserve role of that class of proteins that thus is not affected by environmental factors). On the other hand, statistically significant differences are present among the five varieties tested, especially for what concern immunogenic peptides. Since the cultivar selection operated by breeders in the last years in order to achieve the desired rheological properties has led to a decrease in the genetic biodiversity of durum wheat varieties present nowadays on the market, 25 durum wheat accessions were selected from a durum wheat panel in order to maximize the genetic biodiversity of the samples (and thus eventual differences in immunotoxic peptides production upon digestion). Results obtained from every single accessions were mediated in five groups on the basis of phylogenetic affinity on dendrogram.

For toxic peptides no significant differences are present while strong variability emerged for immunogenic peptides, with accessions of the second groups (ICARDA accessions for temperate areas) showing a significantly lower content of peptides eliciting adaptive immune response.

The higher variability of immunogenic peptides compared to toxic peptides can be explained on the basis of gliadins sequence variability: in fact, toxic peptides usually derive from the N-term region of the protein, that is the most conserved. On the opposite, immunogenic peptides derive from a region of the protein showing a much higher variability. So, different wheat genotypes can express different gliadins isoforms thus showing a different final content of immunogenic sequences. Then, it is possible to select wheat varieties with a good gluten content (and good rheological properties) but with a reduced amount of immunogenic sequences in order to reduce the exposure of people and decreasing the risk of new cases of disease.

Finally, it was recently discovered a new non gluten trigger for celiac disease patients. In fact, it has been recently discovered that α-amylase/trypsin inhibitors, already known in literature to be involved in wheat allergies and baker's asthma, can bind the Toll-Like Receptor 4, thus activating a strong innate immune response. The two α-amylase/trypsin inhibitors mainly involved are CM3 and 0.19. In this PhD thesis the α-amylase/trypsin inhibitor CM3 was identified in durum wheat salt soluble extracts using a bottom-up proteomic approach and a liquid chromatography-mass spectrometry method for its detection was developed. The α-amylase/trypsin inhibitor CM3 was quantified using the isotopically labelled internal standard method using a marker proteolytic peptide. In order to evaluate varietal and environmental differences the samples analyzed belong to three different durum wheat varieties (Svevo, Levante and D240) harvested in three different Italian areas (Argelato and Poggio Renatico in the North and Lucera in the South).

While the difference among the varieties is in this case negligible, the differences among the different cultivation areas are more evident and statistically significant, even if some location (Argelato and Poggio Renatico) were very closed together. This finding is consistent with the defensive function that α-amylase/trypsin inhibitors have *in vivo* in the plant: in fact, being defence proteins, the different agronomical practices and/or the different pedoclimatic conditions can affect the pest infestations, so the need for the plant to express defence proteins. In this part of the work a novel non gluten trigger for celiac disease patient was accurately quantified, demonstrating a strong influence of environmental factors on its final amount.

Given the severity of allergic reactions, analytical tools for allergen detection in food products are becoming more and more important. So, in Chapter 6 of the thesis, PCR-based methods, immunoenzymatic assays and MS based methods were compared for gluten detection in different food matrices. All these methods present several weakness depending on the matrix and on the gluten amount to be detected, highlighting a very important issue in the field of food safety: different detection methods give different results and this problem is significant for samples that have a gluten content close to the legal limit of 20 ppm. The same peptidomic approach used for gluten quantification with mass spectrometry can be used also for quality assessment: in Chapter 7 it has been used to detect and quantify common wheat presence in durum wheat flours, due to contamination during processing or food adulteration.

11 CURRICULUM VITAE

PERSONAL INFORMATION

Barbara Prandi

Via Matteotti 7, 42027, Montecchio Emilia (RE), Italy

+390522866916 +393333485981

barbara.prandi@unipr.it

Sex Female | Date of birth 08/12/1986 | Nationality Italian

CURRENT POSITION

PhD student Food Science and Technology, Research Fellow

WORK EXPERIENCE

From January 2011 to date

PhD student in Food Science and Technology

At Department of Organic and Industrial Chemistry (University of Parma) from January 2011 to January 2013, Department of Food Science (University of Parma) from January 2013 to June 2013 and at ILVO (Instituut voor Landbouw- en Visserijonderzoek, Belgium) from June 2013 to December 2013

- Study of molecular features involving food allergies and intolerances
- Nitrogen fraction extraction (amino acids, peptides, proteins) from different food matrices
- Nitrogen fraction characterization using Kjeldahl, chromatographic techniques (LC and GC), UV, fluorescence, mass spectrometry, MS/MS
- Bottom-up proteomic: mono- and bidimensional electrophoresis, in-gel digestion, MS/MS
- Simulated gastrointestinal digestion of different food matrices and characterization of the peptides generated
- Solid phase peptide synthesis of peptides of interest in food or immunology and subsequent purification
- Esterification of peptides and protein hydrolysates with food grade alcohols and assessment of hydrophilic-lipophilic antioxidant activity
- Lipid fraction characterization using LC-MS
- Allergens detection using RT-PCR, ELISA and LC-MS/MS

