University of Parma

Faculty of Agriculture

Ph.D. in Food Science and Technology

Cycle XX (2005-2007)

Peptides and proteins as markers of quality and safety in food: an approach by Mass Spectrometry

Candidate:

Valeria Cavatorta

Tutors:

Prof. Rosangela Marchelli

Prof. Stefano Sforza

Ph.D. Coordinator

Prof. Giuliano Ezio Sansebastiano

Peptides and proteins as markers of quality and safety in food: an approach by Mass Spectrometry

by

Valeria Cavatorta

A thesis submitted for the degree of
Doctor of Philosophy
in

Food Science and Technology

This thesis will be defended in public at 9 a.m. on march 28th, 2008
Centro Sant'Elisabetta Auditorium, University Campus, Parma –IT

Per i miei genitori

Per mia sorella

Per i miei nonni

Per Cristian

Preface

Mass Spectrometry (MS) has gained an outstanding momentum in many fields of science, as a really interdisciplinary technique. It nowadays contributes to many sectors of Chemistry, Biology and Medicine, and also in Food Science MS is becoming largely used.

This Ph.D. thesis describes several cases where Mass Spectrometry has been applied to the study of food proteins and peptides for assessing food safety and food quality. The thesis is organized in three parts: a general introduction and two parts concerning the results, one devoted to food safety issues and the other to food quality issues. The extensive introduction is mainly intended for newcomers in the subjects of protein and peptides studies by Mass Spectrometry. Indeed, all the results have their own short introduction, which allows immediate contextualization of the subjects. The first part of results is devoted to allergenic proteins: MS has been exploited in order to detect trace amounts of hidden caseins in hypoallergenic infants formulas (chapter 1), to elucidate their primary sequence and to localize the protein directly on a fruit section (chapter 2), to purify and characterize peptides derived from allergenic protein after a simulated gastro-intestinal digestion (chapter 3).

The second part of the results concerns peptides in cheeses derived from proteolysis of caseins. Peptides identification by LC/MS has been utilized to elucidate changes occurring during cheese ageing (chapter 4), peptide-like structures naturally occurring in aged cheeses (chapter 5), to describe the peptide pattern in correlations with the different technologies of production (chapter 6), to detect sheep cheese adulteration with cows' milk (chapter 7).

Valeria Cavatorta

Table of contents

General introduction

1 PROTEINS AND PEPTIDES

Proteins are the most abundant biological macromolecules, which occur in all cells ranging in size from relatively small peptides to huge polymers with molecular weights of 10^5 D. Proteins exhibit different biological functions and they are the most important final products of the information pathways as proteins are the molecular instruments through which genetic information is expressed.[1]

1.1 Chemical nature of protein, peptides, amino acids

Proteins are polymers of amino acids, with each amino acid residue joined to its neighbor by a peptide bond. The 20 α-amino acids (including the imino acid proline) commonly found as residues in proteins contain an α-carboxyl group, an α-amino group, and a distinctive R group substituted on the α-carbon atom (Figure 1.1).

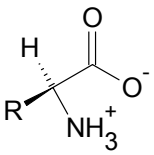


Figure 1.1. Basic structural formula of amino acids.

The α-carbon atom of all amino acids except glycine is asymmetric, and thus amino acids can exist in at least two stereoisomeric forms, D and L enantiomers; only L enantiomers are found in proteins.

Some amino acids have a hydrophobic, others an hydrophilic character (Figure 1.2) and well established acid-base properties (Table 1.1).

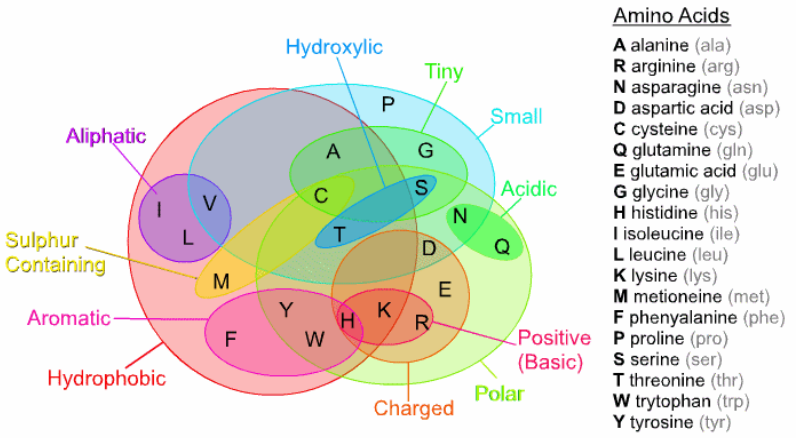


Figure 1.2. Hydrophobic/hydrophilic character, dimension and charge state of side group of 20 fundamental amino acids.

Table 1.1. Properties of standard amino acids.

Amino acid	Abbreviated names		M_r	pK_s values: pK_1 (—COOH)	pK_2 ($—NH_3^+$)	pK_R (R group)	pI
Nonpolar, aliphatic R groups							
Glycine	Gly	G	75	2.34	9.60		5.97
Alanine	Ala	A	89	2.34	9.69		6.01
Valine	Val	V	117	2.32	9.62		5.97
Leucine	Leu	L	131	2.36	9.60		5.98
Isoleucine	Ile	I	131	2.36	9.68		6.02
Methionine	Met	M	149	2.28	9.21		5.74
Aromatic R groups							
Phenylalanine	Phe	F	165	1.83	9.13		5.48
Tyrosine	Tyr	Y	181	2.20	9.11	10.07	5.66
Tryptophan	Trp	W	204	2.38	9.39		5.89
Polar, uncharged R groups							
Serine	Ser	S	105	2.21	9.15		5.68
Proline	Pro	P	115	1.99	10.96		6.48
Threonine	Thr	T	119	2.11	9.62		5.87
Cysteine	Cys	C	121	1.96	10.28	8.18	5.07
Asparagine	Asn	N	132	2.02	8.80		5.41
Glutamine	Gln	Q	146	2.17	9.13		5.65
Positively charged R groups							
Lysine	Lys	K	146	2.18	8.95	10.53	9.74
Histidine	His	H	155	1.82	9.17	6.00	7.59
Arginine	Arg	R	174	2.17	9.04	12.48	10.76
Negatively charged R groups							
Aspartate	Asp	D	133	1.88	9.60	3.65	2.77
Glutamate	Glu	E	147	2.19	9.67	4.25	3.22

The peptide bon is planar, since it has a partial double bond character and in linear peptides and proteins is in the trans-configuration (Figure 1.3).

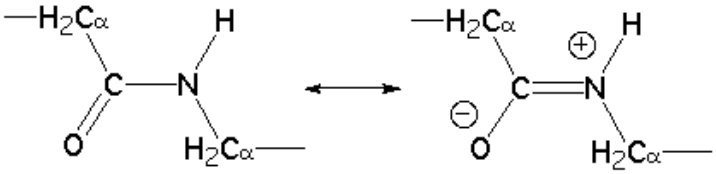


Figure 1.3. The Peptide Bond partial double bond character and partial charges.

When a few amino acids are joined together (up to 12 amino acids), the species is called “oligopeptide”. When the sequence contains many amino acids (up to 50), the product is called “polypeptide”; with more than 50 amino acids, “protein”. Proteins may have thousands of amino acid residues, although the terms “protein” and “polypeptide” are sometimes used interchangeably.

No generalizations can be made about the molecular weights of biologically active peptides and proteins in relation to their functions. Even the smallest peptides can have biologically important effects.

1.2 "Structure in structure" of a protein

Protein structures are at several levels of complexity, arranged in a kind of hierarchy. Four levels of protein structure are commonly defined (Figure 1.4) and each step of organization has its proper methods of analysis.

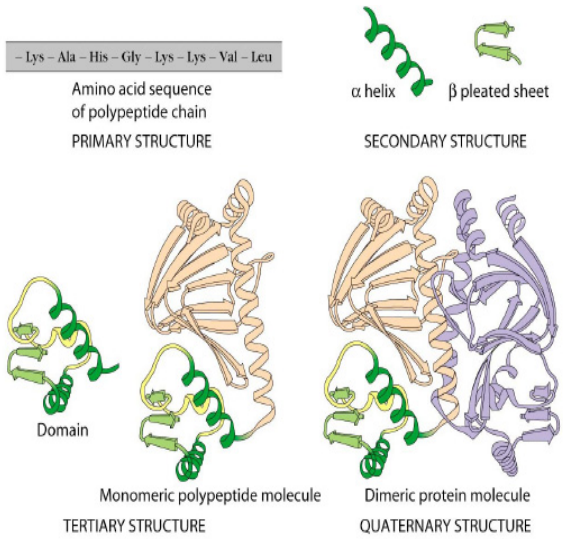


Figure 1.4. Primary, secondary and tertiary structures of a peptidic chain.

1.2.1 Primary structure of peptides and proteins

The primary structure is the sequence of amino acid residues. The knowledge of the sequence of amino acids in a protein can offer insights into its three-dimensional structure and its function, cellular location, and evolution. Thousands of sequences are known and available in databases accessible through the Internet. A comparison of a newly obtained sequence with this large bank of stored sequences often reveals relationships both surprising and enlightening.

i Amino Acid Sequences Deduced by Edman degradation

The sequence of short polypeptides is obtained using automated procedures: several protocols are available but a chemical method devised by Pehr Edman[2] is usually employed. The Edman degradation procedure labels and removes only the amino-terminal residue from a peptide, leaving all other peptide bonds intact.

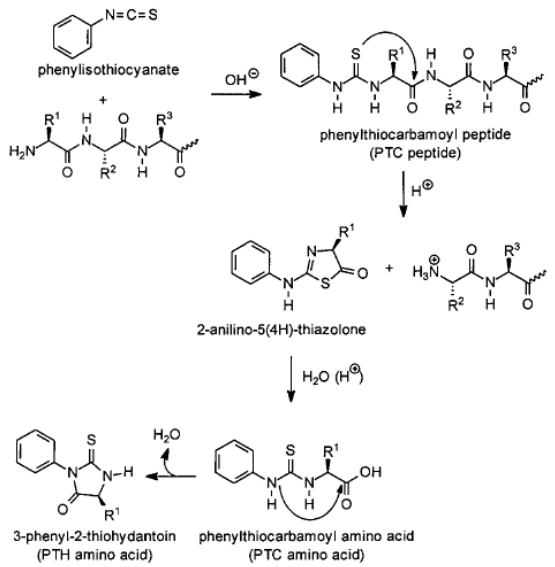


Figure 1.5. The Edman method for the N-terminal stepwise peptide degradation.

The overall accuracy of amino acid sequencing generally declines as the length of the peptide increases. The very large polypeptides or proteins must be broken down into smaller pieces to be sequenced efficiently. There are several steps in this process. First, eventual disulfide bonds must be irreversibly broken, (if not, the peptide containing a cysteine involved in the S-S bonds, may remain attached to the other polypeptide strand containing the second cysteine residue). Then, the protein is cleaved into a set of specific fragments by chemical or enzymatic methods. Several methods can be used for fragmenting the polypeptide chain. Enzymes called proteases catalyze the hydrolytic cleavage of peptide bonds. Some proteases cleave only the peptide bond adjacent to particular amino acid residues and thus fragment a polypeptide chain in a predictable and reproducible way. A number of chemical reagents also cleave the peptide bond adjacent to specific residues.

Among proteases, the digestive enzyme trypsin catalyzes the hydrolysis of only peptide bonds in which the carbonyl group is contributed by either a Lys or an Arg residue, regardless of the length or amino acid sequence of the chain. The number of smaller peptides produced by trypsin cleavage can thus be predicted from the total number of Lys or Arg residues in the original polypeptide, as determined by hydrolysis of an intact sample. A polypeptide with five Lys and/or Arg residues will usually yield six smaller peptides on cleavage with trypsin. Moreover, all except one of these will have a carboxyl-terminal Lys or Arg.

Table 1.2. Specificity of some common methods to cleave polypeptide chains (* all reagents except cyanogen bromide are proteases; † residues furnishing the primary recognition point for the protease or reagent: peptide bond cleavage occurs on either the carbonyl (C) or the amino (N) side of the indicated amino acid residues.

*Reagent (biological source)**	*Cleavage points*†
Trypsin (bovine pancreas)	Lys, Arg (C)
Submaxillarus protease (mouse submaxillary gland)	Arg (C)
Chymotrypsin (bovine pancreas)	Phe, Trp, Tyr (C)
Staphylococcus aureus V8 protease (bacterium *S. aureus*)	Asp, Glu (C)
Asp-*N*-protease (bacterium *Pseudomonas fragi*)	Asp, Glu (N)
Pepsin (porcine stomach)	Phe, Trp, Tyr (N)
Endoproteinase Lys C (bacterium *Lysobacter enzymogenes*)	Lys (C)
Cyanogen bromide	Met (C)

The fragments produced by cleavage are then separated by chromatographic or electrophoretic methods. Each peptide fragment resulting from the action of trypsin is sequenced separately by the Edman procedure.

Finally, the order in which the fragments appear in the original protein must be determined and disulfide bonds (if any) located: ordering peptide fragments is done by cleaving into fragments the same polypeptide using a different enzyme or reagent. The fragments resulting from this second procedure are then separated and sequenced as before. Overlapping peptides obtained from the second fragmentation yield the correct order of the peptide fragments produced in the first.

ii Amino Acid Sequences Deduced by c-DNA

With the development of rapid DNA sequencing methods, the elucidation of the genetic code, and the development of techniques for isolating genes, researchers can deduce the sequence of a polypeptide by determining the sequence of nucleotides in the gene that codes for it. The techniques used to determine protein and DNA sequences are complementary. When the gene is available, sequencing the DNA can be faster and more accurate than sequencing the protein. Most proteins are now sequenced in this indirect way. DNA sequencing does not provide every information (the location of disulfide bonds, for example).

iii Amino Acid Sequences Deduced by Mass spectrometry

New methods based on mass spectrometry permit the sequencing of short polypeptides (20 to 30 amino acid residues) in a few min. [3,4] Peptide sequencing is usually achieved by tandem mass spectrometry. (see chapter 4).

1.2.2 Secondary structure of peptides and proteins

Secondary structures are the recurring structural patterns and arrangements of amino acid residues in a segment of a polypeptide chain, in which each residue is spatially related to its neighbors in the same way. The most common secondary structures are the α-helix, the ß-platelet sheets and ß-turns. The nature of the peptide bond in the polypeptide backbone generates constraints on the structure. It has a partial double-bond character that keeps the entire six-atom peptide group in a rigid planar configuration. The N-C_α and C_α-C bonds can rotate to assume bond angles of φ and ψ, respectively. The secondary structure of a polypeptide segment can be completely defined if the ψ and φ angles are known for all amino acid residues in that segment.[5] (Figure 1.6)

Regular secondary structures, α helix and ß sheet, are a direct consequence of the amino acid chain which collapse into a compact space. The α helix is a type of regular secondary structure in which successive amino acids adopt the same φ and ψ angles (peptide bonds all trans). It is a coiled structure characterized by 3.6 residues per turn, and translating along its axis 1.5 Å per amino acid. The screw sense of alpha helices is always right-handed. (Figure 1.6, B). ß sheet is also a regular secondary structure formed by successively repeated φ and ψ angles. Importantly, however, the H-bonding pattern is not regularly spaced with respect to the amino acid sequence. H-bonds span between amino acids on separate beta strands, which may be quite distant from each other in the sequence. (Figure 1.6, C)

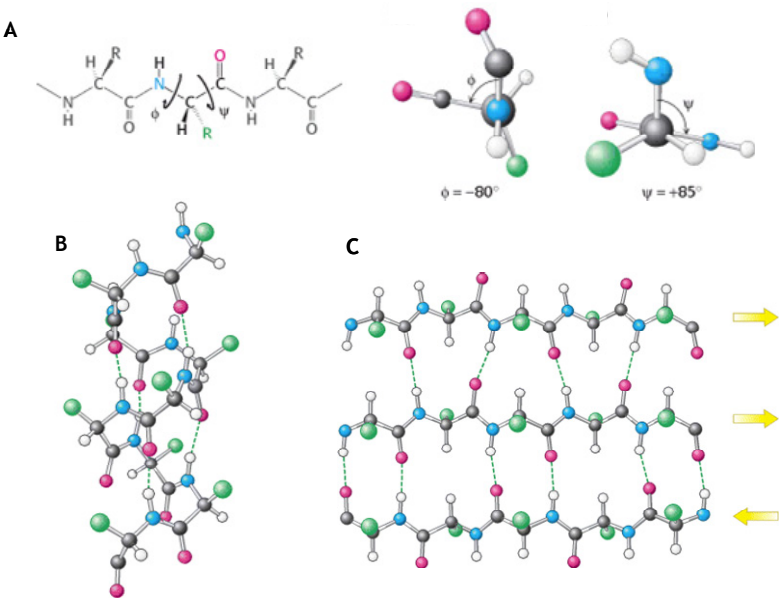


Figure 1.6. (A) Planes formed on the two sides of a Cα and Angles of Rotation of Peptide Bonds (φ and ψ). (B) α-helix. (C) ß-sheets (parallel and anti-parallel.

The secondary structure may be determined by circular dichroism (CD), infrared spectroscopy (IR), NMR, etc.

i. Circular Dichroism

Linear polarized light consists of two circular polarized components of opposite helicity but identical frequency, speed, and intensity. When linear polarized light passes through an optically active medium, for instance a solution containing one enantiomer of an optically active compound, the speed of the light in matter is different for the left and right circular polarized components (different refractive indexes).[6] In such a case a net rotation of the plane of polarization is observed for the linear polarized light. Consequently, enantiomerically pure or enriched optically active compounds can be characterized by the optical rotation index and optical rotatory dispersion. CD is a method of choice for the quick determination of protein and peptide mean secondary structure content. Proteins are often composed of the two classical secondary structure elements, α-helix and ß sheet, in complex combinations. Besides these ordered regions, other parts of the protein or peptide may exist in a random coil state. CD spectroscopy is a highly sensitive method that is able to distinguish between α -helical, ß -sheet and random coil conformations.

i Infrared Spectroscopy

Fourier transform infrared spectroscopy (FT-IR) is nowadays commonly applied to peptides and proteins, and is mainly used to estimate the content of secondary structure elements. IR spectroscopy allows monitoring of the exchange rate of amide protons, and hence provides collective data for all amino acid residues present in a protein or peptide. Polarized IR spectroscopy provides information on the orientation of parts of a protein molecule present in an ordered environment.

1.2.3 Tertiary structure of peptides and proteins

Every protein has a three-dimensional structure that reflects its function. The overall three-dimensional arrangement of all atoms in a protein is referred to as the protein tertiary structure. While the term “secondary structure” refers to the spatial arrangement of amino acid residues that are adjacent in the primary structure, tertiary structure includes longer-range aspects of amino acid sequence. Amino acids that are far apart in the polypeptide sequence and that reside in different types of secondary structure may interact within the completely folded structure of a protein by different peptide bond types: hydrogen bond, ionic bonds, hydrophobic interactions, dipole-dipole interactions and covalent disulphide bond (Figure 1.7).

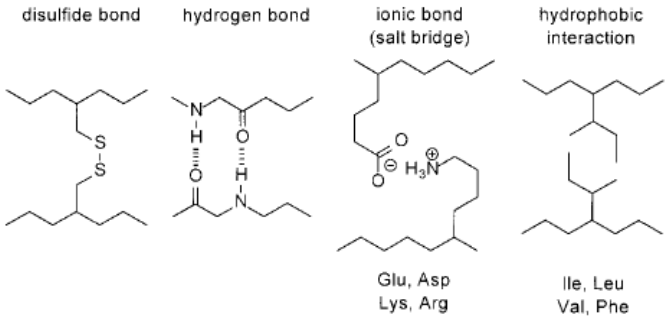


Figure 1.7. Stabilizing interchain interactions between amino acid side chains.

There are two general classes of proteins based on their tertiary structure: fibrous and globular. Fibrous proteins serve mainly structural roles and have simple repeating elements of secondary structure. Globular proteins have more complicated tertiary structures, often containing several types of secondary structure in the same polypeptide chain.

All proteins begin their existence on a ribosome as a linear sequence of amino acid residues. This polypeptide must fold during and following synthesis to take up its native conformation. A loss of three-dimensional structure sufficient to cause loss of function is called denaturation. Some denatured proteins can re-nature spontaneously to re-form biologically active protein, showing that protein tertiary structure is determined by amino acid sequence.[7]

The three-dimensional structure of a peptide or a protein is the crucial determinant of its biological activity. It can be determined by X-ray diffraction, NMR and fluorescence spectroscopy.

i X-Ray Diffraction

X-ray crystallography determines the arrangement of atoms within a crystal from the manner in which a beam of X-rays is scattered from the electrons within the crystal. The method produces a three-dimensional picture of the density of electrons within the crystal, from which the mean atomic positions, their chemical bonds, their disorder and sundry other information can be derived. The wavelength range of X-rays corresponds to the size of the diffracting structures (atomic radii and lattice constants), and the observed diffraction pattern results from a superposition of the diffracted beams. The spacing of atoms can be determined by measuring the locations and intensities of spots produced on a photographic film by a beam of X rays of a given wavelength, after the beam has been diffracted by the electrons of the atoms. Once a crystal of the protein is obtained, it is placed in an X-ray beam between the X-ray source and a detector, and a regular array of spots called reflections is generated. The spots are created by the diffracted X-ray beam, and each atom in a molecule makes a contribution to each spot. An electron-density map of the protein is reconstructed from the overall diffraction pattern of spots

by using a Fourier transform. In effect, the computer acts as a "computational lens." A model for the structure is then built that is consistent with the electron-density map.

ii Nuclear Magnetic Resonance

An important complementary method for determining the three-dimensional structures of macromolecules is nuclear magnetic resonance (NMR).[8] Modern NMR techniques are being used to determine the structures of ever-larger macromolecules, including carbohydrates, nucleic acids, and small to average-sized proteins. An advantage of NMR studies is that they are carried out on macromolecules in solution, whereas X ray crystallography is limited to molecules that can be crystallized. NMR can also provide information on the dynamic side of the protein structure, including conformational changes, protein folding, and interactions with other molecules. NMR is based on the nuclear spin angular momentum, a quantum mechanical property of atomic nuclei. Only certain atoms, such as ^{1}H, ^{13}C, ^{15}N, ^{19}F, and ^{31}P, possess the kind of nuclear spin that gives rise to an NMR signal. Nuclear spin generates a magnetic dipole. When a strong, static magnetic field is applied to a solution containing a single type of macromolecule, the magnetic dipoles are aligned in the field in one of two orientations, parallel (low energy) or antiparallel (high energy). A short (~10 μs) pulse of electromagnetic energy of suitable frequency (the resonant frequency, which is in the radio frequency range) is applied at right angles to the nuclei aligned in the magnetic field. Some energy is absorbed as nuclei switch to the high-energy state, and the absorption spectrum that results contains information about the identity of the nuclei and their immediate chemical environment. The data from many experiments performed on a sample are averaged, increasing the signal-to-noise ratio, and an NMR spectrum is generated. Structural analysis of proteins became possible with the advent of two-dimensional NMR techniques. These methods allow the measurement of the distance-dependent coupling of nuclear spins in nearby atoms through space (the nuclear Overhauser effect (NOE-NOESY) or the coupling of nuclear spins in atoms connected by covalent bonds (total correlation spectroscopy, or TOCSY).

iii UV Fluorescence Spectroscopy

Fluorescence spectroscopy is widely used in peptide and protein chemistry, either observing intrinsic fluorophors (Trp, Tyr), fluorescent cofactors, or extrinsic fluorophors that are used to label the protein.[9] As fluorescence spectroscopy involves electronic transitions, it can be applied to study very fast processes. Interactions of proteins with other proteins, nucleic acids, small ligands, and membranes can be monitored, as well as protein folding and conformational transitions. The association of peptides can be monitored by fluorescence quenching. Usually, fluorescence intensity, anisotropy, and emission wavelength may be observed.

1.2.4 Quaternary structure of proteins of peptides and proteins

When a protein has two or more polypeptide subunits, their arrangement in space is referred to as quaternary structure. The intermolecular interactions are of the same nature as for the

tertiary structure (hydrogen bond, ionic bonds, hydrophobic interactions, dipole-dipole interactions and covalent disulphide bond). To study a protein in detail, it is necessary to separate it from other proteins and to apply the proper techniques. In considering these higher levels of structure, it is useful to classify proteins into two major groups: fibrous proteins, having polypeptide chains arranged in long strands or sheets, and globular proteins, having polypeptide chains folded into a spherical or globular shape.

1.2.5 Post-translational modification (PTMs) of peptides and proteins

Proteins often undergo several post-translational modification steps in parallel to protein folding. These modifications can be transient or of a more permanent nature. Most modifications are, however, susceptible to alteration during the lifespan of proteins. Post-translational modifications thus generate variability in proteins that are far beyond that provided by the genetic code. Co- and post-translational modifications can convert the 20 specific codon-encoded amino acids into more than 100 variant amino acids with new properties. Post-translational covalent modifications occurring in nature include acetylation, hydroxylation, methylation, glycosylation, sulfatation, iodination, carboxylation, phosphorylation, nucleotidylation, ADP-ribosylation, and numerous other types. This post-translational processing is an essential prerequisite for mature polypeptides or proteins. Normally, the translation product cannot be considered as a functional protein. After the assembly of the complete sequence of a protein, some of the amino acid building blocks may be involved in post-translational modifications. Most modifications are performed after release of the polypeptide from the ribosome, but modifications such as disulfide bridge formation or N-terminal acetylation very often occur in the nascent polypeptide chain. Knowledge of these modifications is extremely important because they may alter physical and chemical properties, folding, conformation distribution, stability, activity, and consequently, function of the proteins. Moreover, the modification itself can act as an added functional group. Conjugated proteins are classified on the basis of the chemical nature of their prosthetic groups; for example, lipoproteins contain lipids, glycoproteins contain sugar groups, and metalloproteins contain a specific metal. A number of proteins contain more than one prosthetic group. Usually the prosthetic group PTMs of a protein can determine its activity state, localization, turnover, and interactions with other proteins, thus PTMs play an important role in the protein biological function.[10],[11]

Various methods used in proteomics, such as 2D gel electrophoresis, 2D liquid chromatography, mass spectrometry, affinity-based analytical methods, interaction analyses, ligand blotting techniques, protein crystallography and structure–function predictions, are all applicable for the analysis of the secondary modifications. [12]

1.3 Peptide and protein synthesis

There are two ways to obtain a peptide with a good yield: (i) Recombinant DNA Techniques (genetic engineering); (ii) Direct chemical synthesis.

i. Recombinant DNA Techniques.

Bacteria such as *Escherichia coli* can accept genetic material from other microorganisms and transmit it to their successors, and can also serve as recipients of genetic material from either plants or animals. Gene expression includes synthesis of the corresponding mRNA (transcription) and synthesis of the protein (translation). The biosynthesis of a foreign gene product (protein) in an organism relies on a recombination of the genetic material of the microorganism with the DNA fragment encoding for the desired protein. The process includes the following steps: isolation of the encoding DNA fragment from the donor organism, insertion of the DNA into a vector, transfection of the vector into the host organism, cultivation of the host organism (cloning), which leads to gene amplification, mRNA synthesis, and protein synthesis and isolation of the recombinant protein. This synthesis can provide a high protein yield, can be automated and is stereospecific.

i Direct chemical synthesis

The major breakthrough in this technology was provided by R. B. Merrifield in 1962. His innovation involved synthesizing a peptide while keeping it attached at one end of a solid support. The support is an insoluble polymer (resin). The free N-terminal amine of a solid-phase attached peptide is coupled to a single N-protected amino acid unit. This unit is then deprotected, revealing a new N-terminal amine to which a further amino acid may be attached. This general principle of Solid-Phase Peptide Synthesis (SPPS) is repeated for each cycle of coupling-deprotection. (Figure 1.8) At each subsequent step in the cycle, protecting chemical groups block unwanted reactions, i.e. at side chains of amino acids. Unlike ribosome protein synthesis, solid-phase peptide synthesis proceeds in a C-terminal to N-terminal fashion.

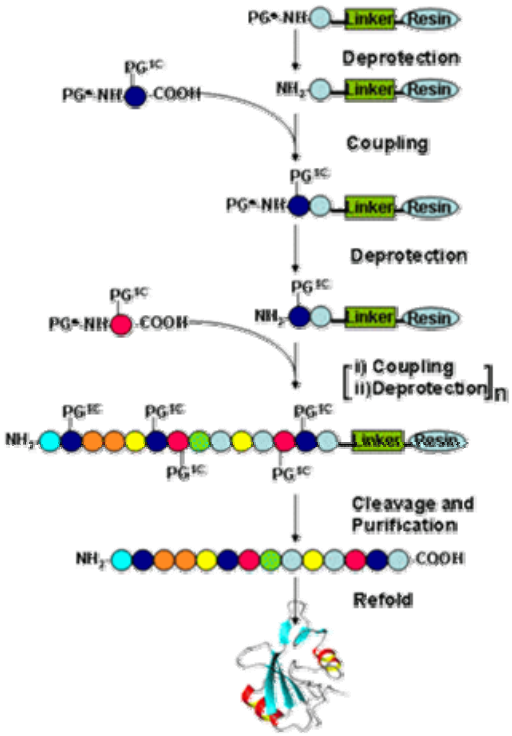


Figure 1.8. Process of solid-phase synthesis. PG=temporary protecting group; each circle, 1 peptide.

There are two main strategies in SPPS: Fmoc and Boc. These two groups can be used to protect the N-termini of amino acid monomers and are to be removed at every cycle, thus they are called "temporary" protecting groups. t-Boc (or Boc) stands for (tert)-Butyloxycarbonyl. In order to remove the Boc group from a growing peptide chain, acidic conditions are used (usually, TFA). The removal of the side-chain protecting groups (usually based on the benzyl moiety, such as benzyl esters or benzyl carbamates, thus stable in TFA) and of the peptide from the resin at the end of the synthesis is achieved by incubating in hydrofluoric acid or, more commonly today, trifluoromethanesulfonic acid; Boc chemistry thus usually involves harsh synthetic conditions. Fmoc stands for 9H-fluoren-9-ylmethoxycarbonyl. In order to remove a Fmoc from a growing peptide chain, basic conditions are used (usually, 20 % piperidine in DMF). The removal of the side-chain protecting groups (usually based on the tert-butyl moiety, such as BOC or t-butyl esters, thus stable in piperidine) and of the peptide from the resin is achieved by incubating in trifluoroacetic acid (TFA), The main advantage of Fmoc chemistry is that milder synthetic conditions can be used.

The technology for chemical peptide synthesis can be also automated and the synthesis of proteins up to 100 amino acid residues is possible.

1.4 References

[1] Berg J. M., Tymoczko J. L., Stryer L., *Biochemistry*, ed. W. H. Freeman, 6th Edition, **2006**

[2] Edman, P. A method for the determination of the amino acid sequence of peptides. *Arch. Biochem. Biophys.* **1949.** 22:475–483.

[3] Domon B. and Aebersold R. "Mass spectrometry and protein analysis", Science, **2006**, 312, 212

[4] Steen H. and Mann M., The abc's (and xyz's) of peptide sequencing, *Nature Reviews, Molecular Cell Biology*, **2004**, 5: 699-711

[5] http://restools.sdsc.edu/biotools/biotools9.html

[6] Woody R.W. *Circular Dichroism, Principles and Applications*, (Eds.)K. Nakanishi, N. Berova, R.W. Woody, VCH, New York, **1994**

[7] Chemical physics of protein folding BROOKS III C. L., GRUEBELE M., ONUCHIC J.N., AND WOLYNES P. G., *Proc. Natl. Acad. Sci. USA* , **1998,** 95: 11037–11038

[8] Montelione G.T., Zheng D., Huang Y. J., Gunsalus K. C., Szyperski T., Protein NMR spectroscopy in structural genomics., *Nature Struct. Biol*., **2000**, 982-5

[9] Brown, M. P., and C. Royer. Fluorescence spectroscopy as a tool to investigate protein interactions. *Curr. Opin. Biotechnol.* **1997.** 8:45–49.

[10] Cohen, P. The regulation of protein function by multisite phosphorylation—a 25 year update. *Trends Biochem. Sci*. **2000** 25: 596–601.

[11] Ghezzi P., Di Simplicio P., Glutathionylation pathways in drug response, *Current Opinion in Pharmacology*, **2007,** 7, Issue 4, 398-403

[12] Baumann M., Meri S., Techniques for studying protein heterogeneity and post-translational modifications, *Expert Review of Proteomics*, **2004**, 1,n.. 2, 207-217

2 PEPTIDES AND PROTEINS IN FOOD

Proteins are fundamental and integral food components, both from a nutritional and functional points of view.

2.1 Nutritional value of peptides and proteins

Next to water, proteins are the major components of body tissues and so they are the essential nutrients for growth. The body is in a dynamic state, with proteins and other nitrogenous compounds being degraded and re-synthesized continuously. More proteins are turned over daily in the body than are ordinarily consumed in the diet. In fact, proteins are regularly digested into their basic components (amino acids) and amino acids can be reassembled in the body by various mechanisms into tissues, organs, enzymes and a host of other protein based regulatory compounds: hormones, enzymes, antibodies, etc., Nutritionally, proteins are a source of energy and amino acids, which are essential for growth and maintenance. The nutrient value of proteins is also due to their use for energy production when required. Specifically, proteins supply approximately 4 calories per gram, about the same as sugars.[1]

Proteins in human diet are derived from two main sources, namely animal proteins (e.g. egg, milk, meat and fish) and plant proteins (e.g. pulses, cereals, nuts, beans and soy products). Animal proteins are more "complete" than vegetable proteins with regards to their amino acid composition. From this point of view, the Biological Value (BV) of a protein is a measure of how efficiently food proteins, once absorbed in the gastrointestinal tract, can be turned into body tissues, so that the term "biologically complete proteins" refers to foods that contain all the essential amino acids needed by the body, whereas, incomplete proteins refers to foods lacking one or more essential amino acids. There are more complete proteins from animal sources than most vegetable proteins, which are "biologically incomplete". However, an incomplete protein can be converted into a complete protein if two incomplete proteins are added together by employing what is called "complementarity of proteins". Two plant proteins such as legumes and grains or legumes and nuts/seeds can be mixed to produce a complete protein from two incomplete ones.

2.2 Biological effect of peptides and proteins

Proteins affect the physicochemical and sensory properties of various proteinaceous foods. Moreover, many dietary proteins possess specific biological properties, which make these components potential ingredients of functional or health-promoting foods.

2.2.1 Bioactive peptides

Biologically active peptides are of particular interest in food science and nutrition because they have been shown to play physiological roles, including opioid-like features, as well as immunostimulating and anti-hypertensive activities, and ability to enhance calcium absorption.

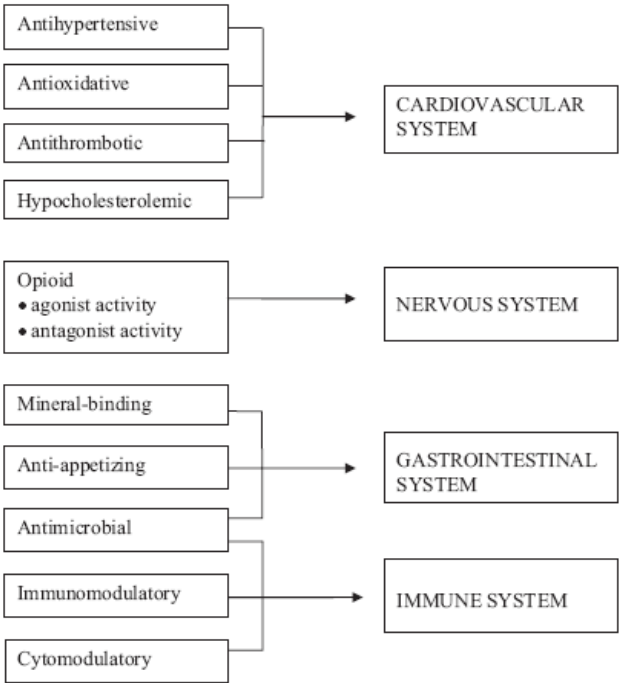


Figure 2.1, Physiological functionality of food derived bioactive peptides

Increasing attention is being focused on physiologically active peptides derived from milk[2], soy[3,4], egg[5,6], meat[7], fish, and other proteins.

Many bioactive peptides have in common structural features that include a relatively short length (e.g. 2-9 amino acids), the presence of hydrophobic amino acid residues in addition to proline, lysine or arginine groups. Bioactive peptides are also resistant to the action of digestion peptidases. Antihypertensive peptides, known as Angiotensin I Converting Enzyme (ACE) inhibitors have been found in milk[8], corn and fish. Peptides with opioid activities are derived from wheat gluten or casein, following digestion with pepsin. Exorphins, or opioid peptides derived from food proteins such as wheat and milk (e.g. exogenous sources) have a similar structure to endogenous opioid peptides, with a tyrosine residue located at the amino terminal site. Immunomodulatory peptides derived from tryptic hydrolysates of rice and soybean proteins act to stimulate superoxide anions (reactive oxygen species-ROS), which trigger non-specific immune defense systems. Antioxidant properties that prevent peroxidation of essential fatty acids have also been shown for peptides derived from milk proteins. The addition of a Leu or Pro residue to the N-terminus of a His-His dipeptide has been shown to enhance the antioxidant

activity and to facilitate further synergy with non-peptide antioxidants (e.g. BHT). The tryptic digests of casein yielding caseinophosphopeptides exhibit both hydrophilic and lipophilic antioxidant activity due to both metal ion sequestering properties and quenching of ROS.

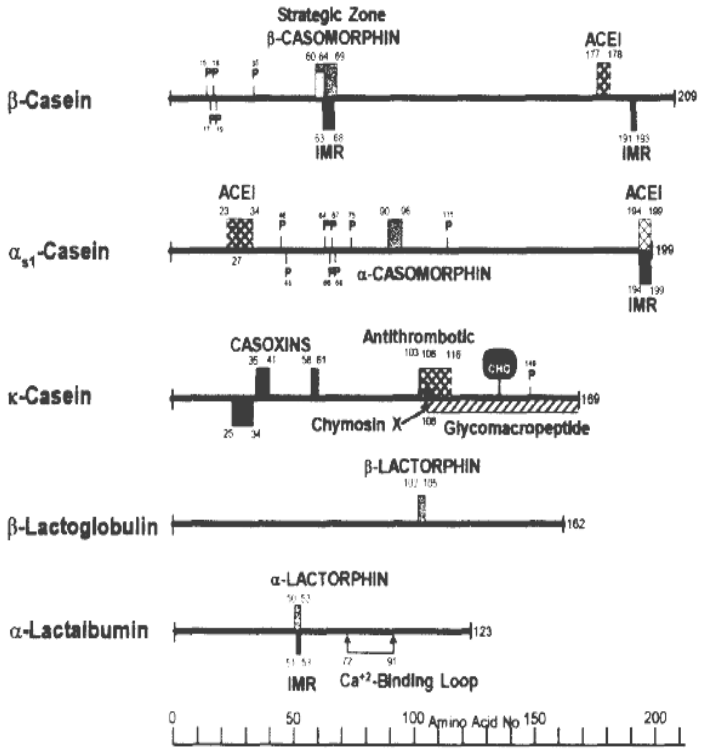


Figure 2.2. Map of some latent bioactive peptides in bovine caseins. PP: mineral binding, ACEI: ACE inhibitory, IMR: immunomodulatory region. (adapted from 8).

These peptides are inactive within the sequence of the parent protein molecule and can be liberated by gastrointestinal digestion, fermentation with proteolytic starter cultures or hydrolysis by proteolytic enzymes obtained by industrial-scale technologies.

2.2.2 Bioactive proteins

There are examples of biologically active food proteins, with physiological significance beyond the pure nutritional requirements that concern available nitrogen for normal growth and maintenance[9]. These biologically active proteins survive in the upper gastrointestinal digestive processes and act downstream both in the lumen and on the mucosal surfaces. These components can have effects not only locally but also systemically beyond the gut.

Biologically active proteins have, for example, antimicrobial actions or growth factor activity. Biologically active proteins have been studied in milk: lactoferrin is a natural and versatile bioactive protein that interacts with the immune system, binding minerals such as iron, and

having antimicrobial properties. Glycomacropeptide from k casein has antimicrobial effects and it is found as intact peptone in whey (ricotta cheese).[10]

Colostrum, the first milk from the cow, is rich in immunologically active proteins such as immunoglobulins and growth factors: immunoglobulins bind to pathogens or toxins, growth factors, including insulin-like growth factor 1 (IGF-1), transforming growth factor b, and related peptides, bind to mucosal receptors and trigger second messengers. This seems to exert effects on mucosal cell growth and renewal.

A biological and important effect of food proteins as intact units, is the trigger of allergic reactions of the immune system.

2.3 Food allergy

Food allergies affect as many as 6% of young children, most of whom "outgrow" the sensitivity, and about 2% of the general population. Although any food may provoke a reaction, relatively few foods are responsible for the vast majority of food allergic reactions: milk, egg, peanuts, tree nuts, fish, shellfish, celery, sesame, lupin, soy, cereals and mustard.

CLASSIFICATION OF ADVERSE FOOD REACTIONS

Adverse reactions to foods, aside from those considered toxic, are caused by a particular individual intolerance towards commonly tolerated foods. Intolerance derived from an immunological mechanism is referred to as Food Allergy, the non-immunological form is called Food Intolerance.[11]

In 1995, the European Academy of Allergy and Clinical Immunology (EAACI) published a position paper[12] that classified adverse reactions to food as:

- Toxic (adverse reactions that occur in any exposed individual provided that the dose is high enough)
- Nontoxic (adverse reactions that depend on individual susceptibility to a certain food)
  - → Immune-mediated (food allergy)
    - IgE-mediated
    - Non-IgE-mediated
  - → non-immune-mediated (food intolerance)
    - enzymatic (e. g. lactase deficiency)
    - pharmacological (abnormal reactivity to substances such as vasoactive amines normally present in some foods)
    - undefined (e.g. food additives intolerance).

Reactions due to *toxic* components occurring naturally in the foodstuff or being present as contaminating agents are developed in *anyone* given a high enough dose of the toxin is high enough. Natural toxins from plants may be both endogenous (such as glucosinolates in cabbage which have a goitrogenic effect) and exogenous (such as aflatoxins, found in peanuts and grains contaminated with mould which can cause encephalopathy, hallucinations, and hepatic disease). [13]

Nontoxic reactions depend on the individual susceptibility and are based on IgE- or non-IgE-mediated immune mechanisms (allergy or hypersensitivity) or non-immune mechanisms (intolerance).

Food allergy is an immunologic reaction resulting from the ingestion of a food or food additive. This reaction occurs only in some individuals, may occur after only a small amount of the substance is ingested, and is unrelated to any physiologic effect of the food or food additive. Food allergic reactions are responsible for a variety of symptoms involving skin, the gastrointestinal tract and the respiratory tract and may be due to IgE-mediated and non–IgE-mediated mechanisms. In fact, with regards to immune mediated reactions, the role of the so-called *type I food allergies* with IgE as mediator has been better investigated. *Non-IgE-mediated* food allergy could consist of immune reactions depending on antibodies of different isotypes from IgE (i.e. IgG, IgM and IgA); immune complexes formed by food and food antibodies; cell-mediated immunity.

Food intolerances are most frequently observed among food adverse reactions and depend on metabolic characteristics of the patient (e.g. milk intolerance due to primary or secondary lactase deficiency), pharmacologic properties of the ingested food (e.g. headaches due to vasoactive mono- or diamines; caffeine in coffee, tyramine in aged cheeses), nonspecific histamine-releasing properties of food (e.g. high contents of lectins, prolamin, peptones, polyamines; e.g. histamine in scombroid fish poisoning, toxins secreted by *Salmonella*, *Shigella* and *Campylobacter*), and undefined intolerances due to idiosyncratic responses.

There are some differences between nomenclature and classifications of adverse food reactions used in Europe and in the USA[14], that is why a specific committee of the World Allergy Organization (WAO) updated the EAACI 2001 position statement in 2003[15]. This document states that the term **"food allergy"** is appropriate when immunological mechanisms have been demonstrated and that the term "IgE-mediated food allergy" should be used if IgE is involved in the reaction. All other reactions should be referred as ‚nonallergic food hypersensitivity'.

There seems to be some crude correlation between the consumption of a food item, and the occurrence of food allergy to that item in certain regions of the world. Among the examples are fish allergy from Scandinavia and coastal North America, allergy to crustaceans on the Mexican Golf coast in Louisiana, peanuts in the USA in general, and *prunoideae* fruits in the Mediterranean region. Despite the globalization of the food market, these differences still exist, and despite the enormous variety of different nutrients found in the daily human diet, only relatively few proteins account for most allergic reactions observed in patients, and this still remains still to be explained.

2.3.1 Mechanisms of food allergy

To date the immunological mechanisms and the site of sensitization to food antigens remain unclear. Several routes can be considered for food antigens to get in contact with the cells of the immune system[16]. There are two types of allergic or hypersensitivity reactions occurring as basic immunological mechanisms involved in food allergies.[17] IgE-mediated reactions, also known as immediate hypersensitivity reactions, involve the formation of IgE antibodies that specifically recognize certain allergens in foods. IgE-mediated reactions are the most important type of food allergy because these reactions involve a wide variety of different foods and the reactions can be severe in some individuals. IgE-mediated mechanisms are also responsible for allergic

reactions to pollens, mold spores, animal dander, insect venoms, and certain drugs; only the source of the allergen differs. The substances which lead to secretion of IgE antibodies are called allergens.

Cell-mediated reactions or delayed hypersensitivities probably play an important, although as yet undefined, role in food hypersensitivity.

2.3.1.1 IgE-mediated allergic reaction (immediate hypersensitivity)

Immunoglobulin E is one of five classes of antibodies that are present in the human immune system (the others being IgG, IgM, IgA, and IgD). The normal function of IgE antibodies is protection from parasitic infections. Although all humans have low levels of IgE antibodies, individuals predisposed to the development of allergies are most likely to produce IgE antibodies that are specific for and recognize certain environmental antigens, typically proteins.