Business or sector Food chemistry – Organic chemistry

From October 2013 to date

Research Fellow Organic Chemistry – Food Chemistry (CHIM/06-CHIM/10)

At Interdepartmental Center SITEIA.PARMA (University of Parma)

- Nitrogen fraction characterization in complex food matrices and related properties
- Advanced proteomic and peptidomic methods applied to the nitrogen fraction of food matrices
- Assessment of modifications occurred after simulated gastrointestinal digestion
- Study of chemical and biological properties of the peptides present
- Pectin oligomers characterization using LC-MS
- MS/MS methods development for allergen detection in food

Business or sector Food chemistry – Organic chemistry

From June 2011 to June 2013

Research Fellow Organic Chemistry

At Department of Organic and Industrial Chemistry (University of Parma) from June 2011 to January 2013, Department of Food Science (University of Parma) from January 2013 to June 2013

- Gluten digestion: pathogenic peptides involved in celiac disease
- Nitrogen fraction extraction (proteins and peptides) from different food matrices
- Nitrogen fraction characterization using Kjeldahl, chromatographic techniques, UV, fluorescence, mass spectrometry, MS/MS, electrophoresis and in-gel digestion
- Simulated gastrointestinal digestion of different food matrices and characterization of the peptides generated
- Solid phase peptide synthesis of peptides involved in celiac disease and subsequent purification

Business or sector Organic chemistry

From March 2010 to October 2010

Master thesis internship

At Department of Organic and Industrial Chemistry (University of Parma)

- Peptides derived from simulated gastrointestinal digestion of the prolamin fraction extracted from different wheat varieties: implications for celiac disease
- Extraction of the prolamin fraction from wheat
- Gluten proteins electrophoresis
- Simulated gastrointestinal digestion of the prolamin extracts
- Characterization of the peptides generated using LC-MS and LC-MS/MS

Business or sector Food chemistry – Organic chemistry

From April 2008 to July 2008

Bachelor thesis internship

At Department of Organic and Industrial Chemistry (University of Parma)

- Study of low molecular weight amino acid derivatives in meat products
- Peptide fraction extraction from meat products
- Characterization of histidyl-dipeptides (anserine, carnosine and balenine), guanidine compounds (creatine and creatinine) and ATP degradation products (inosine and hypoxanthine) using LC-MS, LC-MS/MS and H-NMR

Business or sector Food chemistry – Organic chemistry

From July 2004 to August 2004

Stage

At AGAC (Azienda Gas-Acqua Consorziale) Reggio Emilia

- Waste water quality control
- Chemical laboratory: COD, BOD, phosphors, total nitrogen, DH, suspended solids, ashes, pH

Business or sector Analytical chemistry

EDUCATION AND TRAINING

From October 2008 to October 2010

Master Degree in Food Science and Technology (110/110 cum laude)

University of Parma, Italy

- Chemistry, biochemistry and nutrition applied to the food sector
- Food microbiology
- Food businesses management
- Food products and processes design

From October 2005 to July 2008

Bachelor Degree in Food Science and Technology (110/110 cum laude)

University of Parma, Italy

- Basic subjects (chemistry, biology, physic, English, economy)
- Food chemistry, food microbiology, hygiene, nutrition
- Quality management and food technologies

From September 2000 to June 2005

High School diploma (chemical-biological) 100/100

Industrial Technical Institute L. Nobili (Reggio Emilia, Italy) from September 2000 to June 2002 and Agricultural Technical Institute A. Zanelli (Reggio Emilia, Italy) from September 2002 to June 2005

- Biology, microbiology, biochemistry, general, organic and analytical chemistry
- Laboratories of chemistry, physic and biology

PERSONAL SKILLS

Mother tongue(s) Italian

Other language(s)

UNDERSTANDING		SPEAKING		WRITING	
Listening	Reading	Spoken interaction	Spoken production		
English	B2	B2	B2	B2	B2	
B2 level (certified by Gent University in 2013)					
Dutch	A1		A1	A1	A1	A1

Levels: A1/2: Basic user - B1/2: Independent user - C1/2 Proficient user
Common European Framework of Reference for Languages

Communication skills
- Laboratory of organic and food chemistry lessons for the bachelor degree in Food Science and Technology
- Exercises of organic chemistry for the bachelor degree in Food Science and Technology

Organisational / managerial skills Laboratory activity, correlator of master degree thesis in food science and technology, of bachelor thesis in food science and technology and in chemistry.