IgE antibodies appear in very low serum concentration (0.00005 mg/ml). These allergen-specific IgE antibodies can sensitize mast cells and basophilic granulocytes by binding via Fc portion to high-affinity receptor (FceR1). The ability of multivalent allergens to cross-link these bound IgE molecules will initiate the mediator release from these cells. Soluble IgE antibodies have a short half-life in serum (less than a day), but a markedly prolonged half-life (up to 14 days) when bound to Fc receptors. This binding thus protects IgE antibodies from proteolytic cleavage and clearance. IgE antibody molecules are rather non-flexible: this poses restrictions to the three-dimensional orientation of the epitopes on the allergen molecule. The amino-acid residues responsible for the interaction with the specific IgE molecules can be shown in relation to the overall structure of the allergen. It is clear that the conformational but also linear IgE-binding epitopes can be present for IgE binding on a backbone of structural amino acids that ensure the proper three-dimensional structure of the allergen

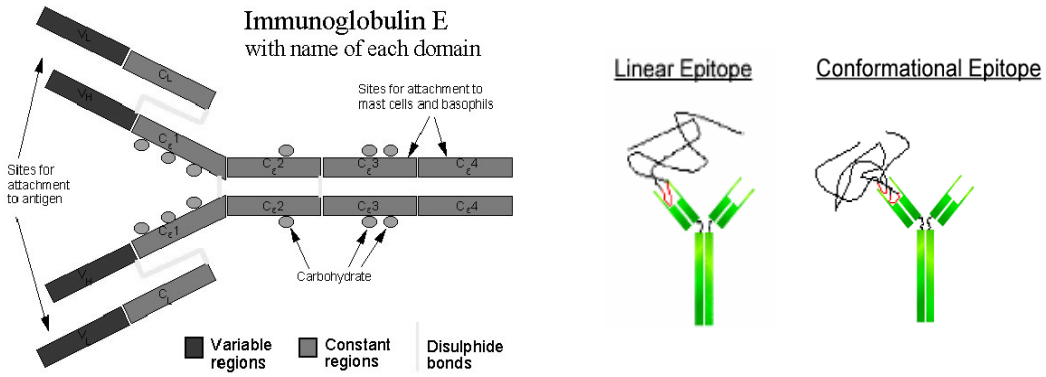


Figure 2.3. Immunoglobulin E structure and Allergen recognition.

i The sensitization phase of an allergic reaction

Allergic sensitization describes the genetically determined propensity of certain individuals to react with the induction of Th2 cells and subsequent allergen-specific IgE-antibody formation upon repeated low-dose exposure of allergens at mucosal surfaces. This allergic sensitization can last for many months to years. Allergic sensitization does not necessarily lead to immediate allergic reactivity.

In IgE-mediated food allergies, allergen-specific antibodies are first produced in response to stimulus of the antibody forming B cells in response to the immunological stimulus created by exposure of the immune system to a specific food allergen, usually a naturally-occurring protein present in the food. The immune response in the small intestine which is responsible for the dominance of the IgE antibody generation is quite complex and involves T helper type 2 cells, interleukin-4 (IL-4), and other factors[18]. IgEs bind to high affinity specific receptors present on the mast cells, fixed cells present in the mucosa and skin, and basophils, circulating in the blood, in a process known as sensitization. The sensitization phase of the allergic reaction is symptomless. In fact, sensitization can occur without the development of clinical reactivity so the demonstration of IgE antibodies directed against a particular food in human blood serum is insufficient evidence for the diagnosis of a food allergy unless it is coupled to a strong history of food allergy or a positive double-blind, placebo-controlled food challenge.

ii The elicitation phase of an allergic reaction

Once sensitized, exposure to the same food allergen on a subsequent occasion can result in an allergic reaction. When this happens, the allergen associates with the IgEs bound to mast cell or basophil and this conjugation triggers a stimulus to these cells, which degranulate, release mediators in the surrounding microenvironment and synthesize new mediators (prostaglandins, leukotriens, cytokines). An immediate reaction follows a few minutes after the contact with the allergen, due mainly to the histamine. At the base of the reaction there is vasodilatation, tissue fluid exudation, smooth muscle contraction and mucous secretion. A late-phase response follows the immediate reaction which begins 4–6 hours after contact with the allergen and continues for several days. This response is caused by chemotactic mediators released at the same time as the immediate reaction which promotes selective recruitment of inflammatory cells, mainly eosinophils and neutrophils, which infiltrate the tissue producing an inflammation lasting a few days. The two clinical elements required to support an IgE-mediated food allergy are the presence of IgE specific antibodies to the culprit food and a proven relationship between ingestion of the food and the appearance of the symptoms.

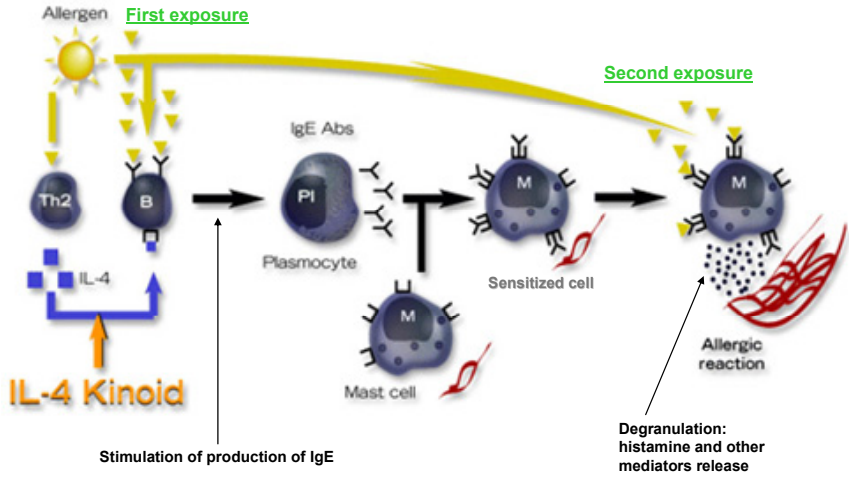


Figure 2.4. Allergic IgE mediated reaction at a cellular level.

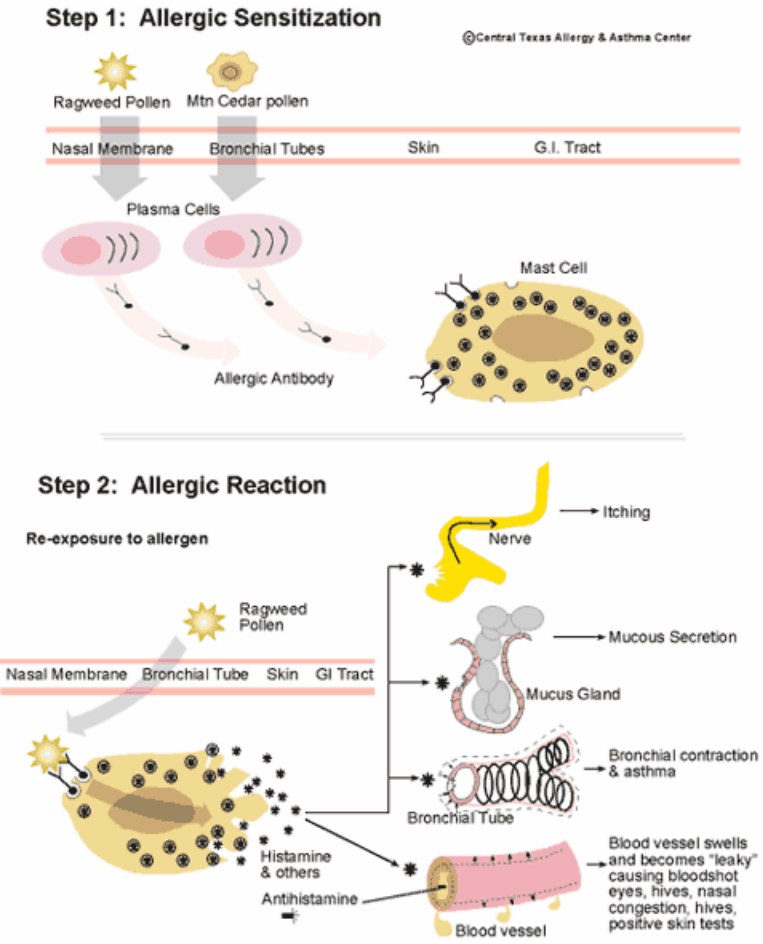


Figure 2.5.Allergic IgE mediated reaction at an organism level.

On the basis of the clinical presentation, physical/chemical features of allergens responsible for the allergic reaction, and underlying immunologic mechanisms, two forms of food allergy can be distinguished. In Class 1, food allergy sensitization occurs through the intestinal tract and is often caused by stable proteins. In contrast, Class 2 food allergy develops as a consequence of sensitization to airborne allergens.

2.3.1.2 Cell-mediated reactions (delayed hypersensitivity)

These reactions develop slowly, reaching a peak at approximately 48 hours and subsiding after 72–96 hours. The mechanisms of cell-mediated food allergies are not well understood. They involve an interaction between specific food allergens and sensitized T lymphocytes. Lymphocyte stimulation initiates the release of cytokines and lymphokines which produces a localized inflammatory response. Antibodies are not involved in these reactions.

T lymphocytes are a major component of the gut-associated lymphoid tissue. Except for celiac disease, evidence for the involvement of cell-mediated immune reactions in food allergies remains incomplete. However, cell-mediated reactions appear to be involved in some cases of cows' milk allergy occurring especially in infants and with symptoms confined primarily to the gastrointestinal tract. No estimates of the prevalance of cell-mediated food allergies have been made. Celiac disease, also known as celiac sprue or gluten-sensitive enteropathy, appears to be an example of a cell-mediated reaction. Celiac disease is a malabsorption syndrome occurring in sensitive individuals upon the consumption of wheat, rye, barley, triticale, spelt, and kamut.

2.3.2 Absorption of food allergens

Intestinal absorption of food antigens and immune responses are highly dependent, and the nature of the antigen can dictate the route and the type of immune response generated. The gastrointestinal tract with its 400 m^2 of surface area is the largest immune organ of the human body and represents by far the largest site of exposure with pathogens and exogenous soluble antigens. Digestion and absorption of nutrients are the main functions exerted in the gastrointestinal lumen or at mucosal sites.

The gastrointestinal tract forms an extensive barrier to the outside environment and provides a surface to process and absorb ingested food and to discharge waste products. The immune system associated with this barrier, the gut-associated lymphoid tissue, is capable of discriminating among harmless foreign proteins or dangerous pathogens.[19]

Despite the evolution of an elegant barrier system, about 2% of ingested food antigens are absorbed and transported throughout the body in an "immunologically intact" form, even through the mature gut.[20]

Clinical aspects of food allergy

Food allergies cause a number of clinical conditions involving the gastrointestinal tract, the skin, the airways or the most dangerous of all allergic reactions, anaphylaxis.21

Oral allergy syndrome (OAS)

Oral allergy syndrome is a particular type of IgE-mediated contact urticaria involving lips, oral mucosa, and pharynx. Symptoms develop within minutes and typically include local itching of lips, tongue, palate, throat, and/or ears and nose and/or swelling (angioedema) of the same areas.

Gastrointestinal food hypersensitivity reactions

IgE-mediated gastrointestinal problems may present with a variety of symptoms including nausea, vomiting, gastric retention, intestinal hyper-motility, abdominal pain due to colonic spasms and diarrhoea. Symptoms usually develop within minutes to 2 h of the ingestion of the offending food. Food allergens causing gastrointestinal symptoms are generally pepsin-stable, and hence able to reach the gastrointestinal tract in an almost unmodified form or as (assembled) fragments with sufficient residual allergenicity.

Skin disorders

The skin is frequently involved in IgE-mediated food allergy. Cutaneous symptoms may vary from pruritus, urticaria, and angioedema to morbilliform rashes. Acute urticaria, with or without angioedema, is the most common skin disorder in adult patients with food allergy. Atopic dermatitis is a chronically relapsing inflammatory skin disease commonly associated with the presence of IgE specific for airborne and/or food allergens.

Respiratory disorders

Respiratory symptoms (rhinoconjunctivitis and bronchospasm) may occur in food-allergic patients following the ingestion of the offending foods in association with gastrointestinal and skin disorders but are rarely present as the only symptom.

Anaphylaxis

According to the recent EAACI position paper on nomenclature, anaphylaxis is defined as a severe, life-threatening, generalized or systemic hypersensitivity reaction'. The anaphylactic reaction is the most dramatic allergic reaction and is always a medical emergency. Along with drugs, foods are one of the most common causes of anaphylaxis. Anaphylaxis is caused by the abrupt, massive release of mediators from mast cells and/or basophils throughout the body, inducing gastrointestinal, skin, and respiratory symptoms, in some cases associated with cardiovascular symptoms including hypotension, collapse and dysrhythmia. Patients may react within minutes or even seconds after contact with (traces of) the food, with a generalized, life-threatening reaction characterized by a combination of the following symptoms: generalized urticaria, erythema, itching, nausea, vomiting, dyspnoea due to oedema of the glottis (throat tightness) and/or bronchospasm, dizziness, palpitations, fainting or even collapse. These reactions may be fatal or near-fatal.

2.3.3 Management of food allergy

There is presently no cure for food allergies. Once the diagnosis of food hypersensitivity is established, the only proved therapy remains elimination of the offending allergen, although therapeutic modalities are on the horizon. Sometimes, elimination from the diet of a food can cause nutritional lacks, and so alternatives should be found, such as hypoallergenic formulas of breast and cows milk in the baby food industry. The management of food allergy depends on the study of the characteristics of the allergen: for example, The studies on the allergens' molecular

weights, their locations in foods and their stability (chemical, enzymatic and thermal) allow the development of treatment to decrease the allergenic potential: i.e., heat-treatments for heat labile allergens[22], or chemical lye peeling combined to ultrafiltration to decrease the allergenicity of peach juice[23] and apricot. The determination of the allergenic epitopes brings to the production of recombinant allergens with possible mutations in correspondence of the epitopes, to produce hypoallergenic proteins.[24] Moreover, complexation of the antigenic protein has been studied, by treatment of apples with polyphenol oxidase (PPO) and/or peroxidase (POD) Oxidative reactions, which catalyze the conversion of phenolic compounds present in apple into o-quinones: this treatment seemed to decrease its allergenicity.[25] Scientific efforts were recently rewarded with the development of a hypoallergenic variety of apple (Santana apple[26]).

However, the problem of thresholds remains to be solved. The individual response to an allergen may also be so different that it is very difficult to establish the distinction safe/toxic in this case. This critical question should drive clinical approaches toward the needs of individual patients.

In all cases, no food treatment showed a complete reduction of allergenicity. Moreover, there is possible contamination endangering the allergic patients, when an industrial line of a product or of an ingredient, can be contaminated with a potential allergen. Strict regulations were introduced in Europe in 2003, which brought important changes[27].

EU regulation

Directive 2000/13/EC10 addresses the labeling, presentation and advertising of foodstuffs, and has superseded the older Directive, 79/112/EEC.11 The major elements of this directive are the "need to inform and to protect the consumer and to enact Community rules of a general nature applicable horizontally to all foodstuffs put on the market"; and to promote the use of "detailed labeling, in particular giving the exact nature and characteristics of the product, which enables the consumer to make his choice in full knowledge of the facts". The amendment includes the requirement to label all the ingredients, additives, processing aids and other substances that may cause adverse reactions in consumers. The most common food allergens are found, intentionally or unintentionally, in a wide variety of processed foods, so labeling rules are necessary to ensure that consumers suffering from food allergies receive appropriate information.

A list of allergenic substances has been drafted, and includes those foodstuffs, ingredients and other substances recognized as causing allergy. The list of allergens of Annex IIIa of this directive has been recently implemented by Commission Directives 2005/26/EC and 2007/68/EC and now the list includes:

1. Cereals containing gluten (i.e. wheat, rye, barley, oats, spelt, kamut or their hybridised strains) and products thereof, except: (a) wheat-based glucose syrups including dextrose; (b) wheat-based maltodextrins; (c) glucose syrups based on barley; (d) 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: (a) fish gelatine used as carrier for vitamin or carotenoid preparations; (b) fish gelatine or Isinglass used as fining agent in beer and wine.
5. Peanuts and products thereof.
6. Soybeans and products thereof, except: (a) fully refined soybean oil and fat (1); (b) natural mixed tocopherols (E306), natural D-alpha tocopherol, natural D-alpha tocopherol acetate, natural D-alpha tocopherol succinate from soybean sources; (c) vegetable oils derived phytosterols and phytosterol esters from soybean sources; (d) plant stanol ester produced from vegetable oil sterols from soybean sources.
7. Milk and products thereof (including lactose), except: (a) whey used for making distillates or ethyl alcohol of agricultural origin for spirit drinks and other alcoholic beverages; (b) 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: (a) 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 SO2.
13. Lupin and products thereof.
14. Molluscs and products thereof.'

An important question is whether hidden allergens can be identified in processed food. A number of methods for the detection of residues of allergens in foods have been developed by scientist but most of them are not commercially available yet. Such analytical methods will provide important tools for the food industry and for regulatory agencies. These methods will be important quality assurance tools for the food industry who is rather behind in this respect. Shared equipment is an economic necessity in the food industry and is common in a wide variety of food industry sectors (e.g. ice cream, confectionery, bakery, pasta). Proper sanitation of that shared equipment is critical to mitigate potential allergen hazards, and analytical tools will facilitate the validation of sanitation practices. Clearly, analytical methods must be specific, allowing the detection of an allergenic food in the presence of others, and must be appropriately sensitive, but also easy to use.

2.3.4 Major plant food allergen families

Plant proteins can be classified into families and superfamilies based on their sequence relationships. Comparison of these families with accepted lists of plant food-protein allergens demonstrates that only a small number of families contain characterized allergens, most falling into three large super-families (prolamin, cupin and the family 10 of pathogenesis-related

proteins).[28] Thus, about 129 food-allergen sequences are described in the literature, distributed over 20 protein families of the 3849 possible families.

2.3.5 Cross-Reactivity between allergens

Structural homology can explain cross-reactive IgE binding, but does not necessarily imply similar chemical stability or activity in the induction of allergic symptoms in allergic individuals. [29]

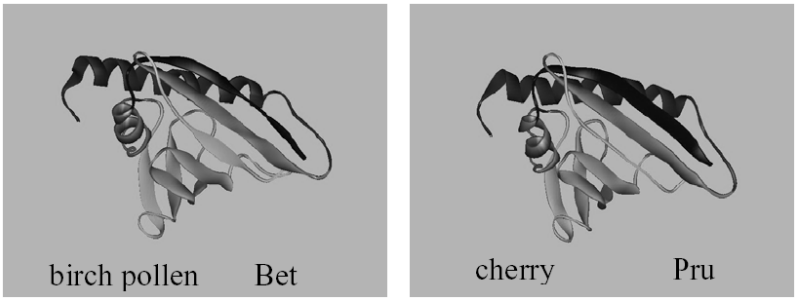


Figure 2.6. Structural homology between the major allergens in birch pollen and cherry: Bet v1 and Pru av1, respectively.

Structural homology and thus cross reactivity can explain elicitation/sensitization mechanisms: a pollen allergen can sensitize via the respiratory immune system and the food allergen, in a second phase, can elicit the allergic response. In fact, only a TRUE food allergen can sensitize via the gastro intestinal tract, but a great part of food allergens are protein cross-reacting with other sensitizers (Figure 2.7). [20]

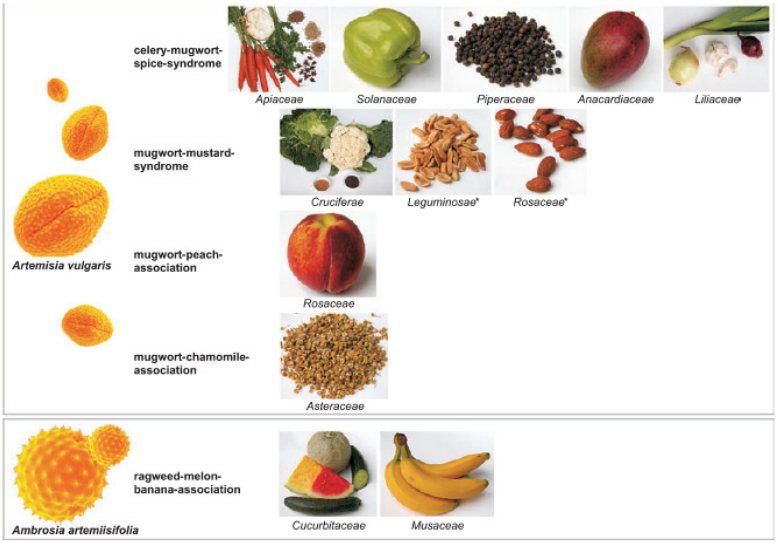


Figure 2.7. Examples of food allergens cross reacting with pollen allergens (from ref. 29).

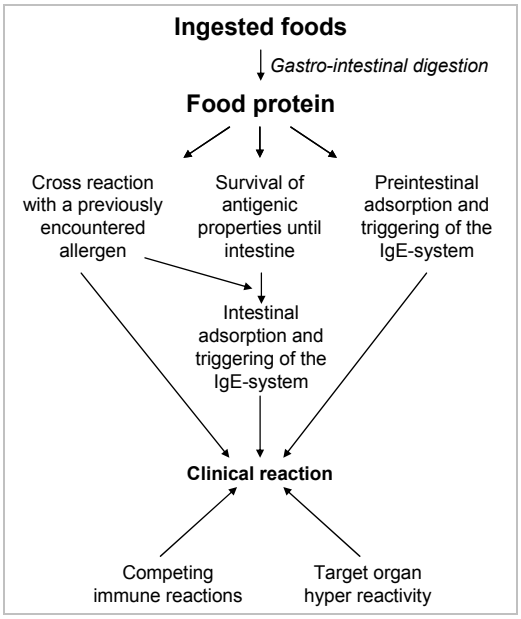


Figure 2.8. Mechanism proposed for the sensitization of true food allergens via the gastro intestinal barrier (adapted from ref.20).

Structural characteristics of food allergens in the sensitization and effector phase of food allergy are crucial in determining the capability of a food protein to arise an immunogenic response. Degradation of dietary proteins during the gastrointestinal transit results in a hydrolysis of the proteins to free AAs or oligopeptides, which either induce tolerance or are immunologically ignored. However, when proteins resist gastrointestinal proteases or persist during the transit due to impaired digestion, the epitope structure remains conserved. Thus, proteins reveal sensitization potential or trigger an allergic response by intact conformation .

2.3.6 Common properties and functions to food allergens

The characterization of food allergens at a molecular level has increased our understanding of the immunopathogenesis of many responses and may soon lead to novel immunotherapeutic approaches. The protein structure of many allergens has been resolved and has provided an explanation for the epitope-specific IgE cross-reactivity responsible for the pollen-food syndrome. Potentially important functional and physicochemical characteristics shared by many food allergens that might promote allergenicity: ligand binding, interaction with membranes and other lipid structures, disulfide bonds, glycosylation. All these characteristics enhance stability of the structure, for example a ligand integrated into the 3-dimensional structure of a protein can have the overall effect of reducing mobility of the polypeptide backbone, increasing both thermal stability and resistance to proteolysis, while many proteases require flexibility in substrate proteins. [30]

2.4 References

[1] Young VR, Protein and amino acids. *Present Knowledge in Nutrition*. 8th Edition. (eds). Bowman BA and Russel RM International Life Sciences Institute, Chapter 5, pp. 43-58. **2001**

[2] Korhonen, H. Pihlanto, A. Bioactive peptides from food proteins. *Handbook of Food Products Manufacturing*, **2007**, 2: 5-37.

[3] Gibbs B.F.; Zougman A.; Masse R.; Mulligan C., Production and characterization of bioactive peptides from soy hydrolysate and soy-fermented food. *Food Research International*, **2004**, 37, 2, 123-131

[4] Yoshikawa, M., Tsuruki, T., Bioactive peptides derived from soy protein. *Soy in health and disease prevention*, Editors: Sugano, M., **2006**

[5] Marta M.; Amaya A., Antihypertensive peptides derived from egg proteins. *Journal of Nutrition* **2006**, 136, 6, 1457-1460

[6] Quiros, A.; Chichon, R.; Recio, I.; Lopez-Fandino, R. The use of high hydrostatic pressure to promote the proteolysis and release of bioactive peptides from ovalbumin.. *Food Chemistry* **2007**, 104, 4, 1734-1739.

[7] Arihara K., Functional properties of bioactive peptides derived from meat proteins. *Food Science and Technology (Advanced Technologies for Meat Processing)* **2006**, 158: 245-273.

[8] Schanbacher F. L., Talhouk R. S. and Murray F. A., Biology and origin of bioactive peptides in milk. *Livestock Production Science*, 50, 2: 105-123

[9] Kitts, D.D.; Weiler, K. Bioactive Proteins and Peptides from Food Sources. Applications of Bioprocesses used in Isolation and Recovery *Current Pharmaceutical Design*, **2003,** 9, 16: 1309-1323

[10] Hoerr R. A., Bostwick E. F., Bioactive Proteins and Probiotic Bacteria: Modulators of Nutritional Health, *Nutrition*, **2000**, 16: 711-713

[11] Ortolani C., Pastorello E. A , Food allergies and food intolerances,.*Best Practice & Research Clinical Gastroenterology* **2006**, 20, 3: 467–483

[12] Bruijnzeel-Koomen C, Ortolani C, Aas K et al. Position paper of the European academy of allergology andclinical immunology on adverse reactions to food. *Allergy* **1995**; 12: 357–378.

[13] May CD. Immunologic versus toxic adverse reactions to foodstuffs. *Ann Allergy* **1983**; 51: 267–268.

[14] Sampson, H. A., Food allergy. Part 1: Immunopathogenesis and clinical disorders, *J. Allergy Clin. Immunol*. **1999**, 103: 717–728.

[15] Johansson, S. G. O., Bieber, T., Dahl, R., Friedmann, P. S. et al., Revised nomenclature for allergy for global use: Report of the Nomenclature Review Committee of the World Allergy Organization, *J. Allergy Clin. Immunol.* **2004**, 113: 832–836.

[16] Larché M., Akdis C.A. and Valenta R. Immunological mechanisms of allergen-specific immunotherapy, *Nature Reviews Immunology*, **2006**, 6: 761-771

[17] Taylor, S L and Hefle, S L 'Food allergies and other food sensitivities', *Food Technol*, **2001**, 55, 9, 68–83.

[18] Jeurink P.V. and Savelkoul H. F.J., Induction and regulation of allergen-specific IgE, Springer - Life Sciences Publisher ch. 2, **2005**

[19] Sampson H. A., Food allergy , *J Allergy Clin Immunol*, **2003**, 111, 2: 540-547

[20] Poulsen L. K., In search of a new paradigm: mechanisms of sensitization and elicitation of food allergy, *Allergy*, **2005**: 60: 549–558

[21] Sampson H.A., Update on food allergy, *Current Reviews of Allergy and Clinical Immunology,* **2003**, 113, 5: 805-819

[22] Mondoulet L., Paty E., Drumare M. F., Ah-Leung S., Scheinmann P., Willemot R. M., Wal J. M., and Bernard H. Influence of Thermal Processing on the Allergenicity of Peanut Proteins *J. Agric. Food Chem*. **2005**, 53: 4547-4553

[23] Brenna O, Pompei C, Ortolani C, Pravettoni V, Farioli L, Pastorello EA Technological processes to decrease the allergenicity of peach juice and nectar., *J. Agric. Food Chem*., **2000**, 48, 2: 493 -497,

[24] Luud J.W.J. Gilissen, Suzanne T.H.P. Bolhaar, André C. Knulst, Laurian Zuidmeer, Ronald van Ree, Z.S. Gao, and W. Eric van de Weg, Production of hypoallergenic plant foods by selection, breeding and genetic modification, *Allergy Matters: new approaches to allergy prevention and management* -The International Conference on Allergy Prevention, Wageningen, 2004 http://library.wur.nl/frontis/allergy_matters/toc.html

[25] Garcıa A., Wichers J. H., Wichers H. J., Decrease of the IgE-binding by Mal d 1, the major apple allergen, by means of polyphenol oxidase and peroxidase treatments, *Food Chemistry* **2007**, 103, 94–100

[26] Kootstra H., Vlieg-Boerstra B., van der Maas M., A. Dubois, Assessment of Hypoallergenic Properties of the Santana Apple *J. of Allergy and Clinical Immunology*, 119, 1, **2007**, S119

[27] Humires J. and Wal J.-M., EU Regulation: What's new in terms of labelling of food allergens? *Allergy* **2004**: 59: 1259–1261

[28] Breiteneder H., Clare Mills E.N. Plant food allergens—structural and functional aspects of allergenicity, *Biotechnology Advances*, **2005**, 23, 395–399

[29] Egger M., Mutschlechner S., Wopfner N., Gadermaier G., Briza P., Ferreira F., Pollen-food syndromes associated with weed pollinosis: an update from the molecular point of view, *Allergy* **2006**, 61: 461–476,.

[30] Breiteneder H., Clare Mills E.N., Molecular properties of food allergens *J Allergy Clin Immunol*, **2005**, 115, 1 , 14-23

3 TECHNIQUES OF ANALYSIS AND PURIFICATION OF PROTEINS AND PEPTIDES

Analysis and purification of peptides and proteins relies on a series of separation techniques. In general, separation procedures are either directed to the preparative level, in order to isolate one or more individual components from a mixture for further investigation, or to the analytical level, with the goal of identifying and determining the relative amounts of some or all of the mixture components.

Purification is essential is essential before protein properties and activities can be determined. The methods to be utilized depend on the matrix (tissue, microbial cells, foods). The first step in the protein purification procedures is the extraction, which can include homogenization, centrifugation, sonication with opportune solvents; desalting, concentration and defatting.[1,2] After the final purification, proofs of homogeneity, structural integrity and structural characterization are necessary to ascertain that the desired product, and not structural modifications of it, has been isolated. [3]

3.1 Liquid Chromatography

Once the extract is ready, various methods are available for separating and analyzing the proteins it contains. Only the techniques which have been utilized in the experiments of the present thesis are discussed.

Proteins can be purified according to solubility at different ionic strengths (salting in; salting out), charge (ion-exchange chromatography, Figure 3.1), size (dialysis; gel-filtration chromatography, Figure 3.2) and binding affinity (affinity chromatography, Figure 3.3; HPLC high-pressure liquid chromatography).

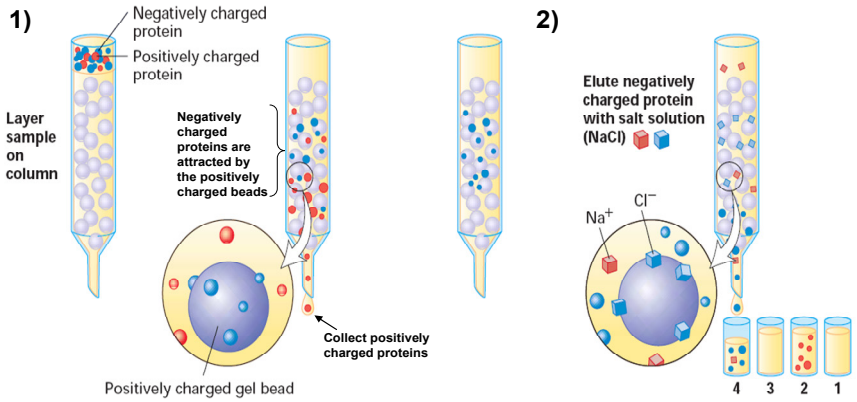


Figure 3.1. Ion exchange chromatography.

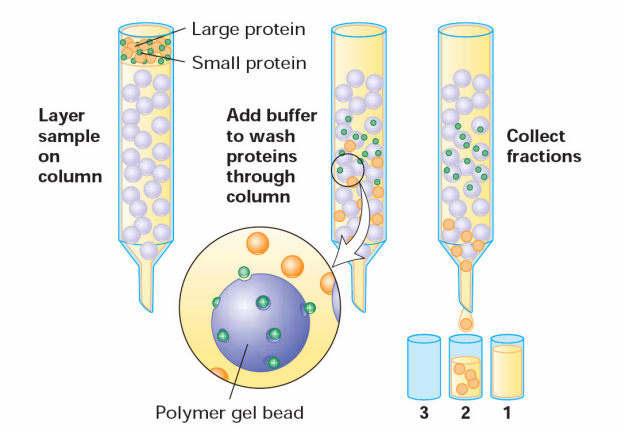


Figure 3.2. Gel filtration chromatography.

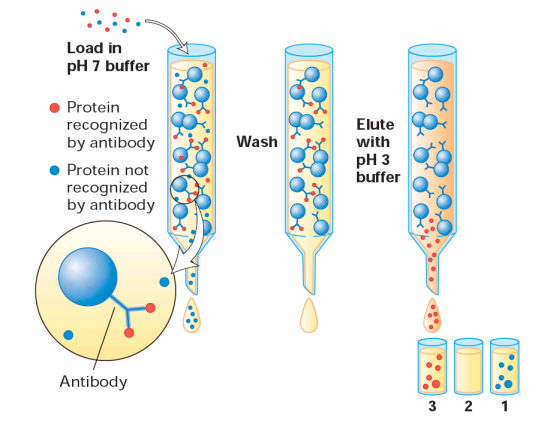


Figure 3.3. Affinity chromatography (by means of specific antibodies).

Methods for separating proteins take advantage of properties which characterize one protein such as size, charge, and binding properties. The partition of components between a solution and a solid surface is the most often exploited principle. The most common methods are reported in Table 3.1. In most cases, several different methods must be used to purify a protein completely.

Table 3.1. Selected methods for separation and purification of peptides and proteins[4].

Method	Remarks
Reversed-phase HPLC (RP-HPLC)	Suitable for the assessment of the level of hydrophilic and lipophilic activity.
Ion-exchange chromatography (IEC)	Most commonly used method for protein purification.
Size-exclusion chromatography (SEC)	Separation based on molecular size; larger molecules elute more rapidly than smaller ones.
Affinity chromatography (AC)	A bioselective ligand chemically bound to an inert matrix retains the target component with selective affinity to the ligand.
Capillary electrophoresis (CE)	Separation of peptides and proteins based on their differential migration in an electric field.
Multidimensional separations[5],[6],[7]	a) 2D gel electrophoresis; b) multidimensional chromatographic approach. They are capable of more accurate quantification of the analyte, and are more compatible with mass spectrometry.
Ultra-filtration (UF) [8],[9]	Method for rapid concentration of protein solutions; lack of selectivity has severely restricted the use of UF for protein fractionation.
Two-phase systems for protein separation and purification [10],[11]	Hydrophobic partitioning of proteins in aqueous two-phase systems containing poly(ethylene glycol) and hydrophobically modified dextrans.

Published purification protocols are available for many thousands of proteins.

3.2 Gel Electrophoresis

This commonly used technique for the separation of proteins is based on the migration of charged proteins in an electric field, a process called electrophoresis. These procedures are not generally used to purify proteins in large amounts, because simpler alternatives are usually available and electrophoretic methods often adversely affect the structure and thus the function of proteins. Electrophoresis is, however, especially useful as an analytical method, allowing for

the determination of crucial properties of a protein such as its approximate molecular weight (and sometimes, its isoelectric point).

Electrophoresis of proteins is generally carried out in gels of polyacrylamide, cross-linked by methylenebisacrylamide, chemically inert and readily formed. In electrophoresis all molecules are forced to move through the same matrix, driven by an electric field. Proteins can be separated on the basis of mass under denaturing conditions: the mixture of proteins is first dissolved in a solution of sodium dodecyl sulfate (SDS), an anionic detergent that disrupts nearly all noncovalent interactions in the native proteins. Mercaptoethanol (2-thioethanol) or dithiothreitol also is added to reduce disulfide bonds. Anions of SDS bind to the protein at a ratio of about one SDS anion for every two amino acid residues. This complex of SDS with a denatured protein has a large net negative charge that is roughly proportional to the mass of the protein. The negative charge acquired on binding SDS is usually much greater than the charge on the native protein, which is thus almost insignificant. The SDS-protein complexes are then subjected to electrophoresis.[12]

In addition, the native conformation of a protein is altered when SDS is bound, and most proteins assume a similar shape. Electrophoresis in the presence of SDS therefore separates proteins almost exclusively on the basis of mass (molecular weight), with smaller polypeptides migrating more rapidly. In electrophoresis, the velocity of migration (v) of a protein (or any molecule) in an electric field depends on the electric field strength (E), the net charge on the protein (z), and the frictional coefficient (f). The electrophoretic mobility reflects in part the protein shape, thus:

$$v = \frac{Ez}{f}$$

The electric force (Ez) driving the charged molecule toward the oppositely charged electrode is opposed by the viscous drag (fv) arising from the friction between the moving molecule and the medium. The frictional coefficient (f) depends on both the mass and shape of the migrating molecule and the viscosity (η) of the medium. For a sphere of radius (r),

$$f = 6\pi\eta r$$

The migration of a protein in a gel during electrophoresis is therefore a function of its size and its shape. After electrophoresis, the proteins are visualized by adding a dye, such as Coomassie blue, which reveals a series of bands. Radioactive labels can be detected by placing a sheet of x-ray film over the gel, a procedure called autoradiography.

By means of electrophoresis, it is possible to monitor the progress of a protein purification procedure as the number of protein bands visible on the gel decreases after each new fractionation step. When compared with the positions to which proteins of known molecular weight migrate in the gel, the position of an unidentified protein can provide a good measure of its molecular weight. If the protein has two or more different subunits, the subunits will be

separated by the SDS treatment with eventual addiction of urea, and two separate bands will appear.

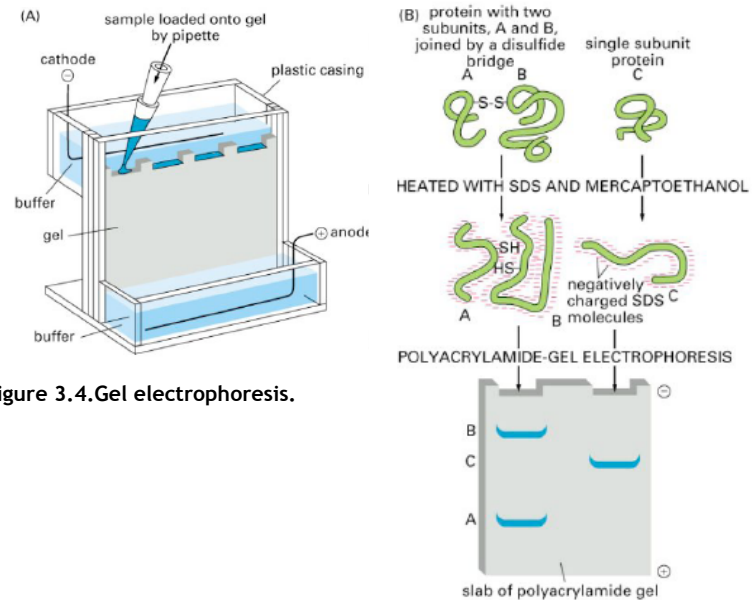


Figure 3.4.Gel electrophoresis.

3.2.1 Isoelectric Focusing.

Isoelectric focusing is a procedure used to separate proteins on the base of their isoelectric point (pI). A pH gradient is established by allowing a mixture of low molecular weight organic acids and bases (ampholytes[13]) to distribute themselves in an electric field generated across the gel. When a protein mixture is applied, each protein migrates until it reaches the pH that matches its pI. Proteins with different isoelectric points are thus distributed differently throughout the gel.

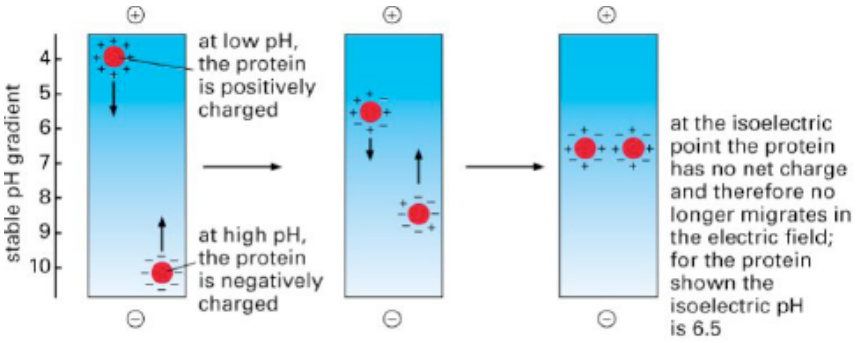


Figure 3.5. Isoelectrofocusing IEF.

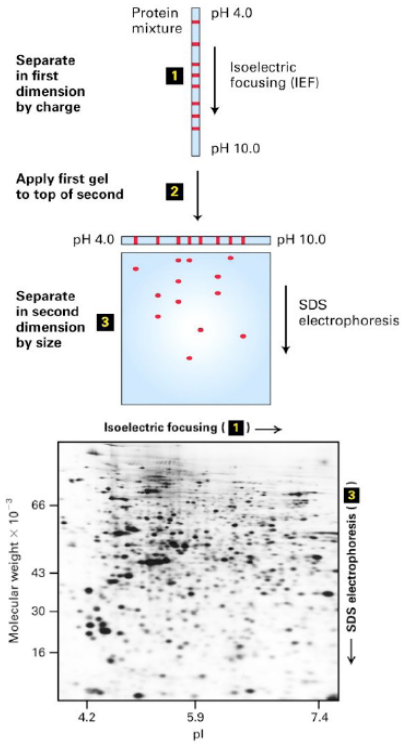


Figure 3.6. 2D electrophoresis.

3.2.2 Two-Dimensional Electrophoresis.

By combining isoelectric focusing and SDS electrophoresis sequentially in a process called two-dimensional electrophoresis it is possible to obtain the resolution of complex mixtures of proteins (Figure 3.6). This is a more sensitive analytical method than the electrophoretic method alone. Two-dimensional electrophoresis separates proteins of identical molecular weight that differ in pI, or proteins with similar pI values but different molecular weights.

3.3 References

[1] Scopes, R K, *Protein Purification Principles and Practice*, Springer Advanced Texts in Chemistry, 3rd ed., **1994**,

[2] *Protein Purification Handbook*, Amersham Pharmacia Biotech, www.genomics.bham.ac.uk/Documents/ProteinPurification.pdf

[3] Sewald N., H.D. Jakubke *Peptides: Chemistry and Biology*. **2002** Wiley-VCH Verlag GmbH & Co.

[4] Larive, C. K.; Lunte, S. M.; Zhong, M.; Perkins, M. D.; Wilson, G. S.; Gokulrangan, G.; Williams, T.; Afroz, F.; Schöneich, C.; Derrick, T. S.; Middaugh, R.; Bogdanowich-Knipp, S. Separation and analysis of peptides and proteins, *Anal. Chem.* **1999**, 71, 389-423.

[5] N.G. Anderson, N. L. Anderson, *Anal. Biochem*. **1978**, 85, 331

[6] T. Okazaki, Y. Yanagisawa, T. Nagai, Clin. Chim. Acta **1997**, 258, 137

[7] G. J. Opiteck, J.W. Jorgenson, M.A. Mosely, R. J. Anderegg, J. Microcolumn Sep. **1998**, 10, 365

[8] M.R. Torres, A. J. Ramos, E. Soriano,Bioprocess Eng**. 1998**, 19, 213

[9] U. Mathias, R. Marco, M. Uwe, J.Membr. Sci. **1996**, 120, 239

[10] F. Hachem, B. A. Andrews, J. A. Asenjo,Enzyme Microb. Technol. **1996**, 19, 507

[11] M. Lu, F. Tjerneld, J. Chromatogr. A **1997**, 766, 99

[12] Berg J. M., Tymoczko J. L., Stryer L., *Biochemistry*, ed. W. H. Freeman, 6th Edition, **2006**

[13] Bjellqvist B, Ek K, Righetti PG, Gianazza E, Görg A, Westermeier R, Postel W. Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods* **1982,** 6, 4:317-39.

4 MASS SPECTROMETRY FOR BIOLOGICAL COMPOUNDS

Some of the most difficult chemical, biological and environmental problems require a combination of instrumental attributes such as speed, sensitivity, wide linear dynamic range, small sample quantity, low detection limit, and capability of on-line analysis in conjunction with separation techniques. Among the analytical techniques, MS holds a special place because it measures an intrinsic property of a molecule, its mass, with a very high sensitivity and therefore it is used in an amazingly wide range of applications. MS has the potential to provide accurate mass measurements from low molecular weight compounds to macromolecules.

Nowadays, mass spectrometry is used in different fields, for both routine and research purposes, in particular in the analysis of proteins, peptides and oligonucleotides.

MS makes possible:

- accurate molecular weight measurements, used for sample confirmation (i.e. quality control of recombinant proteins), for determining the purity of a sample, for verifying amino acid substitutions, for detecting post-translational modifications and any covalent modification that alters the mass of a protein, for calculating the number of disulphide bridges;
- reaction monitoring, for monitoring enzyme reactions, chemical modifications, protein digestion;
- amino acid sequencing, for sequence confirmation, de novo characterization of peptides, identification of proteins by database searching with a sequence "tag" from a proteolytic fragment;
- protein structure studies, by protein folding monitoring through H/D exchange, for the determination of protein-ligand complex formation under physiological conditions, for macromolecular structure determination;
- oligonucleotide sequencing, for the characterization or quality control of natural and synthetic oligonucleotides.

All mass spectrometric techniques have in common the transformation of a molecule, either ionic or neutral, into an ion transferred to the gas phase. These ions are then introduced into an analyzer and their paths depend on their mass-to-charge ratio, m/z, allowing for their physical separation and detection. This measured property of the ionized species can be used to deduce the mass (M) of the analyte with very high precision. Mass spectrometers can be divided into three fundamental parts, namely the ionization source, the analyzer, and the detector (Figure 4.1) preceded by a sample inlet system and followed by a systemwhich elaborates ionic signals into electrical signals.