Job-related skills
- Chromatographic techniques, especially liquid chromatography
- Mass spectrometry and MS/MS
- Protein electrophoresis
- RT-PCR
- Immunoenzymatic assays
- Antioxidant activity
- Peptides and peptide-derivatives synthesis
- UV-Vis spectrophotometry

Computer skills
- good command of Microsoft Office™ tools
- software for chromatographic data acquisition (MassLynx, Empower) and spectrophotometric data acquisition (LambdaBio, Optima, LightCycler)
- software for statistical analysis (SPSS)

Driving licence
- B

Publications

Prandi B, Mantovani P, Galaverna G, Sforza S (2014) Genetic and environmental factors affecting pathogenicity of wheat as related to celiac disease. *Journal of Cereal Science* 59(1): 62-69.

Prandi B, Faccini A, Tedeschi T, Galaverna G, Sforza S (2013) LC/MS analysis of proteolytic peptides in wheat extracts for determining the content of the allergen amylase/trypsin inhibitor CM3: Influence of growing area and variety. *Food Chemistry* 140(1-2):141-6.

Paolella S, Bencivenni M, Lambertini F, Prandi B, Faccini A, Tonetti C, Vineis C, Sforza S (2013) Identification and quantification of different species in animal fibres by LC/ESI-MS analysis of keratin-derived proteolytic peptides. *Journal of Mass Spectrometry* 48(8): 219-226.

Prandi B, Farioli L, Tedeschi T, Pastorello EA, Sforza S (2012) Simulated gastrointestinal digestion of Pru ar 3 apricot allergen: Assessment of allergen resistance and characterization of the peptides by ultra-performance liquid chromatography/electrospray ionisation mass spectrometry. *Rapid Communication in Mass Spectrometry* 26(24):2905-12.

Prandi B, Bencivenni M, Faccini A, Tedeschi T, Dossena A, Marchelli R, Galaverna G, Sforza S (2012) Composition of peptide mixtures derived from simulated gastrointestinal digestion of prolamins from different wheat varieties. *Journal of Cereal Science* 56(2): 223-231.

Prandi B, Bencivenni M, Tedeschi T, Marchelli R, Dossena A, Galaverna G, Sforza S (2012) Common wheat determination in durum wheat samples through LC/MS analysis of gluten peptides. *Analytical and Bioanalytical Chemistry* 403(10): 2909-14.

Posters

Prandi B, Galaverna G, Sforza S (2012) Simulated gastrointestinal digestion of gluten: assessment of different in vitro models. Proceedings of the 17th Workshop on the Developments in the Italian PhD Research on Food Science Technology and Biotechnology, Alma Mater Studiorum - University of Bologna, pag 253-254.

Prandi B, Farioli L, Tedeschi T, Pastorello EA, Sforza S (2012) Simulated gastrointestinal digestion of Pru ar 3 apricot allergen: Assessment of allergen resistance and characterization of the peptides by ultra-performance liquid chromatography/electrospray ionisation mass spectrometry. Proceedings of the EuroFoodChem XVII, 2013, Food Engineering Department of the Hacettepe University, pag. 468, ISBN: 978-605-63935-0-1

Prandi B, Bencivenni M, Galaverna G, Sforza S, Marchelli R. Peptides derived from simulated gastrointestinal digestion of the prolamin fractions extracted from different wheat varieties: implications for the celiac disease. Translating food chemistry into health benefits, proceedings of EuroFoodChem XVI, 2011, Polish Journal of Food and Nutrition Sciences, Volume 61, 133-134, ISBN: 1230-0322.

Lambertini F, Prandi B, Sforza S, Virgili R, Galaverna G, Dossena A, Marchelli R. Inosine, ipoxanthine, creatine, creatinine and histidine dipeptides in “culatello” sausages. Food for the future-the contribution of chemistry to improvement of food quality, proceedings of EuroFoodChem XV, 2009, Department of life sciences, Faculty of life sciences, University of Copenhagen, Volume 1, 118-121, ISBN: 978-87-993033-4-2

Conferences

Prandi B, Faccini A, Tedeschi T, Galaverna G, Sforza S. LC/MS analysis of proteolytic peptides in wheat extracts for determining the content of the allergen amylase/trypsin inhibitor cm3: influence of growing area and variety. Proceedings of the EuroFoodChem XVII, 2013, Food Engineering Department of the Hacettepe University, pag. 154, ISBN: 978-605-63935-0-1

Seminars

LC-MS 2012: Advanced course on the coupling of mass spectrometry with separation techniques in liquid phase. Lucca, Complesso San Micheletto ,11-14 Giugno 2012.