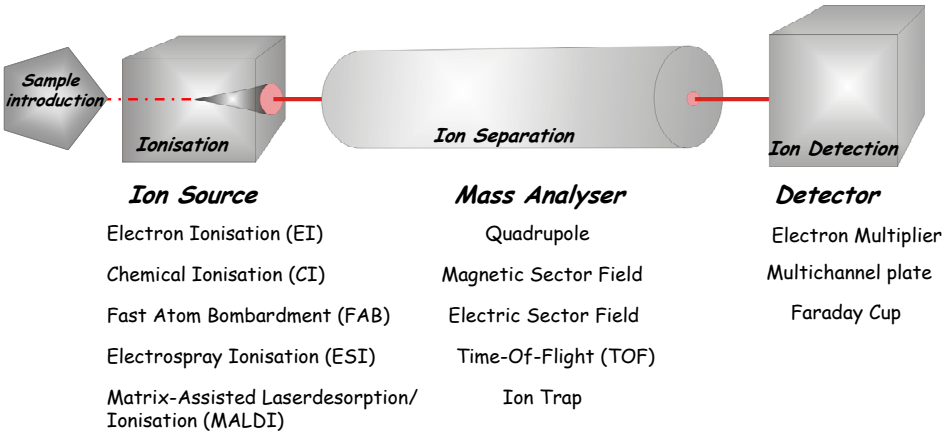


Figure 4.1.Diagram of mass spectrometric analyses.

4.1 Sample introduction

The sample has to be introduced into the ionization source of the instrument, where the sample molecules are ionized. The method of sample introduction to the ionization source often depends on the ionization method being used, as well as the type and complexity of the sample. The sample can be inserted directly into the ionization source, or can undergo some type of chromatography en route to the ionization source. This latter method of sample introduction usually implies that the mass spectrometer is coupled directly to a high pressure liquid chromatography (HPLC), gas chromatography (GC) or capillary electrophoresis (CE) separation column, and hence the sample is separated into a series of components which then enter the mass spectrometer sequentially for individual analysis.

Chromatography has long been the choice of separation in analytical chemistry, and HPLC[1] leads the way for the majority of mass spectrometry applications.

High Pressure Liquid Chromatography (HPLC)

HPLC is a type of chromatography in which the sample dissolved in a solvent flows through the separating column in a liquid mobile phase at high pressure (typically 50-200 bar). Components within the sample come in contact with the stationary phase of the column and are retained with different affinities, they undergo a series of partition equilibra between the stationary and the mobile phase and thus separation occurs. Compounds modulate their affinity for the solid or the mobile phase according to their polarity. The vast majority of HPLC applications carried out uses a non-polar stationary phase and a polar mobile phase (reversed-phase HPLC). When a mobile phase of constant composition is used (e.g. 50:50 acetonitrile:water), it is said to be run under isocratic conditions, whereas when the mobile phase composition changes throughout a run, a gradient elution is performed.

4.2 Ionization methods.

Many ionization methods are available and each has its own advantages and disadvantages.[2] Mechanism of sample ionization are here listed:

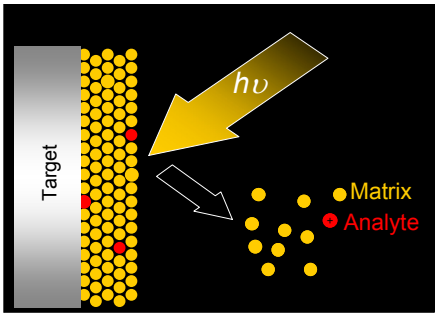
- Protonation: $M + H^+ \rightarrow MH^+$
- Cationization: $M + Cat^+ \rightarrow MCat^+$
- Deprotonation: $MH \rightarrow M^- + H^+$
- Electron Ejection: $M \rightarrow M^{+\cdot} + e^-$
- Electron Capture: $M + e^- \rightarrow M^{-\cdot}$

Due to their "soft" character, able to induce ionization without extensive molecule fragmentation, Matrix Assisted Laser Desorption Ionization (MALDI) and Electrospray Ionization (ESI) are used for many applications involving bioorganic compounds.

4.2.1 Matrix Assisted Laser Desorption Ionization MALDI

Matrix Assisted Laser Desorption Ionization (MALDI)[3] is mostly utilized with thermolabile, non-volatile organic compounds, especially those of high molecular mass, and has been used successfully for the analysis of proteins, peptides, glycoproteins, oligosaccharides, and oligonucleotides. The technique is relatively straightforward to use and reasonably tolerant to buffers and other additives. The mass accuracy depends on the type and performance of the analyzer of the mass spectrometer, but most modern instruments are capable of measuring masses to within 0.01% of the molecular mass of the sample, at least up to ca. 40,000 Da.

MALDI is based on the bombardment of sample molecules with a laser light in order to obtain sample Ionization. The sample is pre-mixed with a matrix at a low concentration of sample into the matrix . The matrix transforms the laser energy into excitation energy for the sample, which leads to the sputtering the analyte and matrix ions from the surface of the mixture. In this way energy transfer is efficient and also the analyte molecules are spared excessive direct energy that may otherwise cause decomposition.



Most commercially available MALDI mass spectrometers use a pulsed nitrogen laser at a wavelength of 337 nm.

The sample to be analyzed is dissolved in an appropriate volatile solvent, usually with a trace of trifluoroacetic acid if positive Ionization is being used, at a concentration of ca. 10 pmol/μL. An aliquot of this solution is mixed with an equal volume of a solution containing a vast excess of the matrix. A wide range of compounds are suitable to be used as matrices: sinapinic acid is common for protein analysis while alpha-cyano-4-hydroxycinnamic acid is often used for peptide analysis. An aliquot of the final solution is applied to the sample target which is allowed to dry

prior to insertion into the high vacuum of the mass spectrometer. After the ionization has been obtained by laser firing, usually a time-of-flight analyzer separates the ions according to their mass(m)-to-charge(z) (m/z) ratios by measuring the time it takes for ions to travel through a field free region known as the flight, or drift, tube. The heavier ions are slower than the lighter ones. The m/z scale of the mass spectrometer is calibrated with a known sample that can either be analyzed independently (external calibration) or pre-mixed with the sample and matrix (internal calibration). MALDI ionization results predominantly in the generation of singly charged ions regardless of the molecular mass, hence the spectra are relatively easy to interpret. Fragmentation of the sample ions does not usually occur.

In positive ionization mode the protonated molecular ions ($M+H^+$) are usually the dominant species, although they can be accompanied by salt adducts, a trace of the doubly charged molecular ion at approximately half the m/z value, and/or a trace of a dimeric species at approximately twice the m/z value. Positive ionization is used in general for protein and peptide analyses. In negative ionization mode the deprotonated molecular ions $(M-H)^-$ are usually the most abundant species, accompanied by salt adducts and possibly traces of dimeric or doubly charged materials. Negative ionization can be used for the analysis of oligonucleotides and oligosaccharides.

4.2.2 Electrospray (ESI)

Electrospray Ionization (ESI) is one of the Atmospheric Pressure Ionization (API) techniques and is well-suited for the analysis of polar molecules ranging from less than 100 Da to more than 100000 Da in molecular mass. During standard electrospray Ionization[4] the sample is dissolved in a polar, volatile solvent and pumped through a narrow, stainless steel capillary (75 - 150 µm internal diameter) at a flow rate between 1 µL/min and 1 ml/min.

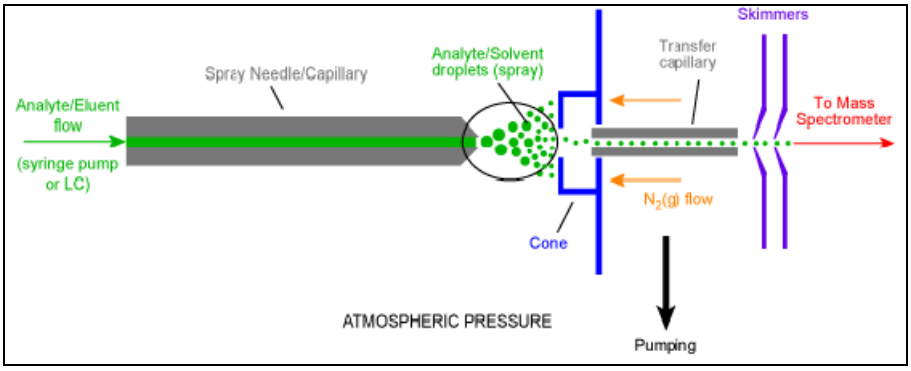


Figure 4.2. A schematic view of ESI interface.

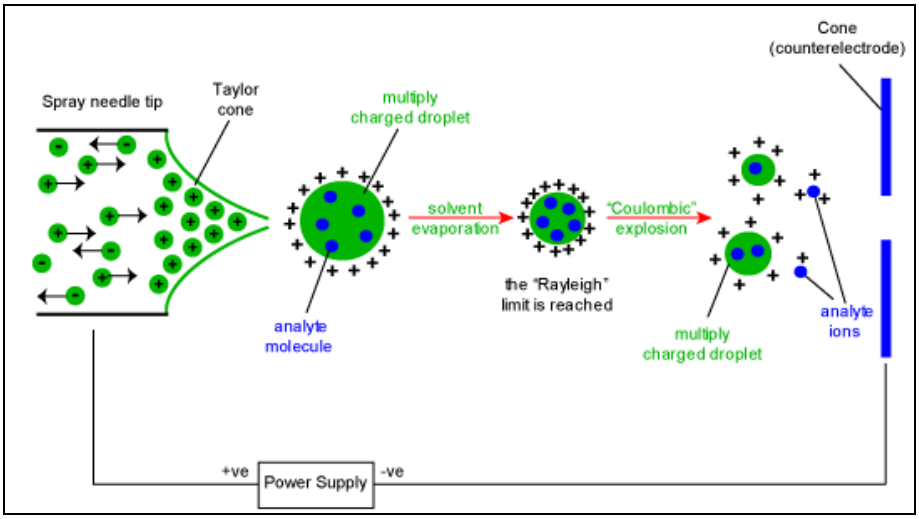


Figure 4.3. A schematic view of the analyte/solvent spray in ES Ionisation.

A high voltage of 3 or 4 kV is applied to the tip of the capillary, which is situated within the Ionization source of the mass spectrometer, and as a consequence of this strong electric field, the sample emerging from the tip is dispersed into an aerosol of highly charged droplets, a process that is aided by a co-axially introduced nebulising gas flowing around the outside of the capillary. This gas, usually nitrogen, helps to direct the spray emerging from the capillary tip towards the mass spectrometer. The charged droplets diminish in size by solvent evaporation, assisted by a warm flow of nitrogen known as the drying gas which passes across the front of the ionization source. As the solvent evaporation occurs, the droplet shrinks until it reaches the point that the surface tension can no longer sustain the charge (the Rayleigh limit) at which point a "Coulombic explosion" occurs and the droplet is ripped apart. This produces smaller droplets, that can repeat the process, as well as naked charged ions. These charged molecules can be singly or multiply charged. This is a very soft method of Ionization as very little residual energy is retained by the analyte upon Ionization. This is why ESI-MS is such an important technique in biological studies where the analyst often requires that non-covalent molecule-protein or protein-protein interactions are representatively transferred into the gas-phase. In this technique nearly no fragmentation is produced but this is useful for structural studies where the analyte molecules are chosen before being fragmented. Eventually charged ions, free from solvent, released from the droplets, pass through a sampling cone or orifice into an intermediate vacuum region, and from there into the analyzer of the mass spectrometer, which is held under high vacuum. The lens voltages are optimized individually for each sample.

ESI Ionization does not tolerate non-volatile buffers and other additives, which should be avoided as far as possible. In positive Ionization mode, a trace of formic acid is often added to aid protonation of the sample molecules; in negative Ionization mode a trace of ammonia solution or a volatile amine is added to aid deprotonation of the sample molecules. Proteins and

peptides are usually analyzed under positive Ionization conditions and saccharides and oligonucleotides under negative Ionization conditions.

Several variations on the electrospray process have been developed such as nanospray, microspray and ionspray (also referred to as pneumatically assisted electrospray)[5].

Nanospray Ionization is a low flow rate version of electrospray Ionization. A common application of this technique is for a protein digest mixture to be analyzed to generate a list of molecular masses for the components present, and then each component to be analyzed further by tandem mass spectrometric (MS-MS) amino acid sequencing techniques. [6]

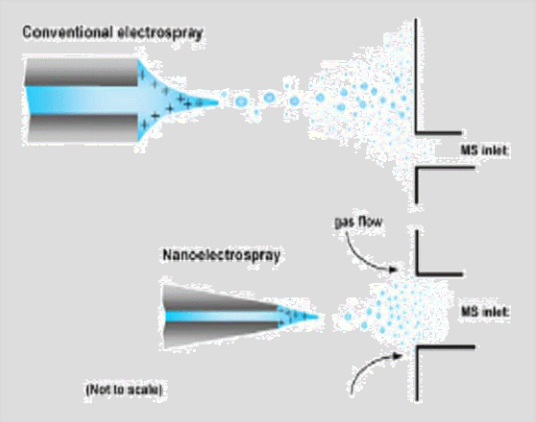


Figure 4.4. Normal flowrate electrospray (top) and a nano flowrate electrospray that produces smaller droplets. By allowing a closer proximity to the MS inlet, the lower flow rate electrospray affords more efficient ion introduction

4.3 Mass analyzers

The main function of the mass analyzer is to separate, or resolve, the ions formed in the Ionization source of the mass spectrometer according to their mass-to-charge (m/z) ratios. There are a number of mass analyzers currently available, the better known of which include quadrupoles , time-of-flight (TOF) analyzers, magnetic sectors, and both Fourier transform and quadrupole ion traps.[7] These mass analyzers have different features, including the m/z range that can be covered, the mass accuracy, and the achievable resolution. The compatibility of different analyzers with different Ionization methods varies. For example, all the analyzers listed above can be used in conjunction with electrospray Ionization, whereas MALDI is not usually coupled to a quadrupole analyzer.

4.3.1 Time-of-Flight Mass Spectrometer

Time-of-flight mass spectrometry (TOF-MS) is probably the simplest method of mass measurement to conceptualize, although there are hidden complexities when it comes to higher

resolution applications. The first commercial TOF instrument was marketed by the Bendix corporation in the late 1950's. Their design was based on the Wiley & MacLaren instrument that was published in 1955[8]. TOF-MS has become widely used in recent years as being an essential instrument for biological analysis applications - this is especially the case with the coupling of TOF-MS to MALDI and ESI Ionization methods and the development of high-resolution and hybrid instruments (for example Q-TOF and TOF-TOF configurations). The inherent characteristics of TOF-MS are extreme sensitivity (all ions are detected), almost unlimited mass range and speed of analysis (modern instruments can obtain full spectra in seconds). This makes TOF-MS one of the most desirable methods of mass analysis.

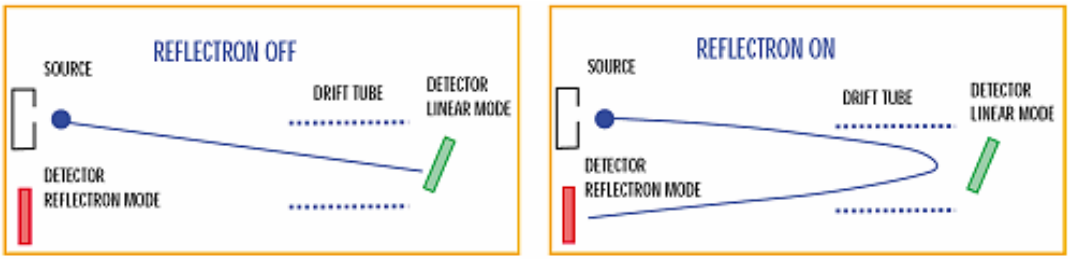


Figure 4.5.A Schematic of a Time-of-Flight mass spectrometer operating in Linear and Reflectron Mode.

The general set up of TOF is shown in Figure 4.5. The ions are introduced either directly from the source of the instrument or from a previous analyzer (in the case of Q-TOF) as a pulse. This results in all the ions receiving the same initial kinetic energy. As they then pass along the field free drift zone, they are separated by their masses, lighter ions travelling faster. This enables the instrument to record all ions as they arrive at the detector and so accounts for the techniques high sensitivity. The equation governing TOF separation is:

$$\frac{m}{z} = 2e\mathbf{E}s\left(\frac{t}{d}\right)^2$$

m/z is the mass-to-charge ratio of the ion
E is the extraction pulse potential
s is the length of the flight tube over which E is applied
d is the length of the field free drift zone
t is the measured time-of-flight of the ion

Theoretically, all ions are given the same initial kinetic energy by the extraction pulse and then drift along the field free drift zone where they will be separated so that all ions at the same *m/z* arrive at the detector at the same time. In practice, the pulse is not felt by all ions to the same intensity and so a kinetic energy distribution for each discrete *m/z* exists. This lowers the resolution by creating a time-of-flight distribution for each *m/z*. This can be corrected quite easily by the application of a reflectron at the end of the drift zone. This consists of a series of electric fields which repulse the ions back along the flight tube - usually at a slightly displaced angle (see figure) - resulting in a refocusing of ions with the same *m/z* value on the reflectron detector.

4.3.2 Quadrupole Mass Spectrometers

A quadrupole mass analyzer consists of four parallel rods (see Figure 4.6) that have fixed (DC) and alternating (RF) potentials applied to them. Ions produced in the source of the instrument are then focused and passed along the middle of the quadrupoles. Their motion will depend on the electric fields so that only ions of a particular m/z will be in resonance and thus pass through to the detector. The RF is varied to bring ions of different m/z into focus on the detector and thus build up a mass spectrum. The trajectory of the ions through the quadrupole is actually very complex.

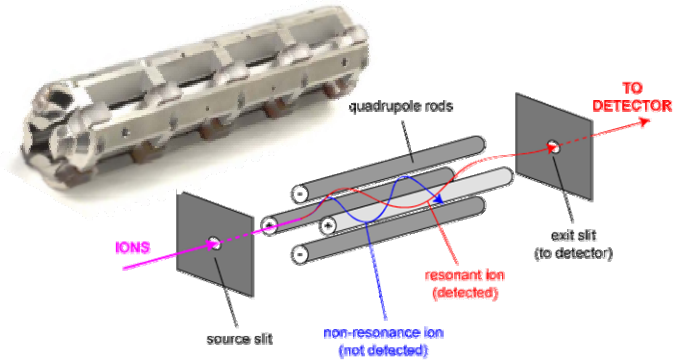


Figure 4.6. A quadrupole mass analyzer and evolution of resonant and NON-resonant ions in the quadrupole.

The two opposite rods in the quadrupole have a potential of +(U+Vcos(ωt)) (labelled '+' on the Figure) and the other two -(U+Vcos(ωt)) where 'U' is the fixed potential and Vcos(ωt) is the applied RF of amplitude 'V' and frequency 'ω'. The applied potentials on the opposed pairs of rods varies sinusoidally as cos(ωt) cycles with time 't'. This results in ions being able to traverse the field free region along the central axis of the rods but with oscillations amongst the poles themselves. These oscillations result in complex ion trajectories dependent on the *m/z* of the ions. Specific combinations of the potentials 'U' and 'V' and frequency 'ω' will result in specific ions being in resonance creating a stable trajectory through the quadrupole to the detector. All other *m/z* values will be non-resonant and will hit the quadrupoles and not be detected (see Figure 4.6). The mass range and resolution of the instrument is determined by the length and diameter of the rods.[9]

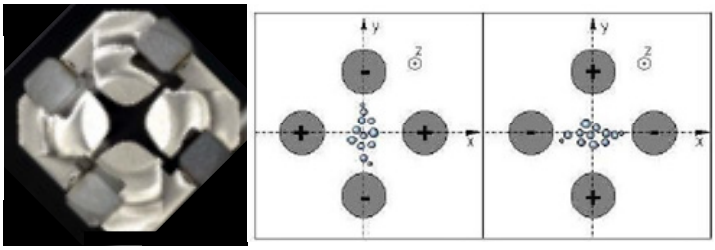


Figure 4.7. Evolution of ions in a Q section during one cycle of alternative current.

4.3.3 Ion-Trap Mass Spectrometers

The quadrupole ion trap (QIT) mass analyzer was developed in parallel with the quadrupole mass analyzer by the mass spectrometry pioneer, Wolfgang Paul in 1950, but only in the 1980's it was QIT-MS started to be commonly used as it is today.[9] A scheme of the basic set up of a QIT mass analyzer is shown in Figure 4.8.

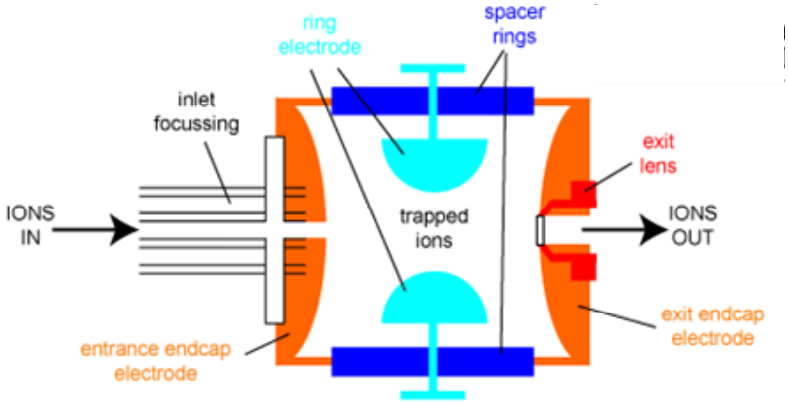


Figure 4.8.A Scheme (cutaway view) of a Quadrupole Ion Trap Mass Analyzer.

The ions, produced in the source of the instrument, enter into the trap through the inlet and are trapped through the action of the three hyperbolic electrodes: the ring electrode and the entrance and exit endcap electrodes. Various voltages are applied to these electrodes which result in the formation of a cavity in which ions are trapped. The ring electrode (RF) potential, an alternatve current potential of constant frequency but variable amplitude, produces a 3D quadrupolar potential field within the trap. This traps the ions in a stable oscillating trajectory. The exact motion of the ions is dependent on the voltages applied and their individual mass-to-charge (m/z) ratios. For detecting the ions, the potentials are altered to destabilize the ion motions resulting in ejection of the ions through the exit endcap. The ions are usually ejected in order of increasing m/z by a gradual change in the potentials. This 'stream' of ions is focused onto the detector of the instrument to produce the mass spectrum. This method for trapping and ejecting ions makes a quadrupolar ion trap especially suited to perform MS^n experiments in structural elucidation studies. In fact, it is possible to selectively isolate a particular m/z in the

trap by ejecting all the other ions from the trap. Fragmentation of this isolated precursor ion can then be induced by collisionally induced dissociation (CID) experiments. The isolation and fragmentation steps can be repeated several times and are only limited by the trapping efficiency of the instrument. Experiments up to MS^5 are fairly routine with this set-up as is the coupling of liquid chromatography to perform LC-MS^n studies.

4.3.4 Magnetic Sectors

The sector mass spectrometer is one of the most common types of mass analyzer and probably the most familiar to the scientist. In the 1950's, the first commercial mass spectrometers were sector instruments. They consist of some combination of a large electromagnetic ('B' sector), and some kind of electrostatic focusing device ('E' sector) - different manufacturers use differing geometries. Figure 4.9 shows a scheme of a standard 'BE' geometry double focusing instrument - that is, a dual sector instrument consisting of a magnetic sector followed by an electrostatic sector.

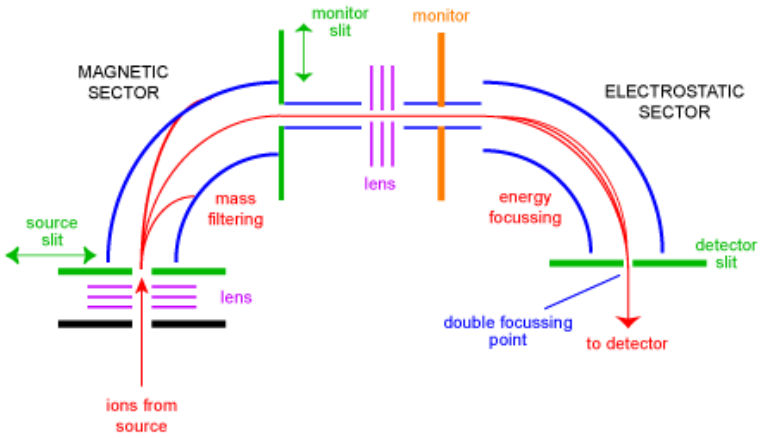


Figure 4.9. A Scheme of a sector mass spectrometer ('BE' geometry).

Ions enter the instrument from the source where they are initially focused. They enter the magnetic sector through the source slit where they are deflected according to the left-hand rule. Higher-mass ions are deflected less than lower-mass ions. Scanning the magnet enables ions of different masses to be focused on the monitor slit. At this stage, the ions have been separated only by their masses. To obtain a spectrum of good resolution - i.e. where all ions with the same m/z appear coincident as one peak in the spectrum, ions have to be filtered by their kinetic energies. After another stage of focusing, the ions enter the electrostatic sector where ions of the same m/z have their energy distributions corrected and are focused at the double focusing point on the detector slit.

Sector instruments are now suited to EI and CI Ionization in GC-MS. Single sector instruments are also used in the specialistic area of isotope ratio mass spectrometry (IRMS).

4.3.5 Fourier-transform Ion Cyclotron Resonance (FT-ICR)

Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS or just FT-MS) was first published in the 1950's where it was demonstrated to be useful for measuring very small mass differences at a very high precision[10]. The technique remained a largely academic tool until the application of FT methods by Alan Marshall and Melvin Comisarow in the early 1970's. It is now one of the most sensitive methods of ion detection in use and with a very high resolution since usually most experiments works in the 10^5 to 10^6 range.[11]

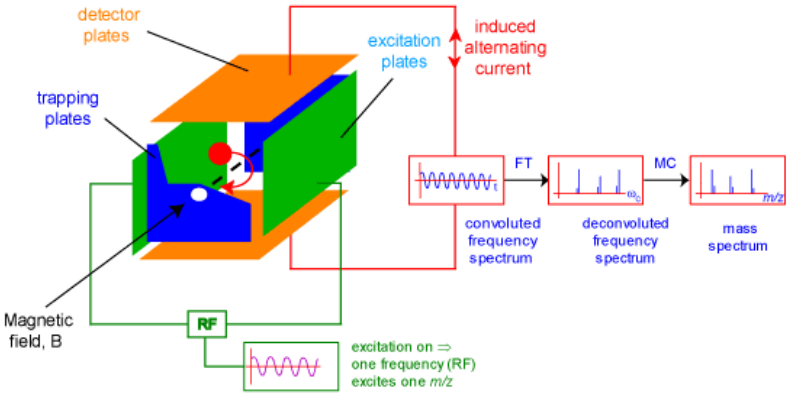


Figure 4.10. Scheme of FT-ICR-MS showing the ion trapping, detection and signal generation.

In the basic FT-MS instrument, the ions are generated in the source (as usual) and then pass through a series of pumping stages under increasingly high vacuum. When the ions enter the cell (ion trap) pressures are in the range of 10^{-10} to 10^{-11} mBar with temperatures close to absolute zero. The cell is located inside a spatial uniform static superconducting high field magnet (typically 4.7 to 13 Tesla) cooled by liquid helium and liquid nitrogen. When the ions pass into the magnetic field they are bent into a circular motion in a plane perpendicular to the field by the Lorentz Force (see equation *1*). They are prevented from precessing out of the cell by the trapping plates at each end.

$$\mathbf{F} = z\mathbf{v} \times \mathbf{B} \quad \text{.....(1)}$$

$$\omega_c = \frac{zB}{2\pi m} \quad \text{.....(2)}$$

$$m/z = \frac{B}{2\pi\omega_c} \quad \text{.....(3)}$$

F is the Lorentz Force observed by the ion when entering the magnetic field
B is the magnetic field strength (constant)
v is the incident velocity of the ion
$\boldsymbol{\omega_c}$ is the induced cyclotron frequency
m is the mass of the ion
z is the charge on the ion

The frequency of rotation of the ions is dependent on their m/z ratio (equation 2). At this stage, no signal is observed because the radius of the motion is very small. Excitation of each individual m/z is achieved by a swept RF pulse across the excitation plates of the cell. Each individual excitation frequency will couple with the ions natural motion and excite them to a higher orbit

where they induce an alternating current between the detector plates. The frequency of this current is the same as the cyclotron frequency of the ions and the intensity is proportional to the number of ions. When the RF goes off resonance for that particular m/z value, the ions drop back down to their natural orbit (relax) and the next m/z packet is excited. Although the RF sweep is made up of a series of stepped frequencies, it can be considered as all frequencies simultaneously. This results in the measurement of all the ions in one go producing a complex frequency vs. time spectrum containing all the signals, the FID. Deconvolution of this signal by FT methods results in the deconvoluted frequency vs. intensity spectrum which is then converted to the mass vs. intensity spectrum (the mass spectrum) by equation *3*. It is also usual to correct for mass errors at this stage by applying a calibration.

4.4 Detectors

The detector monitors the ion current, amplifies it and the signal is then transmitted to the data system where it is recorded in the form of mass spectra . The m/z values of the ions are plotted against their intensities to show the number of components , the molecular mass of each component, and the relative abundance of the various components in the sample. The type of detector is supplied to suit the type of analyzer; the most common are the photomultiplier, the electron multiplier and the micro-channel plate detectors.

4.4.1 The Photomultiplier or Scintillation Counter

In the photomultiplier (or scintillation counter) the ions initially strike a dynode which results in electron emission. These electrons then strike a phosphorous screen which in turn releases a burst of photons. The photons then pass into the multiplier where amplification occurs in a cascade fashion - much like with the electron multiplier. The main advantage of using photons is that the multiplier can be kept sealed in a vacuum preventing contamination and greatly extending the lifetime of the detector. Photomultipliers are now probably the most common detectors in modern mass spectrometers.

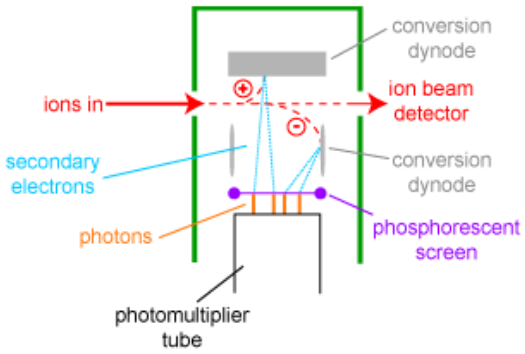


Figure 4.11. Scheme of a photomultiplier, showing the conversion of the ion/electron signal into photon(s) which are then amplified and detected by the photomultiplier.

4.4.2 The Electron Multiplier

Electron multipliers are probably the most common means of detecting ions, especially when positive and negative ions need to be detected on the same instrument. There are two types of electron multiplier, both working essentially by extending the principle of the Faraday cup. A Faraday cup uses one dynode and as a result produces one level of signal amplification. One type of electron multiplier (on the left) has series of dynodes maintained at increasing potentials resulting in a series of amplifications. The other type (the channel multiplier, on the right) has a curved ('horn' shaped) continuous dynode where amplifications occur through repeated collisions with the dynode surface. In both cases, ions pass the conversion dynode (depending on their charge) and strike the initial amplification dynode surface producing an emission of secondary electrons which are then attracted either to the second dynode, or into the continuous dynode where more secondary electrons are generated in a repeating process ultimately resulting in a cascade of electrons. Typical amplification is of the order of one million to one.

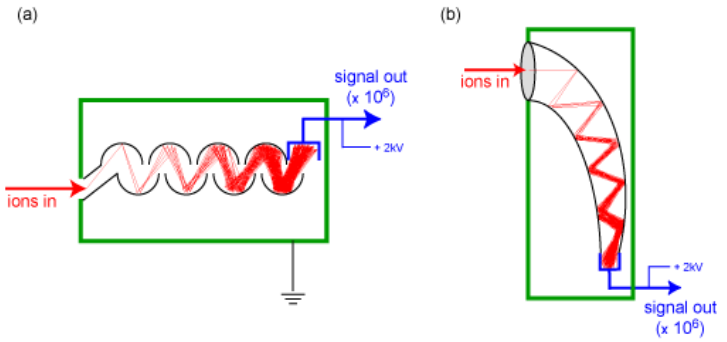


Figure 4.12. Schemes of the two types of electron multiplier, showing the cascade of electrons that results in amplification.

4.4.3 Micro Channel Plate (MCP)

A micro-channel plate (MCP) is closely related to an electron multiplier, as both intensify single particles or photons by the multiplication of electrons via secondary emission. A microchannel plate (MCP) is an array of 104-107 miniature electron multipliers oriented parallel to one another.[12]

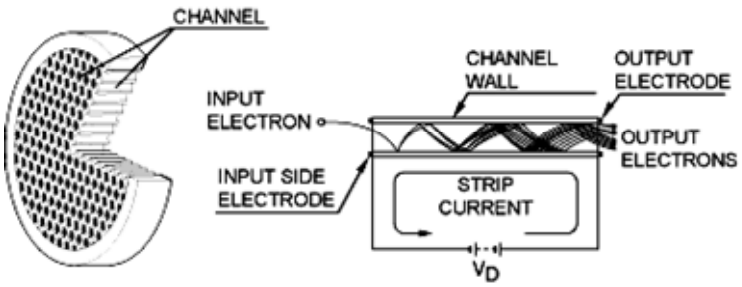


Figure 4.13. A micro channel plate, cutaway view and electron multiplier system incorporated.

4.4.4 The Faraday Cup or Cylinder

The Faraday cup or cylinder electrode detector is very simple. The basic principle is that the incident ion strikes the dynode surface which emits electrons and induces a current which is amplified and recorded. The dynode electrode is made of a secondary emitting material like CsSb, GaP or BeO. The Faraday cup is a relatively insensitive detector but is very robust. It is ideally suited for isotope analysis and IRMS.

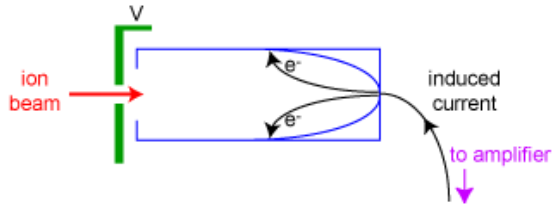


Figure 4.14.A schematic of a Faraday Cup or Cylinder Electrode.

4.5 Tandem mass spectrometry (MS-MS)

Tandem (MS-MS) mass spectrometry involves the use of more than one analyzer and it is suitable for structural and sequencing studies. The analyzers, usually two, are not necessarily of the same type, in which case the instrument is a hybrid one. The two analyzers are separated by a collision cell into which an inert gas (e.g. argon, xenon) is allowed to collide with the selected sample ions and bring about their fragmentation (CID, Collision Induced Dissociation).

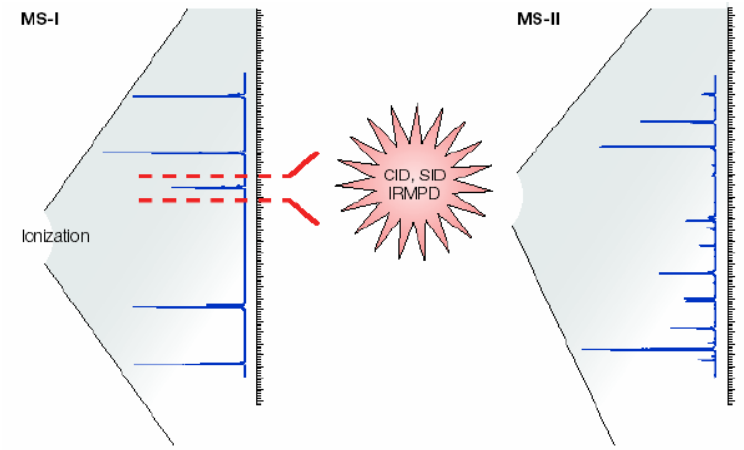


Figure 4.15. Scheme of tandem mass spectrometry. In MS/MS ions with the mass-to-change m/z ratio of interest (parent or precursor ion) are selectively reacted to generate a mass spectrum of product ions. Possible mechanisms of collision are: collision induced dissociation (CID), infrared multi-photon dissociation(IRMPD), surface induced dissociation (SID).

In MS/MS, the precursor is fragmented in a collision cell or chamber before the mass spectrum is acquired. This produces a product ion scan for the particular precursor ion. This type of data is

used for sequencing (peptides and sugars), structural elucidation and analyte identification through fragment fingerprinting.

Collision Induced dissociation (CID)

In the "hard" Ionization methods for mass spectrometry, like EI and CI, spectra usually contain a good amount of fragment ions that can be used to help to confirm or to elucidate chemical structures. In the more modern methods of Ionization, like ESI or MALDI, spectra often only contain the ionized molecule with very little fragmentation data and consequently the spectra are of little use for structural characterization. In these cases, induced fragmentation is required using collision induced dissociation (CID).

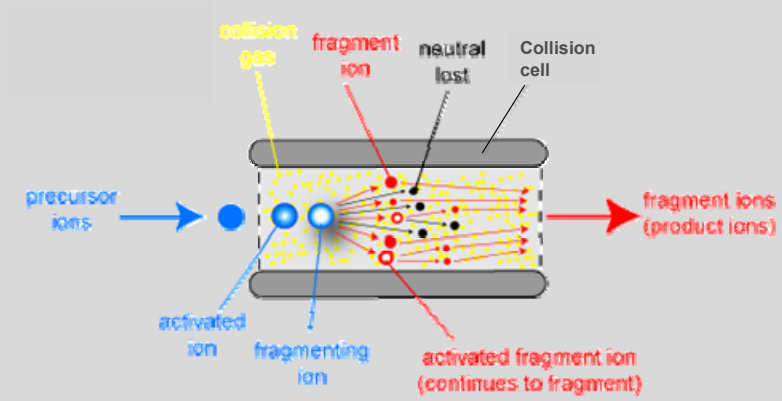


Figure 4.16. Scheme of a collision induced dissociation.

The figure is a schematic representation of the processes involved in CID. The precursor ion enters the collision cell (or in the case of ion-traps, the precursor ion is isolated in the trap) containing a high pressure of an energized, chemically inert collision gas - Ar, He, N_2, CO_2 etc.

The precursor ion undergoes (repeated) collisions with the collision gas, building up potential energy, until when the fragmentation threshold is reached and the product ions are formed. The types of fragmentation that occur vary considerably with the type of product ion and the amount of energy involved. At lower energies (close to the threshold), fragmentation reactions are often limited to neutral losses (H_2O, MeOH, CO, CO_2, CH_3CN etc.) depending on the nature of the precursor ion. These neutral losses are often not considered structurally significant, although they can be used to obtain information about functional groups. At higher energies, retro-synthetic type reactions are often observed. These are much more structurally significant, and often result in cleavage of the molecule at characteristic positions. If the energy is too high, C-C bong cleavage can occur leading to uncontrolled fragmentation and this should be avoided. Usually it is better to work at around the fragmentation threshold, or just above, to maintain the control over the fragmentation processes. Ion-trap and FT-MS instruments allow for the best control over CID, but also tend to produce less energetic reactions. Triple quadrupole and Q/Tof instruments tend to produce more energetic CID with more fragmentation, but less operator control. Ion-trap and FT-MS allow multistage fragmentation experiments to be conducted, which is essential for structural elucidation studies.

The most common combinations of the two analyzers of a MS/MS system are[13]:

- quadrupole – quadrupole, MS^2 in space (Figure 4.17)
- magnetic sector – quadrupole[14],
- magnetic sector - magnetic sector,
- quadrupole - time-of-flight, (Q/Tof, in Figure 4.18),
- time-of-flight - time-of-flight, TOF-TOF;

and the most recent combinations:

- linear trap-ICR (FT-MS),
- linear trap-ORBITRAP[15,16,17] ,
- 3D trap (iontrap).

One of the most commonly available tandem mass spectrometers is the triple quadrupole (QQQ) instrument, whose basic modes of data acquisition are reported in Figure 4.17.

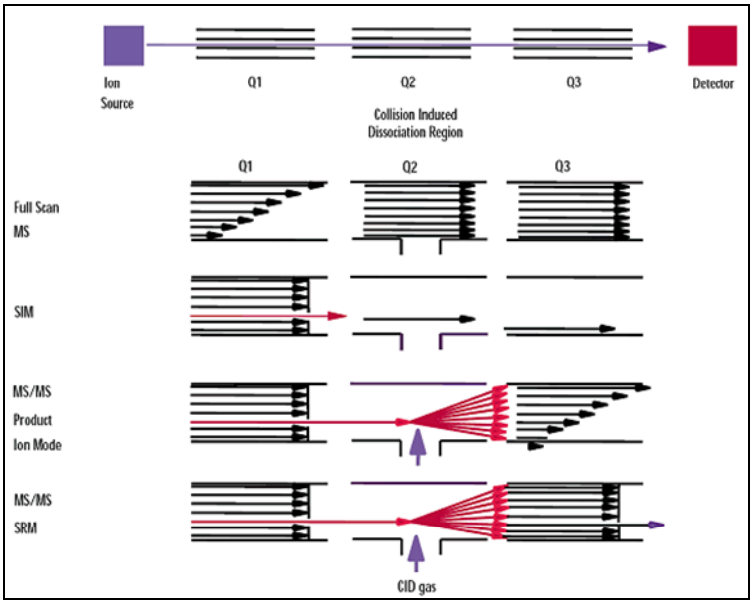


Figure 4.17. Schematic view of a triple quadrupole QQQ and its main ways of working.

Product or daughter ion scanning:

The first analyzer is used to select user-specified sample ions arising from a particular component; usually the molecular-related (i.e. (M+H)+ or (M-H)-) ions. These chosen ions pass into the collision cell, are bombarded by the gas molecules which cause fragment ions to be formed, and these fragment ions are analyzed i.e. separated according to their mass to charge ratios, by the second analyzer. All the fragment ions arise directly from the precursor ions specified in the experiment, and thus produce a fingerprint pattern specific to the compound under investigation. This type of experiment is particularly useful for providing structural information concerning small organic molecules and for generating peptide sequence information.

Precursor or parent ion scanning:

The first analyzer allows the transmission of all sample ions, whilst the second analyzer is set to monitor specific fragment ions, which are generated by bombardment of the sample ions with the

collision gas in the collision cell. This type of experiment is particularly useful for monitoring groups of compounds contained within a mixture which fragment to produce common fragment ions, e.g. glycosylated peptides in a tryptic digest mixture, aliphatic hydrocarbons in an oil sample, or glucuronide conjugates in urine.

Constant neutral loss scanning:

This involves both analyzers scanning, or collecting data, across the whole m/z range, but the two are off-set so that the second analyzer allows only those ions which differ by a certain number of mass units (equivalent to a neutral fragment) from the ions transmitted through the first analyzer. e.g. This type of experiment could be used to monitor all the carboxylic acids in a mixture. Carboxylic acids tend to fragment by losing a (neutral) molecule of carbon dioxide, CO_2, which is equivalent to a loss of 44 Da or atomic mass units. All ions pass through the first analyzer into the collision cell. The ions detected from the collision cell are those from which 44 Da have been lost.

Selected/multiple reaction monitoring

Both analyzers are static: selected specific ions are transmitted through the first analyzer and selected specific fragments arising from these ions are measured by the second analyzer. The compound under scrutiny must be known and have been well-characterized previously before this type of experiment is undertaken. This methodology is used to confirm unambiguously the presence of a compound in a matrix e.g. drug testing with blood or urine samples. It is not only a highly specific method but also has very high sensitivity.

There are many other varieties and configurations of tandem instrument and what here has been described as triple quadrupole, it can be applied with little modification to other systems. Tandem mass spectrometry (MS-MS) is used to produce structural information about a compound by fragmenting specific sample ions inside the mass spectrometer and identifying the resulting fragment ions. This information can then be put together to generate structural information regarding the intact molecule. Tandem mass spectrometry also enables specific compounds to be detected in complex mixtures on account of their specific and characteristic fragmentation patterns. The most common use of MS-MS in biochemical areas is the product or daughter ion scanning experiment which is particularly successful for peptide and nucleotide sequencing. This kind of MS/MS experiments are obtained by Q-TOF, coupled to MALDI or ESI ionization method. MALDI-Q/TOF instrument consists of a MALDI ion source coupled to the quadrupole TOF analyzer. The attraction of this instrument is due to the fact that it combines the ability of MALDI to analyze a large number of samples with the possibility to sequence any selected peptides. However, MALDI produces far fewer ions than electrospray and they are singly charged, which leads to fragmentation patterns that are more complicated than those of the typically multiply charged ES ions.[18] Q-TOF system is shown in Figure 4.18. The molecular ions which enter the first analyzer can be independently selected and transmitted through the quadrupole into the collision cell which lies between the first and second analyzer. An inert gas such as argon is introduced into the collision cell and the sample ions are bombarded by the collision gas molecules which cause them to fragment. The optimum collision cell conditions vary from

analyte to analyte (usually peptides and oligonucleotides) and must be optimized for each. The fragment (or daughter or product) ions are then analyzed by the second (time-of-flight) analyzer, with high resolution if ions are under 2000 Da usually. In this way, an MS-MS spectrum is produced showing all the fragment ions that arise directly from the chosen parent or precursor ions for a given peptide component. Several sequence information can be obtained from each fragmentation spectrum, and the spectra need to be interpreted carefully. Some of the processing can be automated, but in general the processing and interpretation of spectra will take longer than the data acquisition if accurate and reliable data are to be generated.

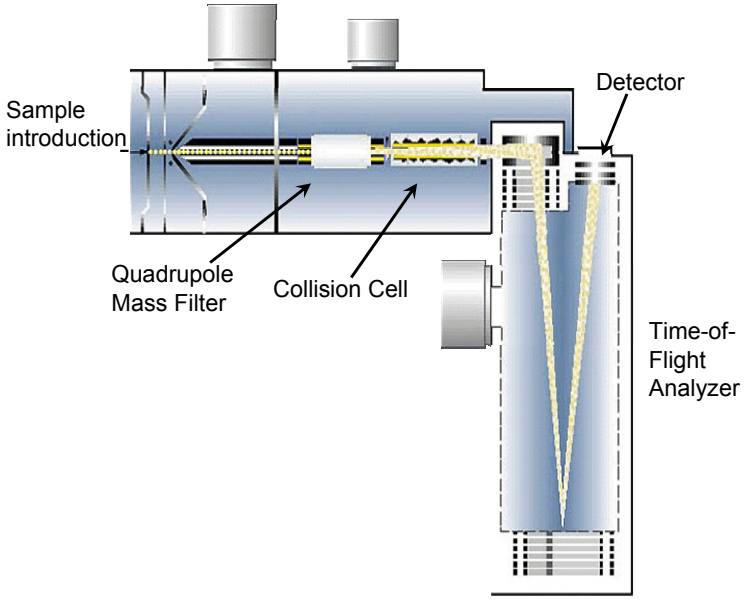


Figure 4.18.Schematic view of a Q-TOF mass analyzer.

MALDI can also be combined with a two-section TOF instrument separated by a fragmentation chamber, a so-called TOF-TOF arrangement. This instrument should have the advantage of very fast analysis time. As in the case of the MALDI quadrupole-TOF, a challenge for this design is the low number of MALDI-produced ions. Additionally, it is more difficult to obtain high-resolution and highly accurate fragmentation spectra after an initial TOF separation compared to the well-defined ion beam produced by the MALDI-quadrupole TOF instrument.[19]

4.5.1 Multistage (Sequential) Tandem Mass Spectrometry (MS^n)

MS^n is a technique that can only be performed on ION-TRAP and FT-ICR instruments, which allow the re-fragmentation of product ions (fragment ions from MS/MS). This kind of instruments allow experiments of MS^2 in time but also MS^n studies. Figure 4.19 shows a schematic view of the MS^n experiment. In the first stage, the normal mass spectrum is produced in the usual way. The isolation of the precursor ion is then performed: it is fragmented by CID in exactly the same way

is in MS/MS. In MS/MS, the experiment would end here as the product ions are consumed by the detector. On the contrary, in MS^n the product ions are trapped allowing another isolation and fragmentation to be performed resulting in the MS^3 spectrum. This process can be repeated a number of times, resulting is a series of MS^n spectra where "n" represents the number of times the isolation-fragmentation cycle is carried out.

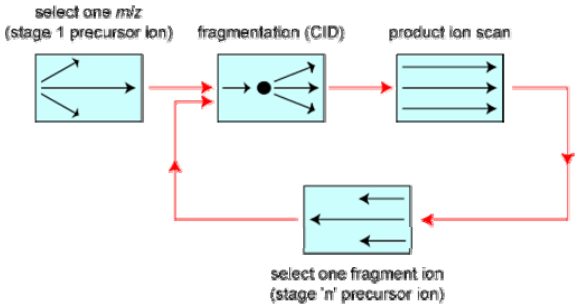


Figure 4.19. A schematic representation of multistage tandem mass spectrometry.

4.6 Structural and sequence information from mass spectrometry.

4.6.1 MW determination

Nominal Mass: is the integral sum of the nucleons in an atom (also called the atomic mass number), calculated from integer mass numbers of the most abundant isotope of each element, e.g. C = 12, H = 1, O = 16. When low molecular weight samples are analyzed using relatively low resolution mass spectrometers, it is common to work with "nominal" mass values, which instead is rarely used in peptide and protein studies because the cumulative error of approximating atomic weights with integers becomes unacceptable.

- **Monoisotopic Mass**: the sum of the exact or accurate masses of the lightest stable isotope of the atoms in a molecule, e.g. C = 12.000000, H = 1.007825, O = 15.994915.
- **Average Mass**: the sum of the average of the isotopic masses of the atoms in a molecule, e.g. C = 12.01115, H = 1.00797, O = 15.9994.

The presence of isotopes at their natural abundances makes it essential to define whether an experimental mass value is an "average" value, equivalent to taking the centroid of the complete isotopic envelope, or a "monoisotopic" value, the mass of the first peak of the isotope distribution.

4.6.2 Resolution

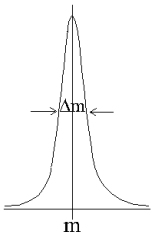
Resolution is the ability of a mass spectrometer to distinguish between ions of different m/z ratios, or the dimensionless ratio of the mass of the peak divided by its width.

$$R=m/\Delta m$$

Δm is the mass difference between two adjacent peaks that are just resolved, m is the mass of the first peak (or the mean mass of two peaks); although this definition is for two peaks, it is acceptable to measure the resolution from a single peak. In that case Δm is the width of the peak at half maxima (FWHM) of the peak corresponding to m.

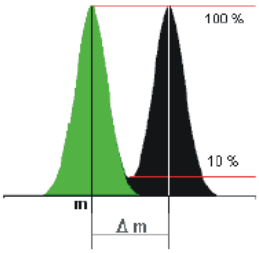
Single Ion method to calculate R

Full Width at Half Maximum (FWHM) or at 5% of the peak height .



Double Ion method to calculate R

2 adjacent ion peaks with a 10% valley max



The mass resolution achievable by a mass spectrometer depends on both the type of analyzer and the experimental conditions. Simple MALDI-TOF instructments may only achieve unit mass resolution over a limited mass range. High performance FTMS systems can achieve resolving powers of 10^5.

4.6.3 Accuracy

Mass accuracy indicates the accuracy of the mass information provided by the mass spectrometer: is the difference which is observed between the theoretical mass and the measured mass.

$$\Delta M \text{ accuracy} = M_{real} - M_{measured}$$

Accuracy is often expressed in parts per million (ppm).

$$ppm = 10^6 * \Delta M \text{ accuracy} / M_{measured}$$

Mass accuracy is linked to the resolution. A low resolution instrument cannot provide a high accuracy. A high resolution instrument (time of flight, sector, FTMS), properly used with a reference compound, provides the mass information with an accuracy better than 5 ppm, which is enough to unambiguously determine the elemental composition.

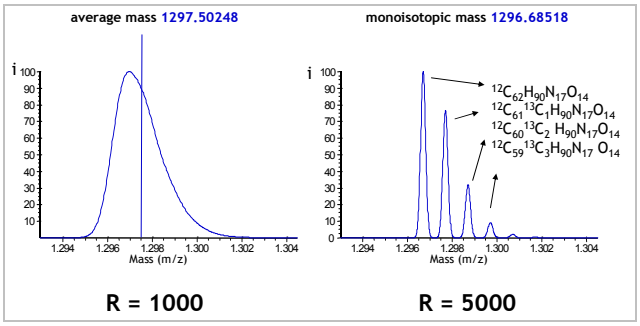


Figure 4.20. Examples of different resolution and consequently determinations of mass, average or monoisotopic.

4.6.4 Calibration

Spectrometers need calibration. Some of them keep the calibration for long time, others need ongoing calibration. Compounds used for calibration include: PEG, PBM, peptides, proteins, PFTBA, CsI

External Calibration: m/z scale is calibrated with a mixture of molecules with different molecular weights; after that the analyte is measured.

Internal Calibration: analyte and a mixture of molecules with different molecular weights are mixed and measured together. Then the spectrum is calibrated by assigning the right masses to the well known calibration standards (perfect: mass of analyte is between the mass of two standards).

4.6.5 Identification of ions formed in mass spectrometry

i m/z

An abbreviation used to denote the dimensionless quantity formed is m/z, obtained by dividing the mass of an ion by the number of charges carried by the ion. It has long been called the mass-to-charge ratio, although m is not the ionic mass nor is z a multiple of the electronic charge. Samples (M) with molecular weights greater than ca. 1200 Da give rise to multiply charged molecular-related ions such as $(M+nH)^{n+}$ in positive ionisation mode and $(M-nH)^{n-}$ in negative ionisation mode. If the sample has functional groups that readily accept a proton (H^+) then positive ion detection is used (e.g. amino groups in proteins and peptides can accept H^+: $R\text{-}NH_2 + H \rightarrow R\text{-}NH_3^+$). If the sample has functional groups that readily loose a proton then negative ion detection is used (e.g. carboxylic acids in saccharides and oligonucleotides can loose H^+: $R\text{-}CO_2H \rightarrow R\text{-}CO_2^-$; alcohols $R\text{-}OH \rightarrow R\text{-}O^-$).

ii Singly-, doubly-, triply-, etc. charged ions

These ions are molecule or molecular moieties which have gained or lost respectively one, two, three or more electrons/protons. In the example it is reported a protonated ion.

$$m/z = \frac{(MW + nH^+)}{n}$$

In MALDI ionization techniques, 1 charge per molecule is generally acquired; in ESI ionization technique, multiple charges are produced in the same molecule, generating spectra as the one in Figure 4.21. In this case, "n" ranges from 21 (giving a measured mass of 808.221) to 12 (corresponding to a measured mass of 1413.631). The peaks with measured masses between these, correspond to the other values of "n" between 12 and 21. By taking successive pairs of measured masses, the relative molecular mass of the myoglobin can be calculated very accurately.

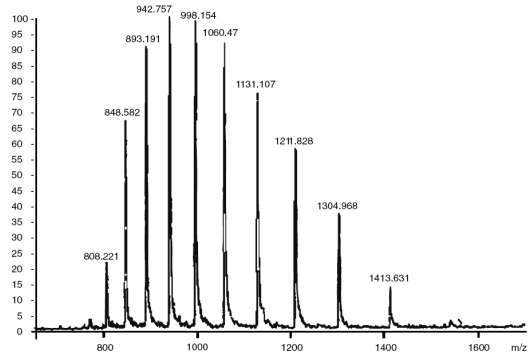


Figure 4.21. [M + nH$^+$] ions from a protein (myoglobin) of molecular mass 16,951.5 Da.

iii MW determination from a given ion.

When the m/z value and the "n" number of charges, positive or negative, acquired are known, the MW is obtained by the following equation:

$$MW = (m/z * n) - nH^+$$

If "n" is not known, we must apply a mathematical system involving two consequent ions, m/z_a and m/z_b, having n_a and n_b numbers of charges H^+ but which took origin from the same protein (thus, same MW).

$$\begin{cases} MW = (m/z_a * n_a) - n_aH^+ \\ MW = (m/z_b * n_b) - n_bH^+ \\ a = b + 1 \end{cases}$$

This "system" is commonly done by opportune software with which all the components (m/z) are transposed onto a true molecular weight (or zero charge state) profile from which molecular weights can be read directly without any amendments or calculations. High resolution spectrometers overcome the necessity of this calculations, as they resolved multiprotonated ions The method in this case is a "Transformation" and it involves the manual or automatic identification of all the components in the mixture from the m/z spectrum, followed by transforming the spectra onto a molecular weight profile by automatically mass-correcting ions from each individual charge state before collating all of them onto one single molecular weight profile.

4.7 Peptide and protein identification by Mass Spectrometry

Peptides fragment in a reasonably well-documented manner[20,21]. The protein can be cleaved in the mass spectrometer according to three different ways (as described in Figure 4.22): the NH-CH, CH-CO, and CO-NH bonds.

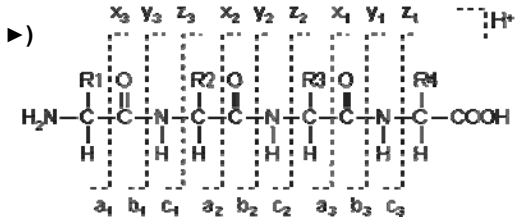


Figure 4.22. The main types of fragment ions observed in an MS/MS spectrum of a peptide.

Each bond breakage gives rise to two species, one neutral and the other one charged, and only the charged species is monitored by the mass spectrometer. The charge can stay on either of the two fragments depending on the chemistry and relative proton affinity of the two species. Hence there are six possible fragment ions for each amino acid residue and these are labelled as in the Figure 4.23, with the a, b, and c" ions with the charge retained on the N-terminal fragment, and the x, y", and z ions with the charge retained on the C-terminal fragment. A subscript indicates the number of residues in the fragment. The most common cleavage sites are at the CO-NH bonds which give rise to the b and/or the y" ions. The mass difference between two adjacent b ions, or y" ions is indicative of a particular amino acid residue. “c” and “y” ions abstract an additional proton from the precursor peptide.

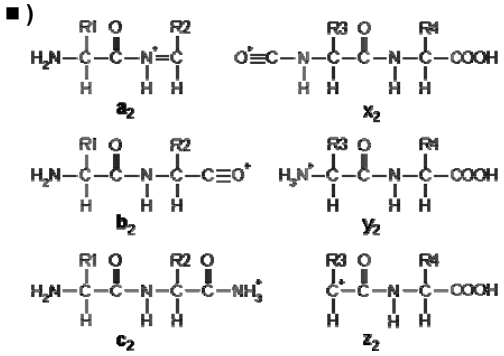


Figure 4.23. Main fragment ions generated by dissociation of a peptide backbone.

In low energy CID (i.e. collision induced dissociation in a triple quadrupole or an ion trap) a peptide carrying a positive charge fragments mainly along its backbone, generates predominantly a, b and y ions. In addition, peaks are seen for ions which have lost ammonia (-17 Da) denoted a*, b* and y* and water (-18 Da) denoted a°, b° and y°. Satellite ions from side chain cleavage are not observed in soft ionizations.

Immonium ions (labelled “i”) appear in the very low m/z range of the MS-MS spectrum. They are internal fragments with just a single side chain formed by a combination of type a and type y cleavages (Figure 4.24). Each amino acid residue leads to a diagnostic immonium ion with the exception of the leucine and iso-leucine, lysine (K) and glutamine (Q), which produce immonium ions with the same m/z ratio, i.e. m/z 86 for I and L, m/z 101 for K and Q. The immonium ions

are useful for detecting and confirming many of the amino acid residues in a peptide (Table 4.2), although no information regarding the position of these amino acid residues in the peptide sequence can be ascertained from the immonium ions.

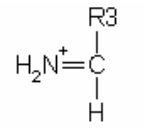


Figure 4.24. Immonium ion.

The extent of side-chain fragmentation detected depends on the type of analyzers used in the mass spectrometer. For example, a quadrupole - quadrupole and quadrupole - time-of-flight mass spectrometers generate low energy fragmentations with few side-chain fragmentations. New, additional ions (Figure 4.25) are observed in high energy collision spectra. Relative abundances are composition dependent.

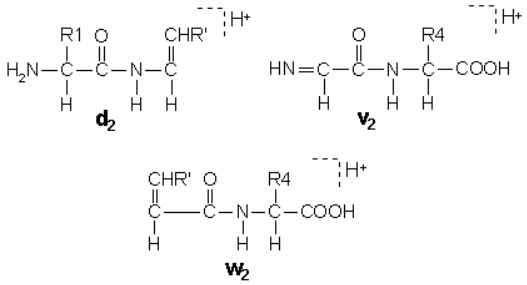


Figure 4.25. CID of ions at high energies can generates additional ion types due to side chain cleavage.

Table 4.1. Formulae to Calculate Fragment Ion m/z values: [N] is the molecular mass of the neutral N-terminal group, [C] is the molecular mass of the neutral C-terminal group, [M] is molecular mass of the neutral amino acid residues.

IonType	Neutral Mr
a	[N]+[M]-CHO
a*	a-NH_3
a°	a-H_2O
b	[N]+[M]-H
b*	b-NH_3
b°	b-H_2O
c	[N]+[M]+NH_2
d	a - partial side chain
v	y - complete side chain
w	z - partial side chain
x	[C]+[M]+CO-H
y	[C]+[M]+H
y*	y-NH_3
y°	y-H_2O
z	[C]+[M]-NH_2
To obtain m/z values, protons are added or subtracted to obtain the required charge, divided by the number of charges	

Table 4.2. AA residues masses and related fragments. Bold indicates strong signals, *italic* indicates weak.

Residue	3letter code	1letter code	Mono-isotopic mass	Average mass	Immonium ion	Related ions
Alanine	Ala	A	71.03712	71.08	*44*	
Arginine	Arg	R	156.10112	156.19	*129*	59,70,73,87,100,112
Asparagine	Asn	N	114.04293	114.10	*87*	*70*
Aspartic acid	Asp	D	115.02695	115.09	*88*	*70*
Cysteine	Cys	C	103.00919	103.14	*76*	
Glutamic acid	Glu	E	129.04260	129.12	*102*	
Glutamine	Gln	Q	128.05858	128.13	101	56,84,129
Glycine	Gly	G	57.02147	57.05	30	
Histidine	His	H	137.05891	137.14	**110**	82,121,123,138,166
Isoleucine	Ile	I	113.08407	113.16	**86**	*44,72*
Leucine	Leu	L	113.08407	113.16	**86**	*44,72*
Lysine	Lys	K	128.09497	128.17	*101*	*70*,84,112,129
Methionine	Met	M	131.04049	131.19	*104*	61
Phenylalanine	Phe	F	147.06842	147.18	**120**	*91*
Proline	Pro	P	97.05277	97.12	**70**	
Serine	Ser	S	87.03203	87.08	60	
Threonine	Thr	T	101.04768	101.10	74	
Tryptophan	Trp	W	186.07932	186.21	**159**	*77*,117,**130**,132,**170,171**
Tyrosine	Tyr	Y	163.06333	163.18	**136**	*91,107*
Valine	Val	V	99.06842	99.13	*72*	41,55,69
Selenocysteine	SeC	U	150.95364	150.03	103	

4.8 High Resolution MS techniques

Some of the most recent MS techniques have been utilized in this work, and are here reviewed.

4.8.1 LTQ-ORBITRAP

LTQ-ORBITRAP has been described recently[15,22] as an hybrid spectrometer which couples a linear ion trap to an Orbitrap mass analyzer. The ion storage and injection into the Orbitrap allows high resolving power, mass accuracy, and transmission over a wide dynamic range and forms the basis for a hybrid mass spectrometer combining these analytical parameters with the MS^n capability of the linear ion trap mass spectrometer. The LTQ-Orbitrap mass spectrometer (Figure 4.26) is constituted by an atmospheric pressure ion source, a transfer octapole (1), a curved rf-only quadrupole (C-trap) (2), a gate electrode (3), a trap electrode (4) and ion optics (5) which send ions between the inner (6) and the outer (7) Orbitrap electrode. A first set of ions is injected and trapped in the C-trap (a); a second set of ions is injected and trapped in the C-trap (b); a mixed ion population is pulsed into the rapidly changing electric field of the Orbitrap (c) and when, finally, a stable electrostatic field is achieved, detection is subsequently performed (d).

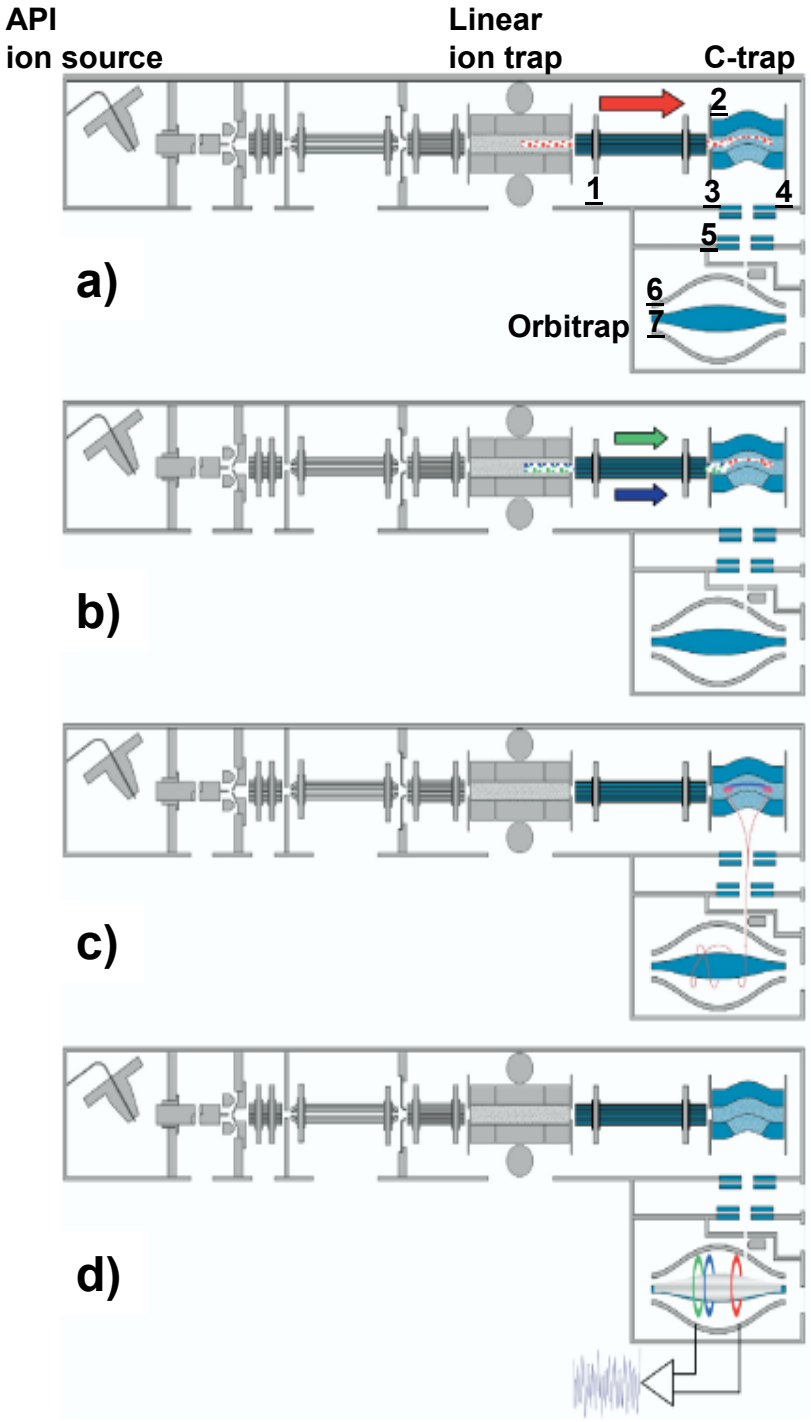
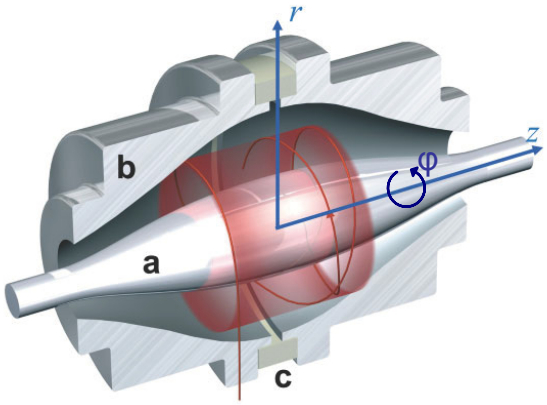


Figure 4.26. Schematic layout of the LTQ Orbitrap mass spectrometer and experimental sequence for measurements of the extent of mass accuracy the Orbitrap mass analyzer. (Adapted from ref.16; letters and numbers explained in text).

As its name suggests, Orbitrap (in detail in Figure 4.27) is a device that is able to store and trap ions[23]. It is not a conventional ion trap as a three-dimensional or a linear ion trap[24], because there is neither RF nor a magnet to hold ions inside, but an electrostatic field that traps ions.[25] The electrostatic attraction towards the central electrode is compensated by a centrifugal force that arises from the initial tangential velocity of ions, which makes ion moving like a satellite on orbit.



Angular frequency

$$\omega_r = \omega_z \sqrt{\left(\frac{R_m}{R}\right)^2 - 2}$$

Radial frequency

$$\omega_\varphi = \frac{\omega_z}{\sqrt{2}} \sqrt{\left(\frac{R_m}{R}\right)^2 - 1}$$

Axial frequency

$$\omega_z = \sqrt{\frac{k}{m/q}}$$

Figure 4.27. Schematic representation of an Orbitrap mass analyzer and characterizing frequencies.

The electrostatic field forces ions to move in complex spiral patterns. The axial component of these oscillations can be detected as an image current on the two halves of an electrode encapsulating the Orbitrap. A Fourier transform is employed to obtain oscillation frequencies for ions with different m/z values, which can be determined from these values. Since ions can be trapped for long times, the frequency of their image current can be registered with high accuracy, allowing to obtain high resolution mass spectrum. The resolution and mass accuracy are very competitive with those achievable in other instruments as FT-ICR mass spectrometers and better than in TOF ones.

The combination of this mass analyzer with linear ion traps is perfect, because the first allows high resolution measurements, whilst the second allows to perform MS^n experiment. Thus, also product ions of these experiments can be registered with high resolution and high mass accuracy, opening new perspectives in a wide range of applications.

Due to its characteristics, Orbitrap mass spectrometers allow for a "*Top Down*" approach in proteomics[26]. In proteome analysis , so far, the use of instrumentation other than FT-ICR allowed only the classical "*bottom up*" approach, a strategy involving the cleavage the protein into peptide fragments sufficiently distinctive to allow protein identification. In these cases,

complete sequence coverage of proteins is rarely achieved. In the "*Top down*" approach, on the other side, intact proteins are ionized and fragmented in mass spectrometer, allowing for their exhaustive characterization, without the need of chemical or enzymatic proteolysis, but it requires high resolution mass spectrometers in order to determine charge states of multiple charged protein ions. The Orbitrap has a sufficiently high resolving power to encourage attempts for analyzing intact proteins. The typical mode of operation involves measuring the intact protein mass in the Orbitrap, selecting one of the charge states of the protein and fragmenting it in the ion trap, then measuring the fragments in the Orbitrap. This allows an unambiguous charge state determination of fragment ions and identification of most of the unmodified and modified proteins by database searching.

4.8.2 MALDI Mass Spectrometry Imaging (MSI)

Recent developments in mass spectrometric technologies led to Mass Spectrometry Imaging (MSI), a powerful alternative to classical techniques commonly used to localize biomolecules in biological tissues[27], as immunohistochemistry: even if it achieves a lower resolution, (15 μm against < 1 μm), it has the great advantage to localize in a fast way more than one protein at time, avoiding the need of a specific antibody for each protein under analysis. Moreover, in immunohistochemistry analysis, where there is the risk of cross-reactivity leading to false positives, MSI localizes a given molecule by means of one of its intrinsic property, i.e. its molecular mass.[28] MALDI Mass Spectrometry Imaging (MALDI MSI)[29] is able to generate molecular profiles and two-dimensional ion density maps of biomolecules directly from the surface of tissue sections, combining the chemical specificity and parallel detection of mass spectrometry with microscopic imaging capabilities.[30] This allows specific information to be obtained on the relative abundance and spatial distribution of a wide range of molecules both endogenous and xenobiotics.

An outline of the technique is reported in Figure 4.28; typically, a thin tissue slice (10-20μm) is covered with a matrix which co-crystallizes with the molecules present in the tissue. Commonly, a 337 nm N_2 laser is fired in a defined raster over the section; absorption of the UV radiation from the laser pulse by the crystals subsequently causes matrix and analyte molecules to desorb from the sample surface, and a mass spectrum, containing all the mass signals of the compounds desorbed, is acquired for each point of the raster. A data set consisting of an ordered array of mass spectra is created, where each spectrum represents the local molecular composition at known *x*,*y* coordinates. Ultimately, an image can be generated for each of the mass signals detected throughout the section. In particular, the intensities of individual *m*/*z* values in each spectrum, corresponding to the molecular weights of specific compounds, can be extracted to produce images of the areas, within the tissue, in which that particular molecule was located, similarly to the digital imaging in photography where each image is composed of an ordered array of thousands of pixels.

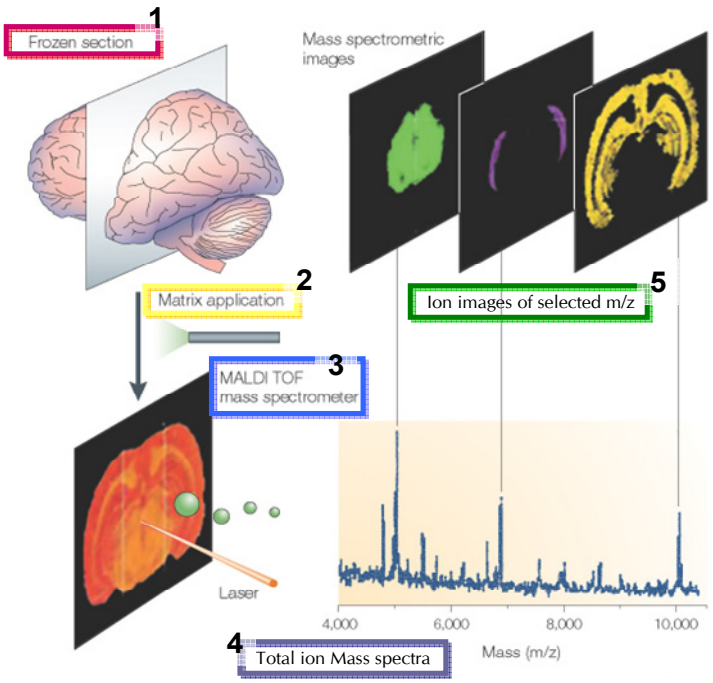


Figure 4.28. The principle of MALDI imaging (adapted from Nature Reviews Genetics, 2005[31]).

Since molecular maps can be obtained for one or more molecules simultaneously without the use of chemical probes, IMS is highly complementary to immunohistochemistry and *in vivo* imaging techniques such as Magnetic Resonance Imaging, Positron Emission Tomography, near IR-fluorescence Imaging, and Whole Body Autoradiography (WBA). Moreover, MSI can be applied to metabolites, drugs, peptides and proteins with no difficulties.

4.9 References

[1] http://hplc.chem.shu.edu/HPLC/index.html

[2] Ashcroft A. E., *Ionization Methods in Organic Mass Spectrometry*, Neil W Barnett (Ed) The Royal Society of Chemistry, UK RCS publisher, **1997**

[3] Hillenkamp F., Karas M., Beavis R. C., Chait B. T., "Matrix-Assisted. Laser Desorption/Ionization Mass Spectrometry of Biopolymers",*Anal. Chem*., **1991**, 63, 1193–1203

[4] Yamashita M. and Fenn J. B., *J. Phys. Chem*., **1984,** 88, 4451.

[5] Lee T. D., Moore R. E., and Young M. K., UNIT 16.8 Introducing Samples Directly into Electrospray Ionization Mass Spectrometers Using a Nanospray Interface AND UNIT 16.9 Introducing Samples Directly into Electrospray Ionization Mass Spectrometers Using Microscale Capillary Liquid Chromatography, *Current Protocols in Protein Science*, ed. John Wiley and Sons, Inc. **1998**

[6] M. Wilm, M. Mann, Analytical Properties of the Nanoelectrospray Ion Source, **1996,** *Anal. Chem*., 68, 1-8

[7] http://www.chm.bris.ac.uk/ms/theory/theory.html

[8] Wiley W.C. & MacLaren I.H.; *The Review of Scientific Instruments*, **1955**; 26; p1150.

[9] W. Paul, *Agewandte Chemie - International Edition*, 29; **1990**, p739.

[10] J.A. Hipple et al.; *Physical Review*, **1949**, 76; p1877.

[11] A.G. Marshall, Accounts of Chemical Research, 29; **1996**, p308.

[12] WIZA J. L. Microchannel plate detectors , *Nuclear Instruments and Methods*, **1979,**162, 587- 601

[13] Burlingame A. L., Boyd R.K., Gaskell S. J. Mass Spectrometry, *Anal. Chem.* **1996,** 68, 599-651

[14] Bateman R. H., Green M. R., Scott G., Clayton E. , A combined magnetic sector-time-of-flight mass spectrometer for structural determination studies by tandem mass spectrometry, *Rapid Communications in Mass Spectrometry* **1995**, 9, 13, 1227–1233

[15] Makarov A., Denisov E., Kholomeev A., Balschun W., Lange O., Strupat K., and Horning S. Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer *Analytical Chemistry*, **2006**, 78, 7, 2113-2120

[16] Makarov A., Denisov E., Lange O., Horning S Dynamic Range of Mass Accuracy in LTQ Orbitrap Hybrid Mass Spectrometer *J Am Soc Mass Spectrom*. **2006.**17, 977–982

[17] Scigelova M. and Makarov A. Orbitrap Mass Analyzer. Overview and Applications in Proteomics *Practical Proteomics* **2006** 1/2,16-21

[18] Shevchenko A, Loboda A, Ens W, Standing KG. Advances in Proteome Analysis by Mass Spectrometry. *Anal. Chem*. **2000** 72:2132–41

[19] Medzihradszky KF, Campbell JM, Baldwin MA, Falick AM, Juhasz P, et al. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal. Chem.* **2000,** 72:552–58

[20] P. Roepstorrf, J. Fohlmann, *Biomed. Mass Spectrom*., **1984,** 11, 601

[21] R. S. Johnson, K. Biemann, *Biomed. Environ. Mass Spectrom*., **1989**, 18, 945

[22] Hu Q., Noll R. J., Li H., Makarov A., Hardman M., Graham Cooks R. The Orbitrap: a new mass spectrometer, *Journal of Mass Spectrometry* **2005** 40, 4: 430 – 443

[23] Makarov, A., Electrostatic axially harmonic orbital trapping: a high-performance. technique of mass analysis., *Anal. Chem.* **2000,** 72, 1156–1162.

[24] Douglas DJ, Frank AJ., Mao D. Linear ion traps in mass spectrometry.*Mass Spectrometry Reviews*, 2005, 24, 1-29

[25] Hardman, M., Makarov, A., Interfacing the orbitrap mass analyzer to an electrospray ion source. *Anal. Chem.* **2003,** 75, 1699–1705.

[26] Kjelden F., Zubarev R.A. *The Encyclopedia of Mass Spectrometry,* **2005,** Elsevier, , vol. 2-Part A, 190-197

[27] Caprioli R.M., Farmer T.B., Gile J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* **1997**, 69, 4751-4760

[28] Chaurand, P., Sanders, M. E., Jensen, R. A., and Caprioli, R. M. Proteomics in diagnostic pathology: profiling and imaging proteins directly in tissue sections. *Am. J. Pathol.* **2004** 165, 1057–1068

[29] McDonnell L. A. and Heeren R. M.A., Imaging Mass Spectrometry *Mass Spectrometry Reviews***, 2007**, 26, 606– 643

[30] Kumar A., Agarwal S., Heyman J.A., Matson S., Heidtman M., Piccirillo S., Umansky L., Drawid A., Jansen R., Liu Y., Cheung KH., Miller P., Gerstein M., Roeder GS., Snyder M., Subcellular localization of the yeast proteome. *Genes & Development,***2002**, 16, 707.

[31] Miniaturization in functional genomics and proteomics Sascha Sauer, Bodo M. H. Lange, Johan Gobom, Lajos Nyarsik, Harald Seitz & Hans Lehrach Nature Reviews Genetics 6, 465-476 (June 2005)

5 PROTEOMICS

The term ''proteome'' (PROTEins expressed by a genOME) refers to all proteins produced by a species, much as the genome is the entire set of genes. This term was coined by Wasinger[1] in 1995, and nowadays means the analysis of the entire PROTEin complement expressed by an organism's genOME, by a cell or tissue type.

Two-dimensional gel electrophoresis (2DE) separation of complex protein mixtures and the subsequent analysis of isolated protein spots by mass spectrometry allow fast and accurate identification of proteins. The comparison of spots from different samples separated on customized 2D gels allows the detection of punctual differences in their mobility and facilitates tracing back differences in protein expression, presence of isoforms, splice variants and posttranslational modifications by mass spectrometry.

The proteomics methodology develops as follows[2]:

1. sample collection, handling and storage.
2. sample preparation,
3. protein separation (2DE).
4. protein identification (peptide mass fingerprinting -PMF- and mass spectrometry).
5. protein characterization (amino acid sequencing).
6. bioinformatics (cross reference of protein informatics with genomic databases).

Mass spectrometry is heavily involved in proteomics, as it is summarized in Figure 5.1.

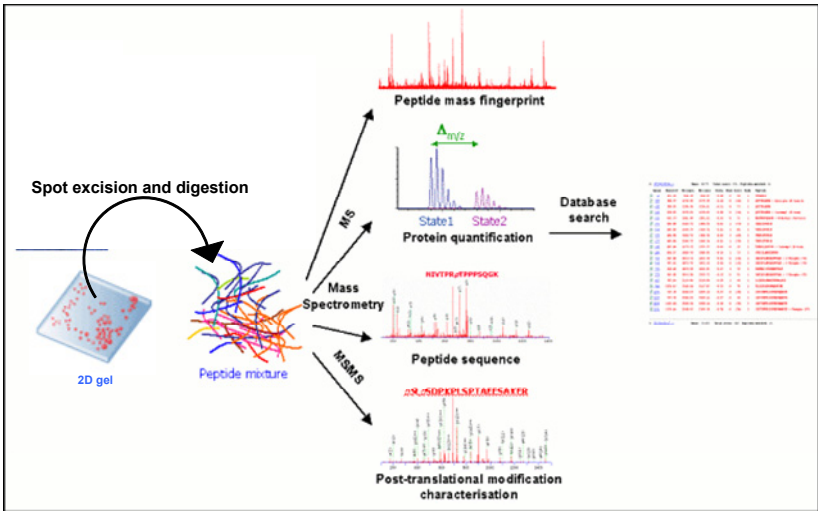


Figure 5.1. Mass-spectrometry-based proteomics.

The 2DE/MALDI-MS technique, although robust and relatively straightforward, does suffer from some fundamental limitations for comprehensive proteome characterization. First, there are

finite limits to the hydrophobicity, isoelectric point (pI), and molecular weight range of proteins resolvable using conventional two dimensional gels. Conventional two-dimensional gels can separate on the order of »1000 components during a single separation; however, incomplete separation of proteins on the gel can lead to overlapping spots and subsequently to problems with protein identification using the MALDI-MS peptide mass mapping approach. A general principle of the 2DE approach is that protein spots must be visible after gel staining in order to be subsequently identified by MALDI-MS.

A powerful alternative approach is the use of tandem mass spectrometry (MS/MS) to induce fragmentation of individual tryptic peptides after online liquid phase separation. Typically this is performed using triple-quadrupole (TQ), quadrupole ion trap (QIT), Fourier transform ion-cyclotron resonance (FTICR), or quadrupole time-of-flight (QTOF) mass spectrometers. It is now well established that low-energy collision-induced dissociation (CID) fragmentation of peptides provides useful information for identification related to the amino acid sequence. These methods can provide highly confident protein identifications in complex mixtures, typically requiring only a single to at most several peptide MS/MS spectra to identify a protein in a constrained database.

One approach for obtaining high-resolution separation of peptides prior to mass spectrometric detection is an orthogonal combination of cation exchange and reversed-phase chromatography, in a microcapillary format (multidimensional liquid chromatography, generally 2D-LC).

5.1 Recent developments in proteomics

5.1.1 Quantitative proteomics

Mass spectrometry is not inherently quantitative because proteolytic peptides exhibit a wide range of physicochemical properties, such as size, charge, hydrophobicity, etc. which lead to large differences in mass spectrometric response. All of the mass-spectrometry-based quantification methods have their particular strengths and weaknesses, but they are beginning to mature to an extent that they can be meaningfully applied to the study of biological systems on a proteomic scale. In contrast, the statistical treatment and subsequent interpretation of quantitative proteomic data are still in their infancy, as the field is only beginning to experience the particular challenges associated with transforming qualitative protein identification and post-translational modification data into reliable quantitative information.

5.1.2 Bottom-Up and Top-Down Proteomics

Strategy of protein identification by Bottom Up approach is the referred to typical peptide mass fingerprinting. The unknown protein is excised from a gel and converted to peptides by the action of a specific protease. The mass of the peptides produced is then measured in a mass spectrometer, or by MS/MS. The mass spectrum of the unknown protein is searched against

theoretical mass spectra produced by computer-generated cleavage of proteins in the database. This strategy depend critically on the conversion of proteins to peptides via enzymatic or chemical digestion prior to mass spectrometric analysis (so called bottom-up approach). It is typically easier to achieve high mass measurement accuracy and to perform routine MS/MS experiments on lower-molecular-weight molecules such as peptides than high-molecular-mass proteins, using conventional mass spectrometers. However, such digestion methods also result in greatly increased sample complexity of global protein extracts. In addition, when using the bottom-up strategy to perform qualitative proteome analysis, complete sequence coverage of proteins is rarely achieved. Given these limitations, there has been increasing interest in top-down protein characterization strategies where individual proteins are selected for analysis by mass spectrometry without the need for prior chemical or enzymatic proteolysis.

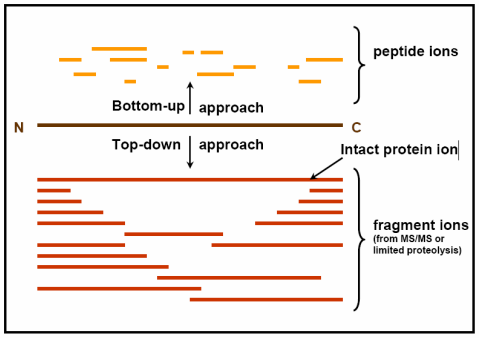


Figure 5.2. Resumed bottom-up and top-down approaches.

5.1.3 Shotgun proteomics

Despite the longstanding success of 2D-PAGE coupled with mass spectrometry, several fundamental issues with the technology, including the challenges of identifying low-abundance proteins, membrane proteins, and proteins with extremes in isoelectric point (pI) and molecular weight (MW), drove researchers to develop alternative approaches for the separation of complex mixtures.

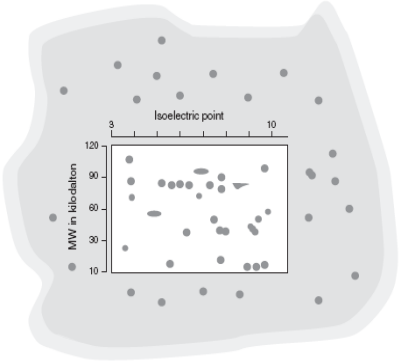


Figure 5.3.Lost proteins in 2D gel electrophoresis.

Shotgun proteomics[3,4] is based on the coupling of high-performance liquid chromatography (HPLC) and mass spectrometry (MS). Similar to the shotgun genomic sequencing approach in which DNA is broken into smaller pieces prior to sequencing and reassembled in silico, proteins are first digested into peptides and then analyzed by multidimensional chromatography coupled to tandem mass spectrometry (MS/MS). Proteomic studies of such complex peptide mixtures result in thousands to millions of MS/MS spectra. The analysis of such large datasets is a daunting task, and interpretation of such complex uninterpreted MS/MS spectra requires sophisticated algorithms. Between the many database searching algorithms for proteomics analysis the most commonly used algorithms are SEQUEST, MASCOT, OMSSA and X!Tandem.[5]
Between proteomics, the maximum increase of the peak capacity can be obtained only by carrying out multidimensional separation based on different separation mechanisms. Two-dimensional column switching liquid chromatography consists in two columns coupled with a switching valve. In column-switching two-dimensional systems a loop or a column is generally used to trap a fraction of interest, where a switching valve is employed to inject this fraction into the second dimensional separation column. The separation of peptide mixtures by LC/LC methods has been performed using several orthogonal combinations such as strong cation exchange/reversed phase liquid chromatography (SCX/RPLC), anion exchange chromatography/reversed phase liquid chromatography (AE/RPLC), size exclusion chromatography/reversed phase liquid chromatography (SEC/RPLC), and affinity chromatography/reversed phase liquid chromatography (AFC/RPLC). In most shotgun proteomics analyses, the second dimension is performed by RPLC because the mobile phase is compatible with the mass spectrometer. For column-switching two-dimensional liquid chromatography (2D-LC), the second dimension should be fast enough to complete analyzing the fraction eluted from the first dimension. To reduce the analysis time of the second dimension, novel stationary phases for fast separation have been used for the second dimensional separation.

5.1.4 Perspectives of mass spectrometry in protein studies

The recent technical advancement in mass spectrometry instrumentation allows this technique to play a key role in protein structural biology, adding information complementary to the genomic/structural ones to better elucidate biological problems. In fact the variety of information that can be given by mass spectrometry is noticeably wide; for example hydrogen/deuterium exchange and limited proteolysis protocols can provide properly structural information, although at low resolution, indicating the regions that are more flexible or exposed to the solvent.[6] Moreover nowadays, advancement in ionization techniques, especially electrospray, allows to explore heterogeneous and large protein assemblies and to obtain information on their stoichiometry and strength of interaction. Another kind of information can be simply the peptide sequence determination and/or the determination of its variation between protein isoforms, without the need of accurate sample purification; also, exploration of

complexes between proteins and their ligands or drugs can be performed by mass spectrometry, opening new possibilities to determine stoichiometry and strength of their interaction. In addition, the ability to analyze inhomogeneous samples and complex protein mixtures offers possibilities for less focused approaches. Finally, the recent advances in data handling software, mass spectrometry instrumentation and histological sample preparation make possible to analyze proteins directly on fresh tissue sample (Mass Spectrometry Imaging)[7]. This approach couples many of the above mentioned mass spectrometry potentialities, allowing a molecular view of the spatial distribution of all the detectable proteins present in the tissue. This opportunity allows classifying different tissues or regions of the same tissue on the basis of their molecular profile, obtained directly from a histological sample, and eventually proceeding with molecular characterization and identification of signals of interest. Moreover this approach is not limited to protein analysis, but it can be also applied to small molecules as drugs and lipids, giving new information on their delivery pathways complementing data obtained from other approaches.

5.2 Proteomics in food quality evaluation

Applied genomics technologies (transcriptomics, proteomics, metabolomics, nutrigenomics, etc.) will contribute to research areas of the nutritional sciences and food technology[2]:

1. Screening for novel functional bioactivities. → Availability of new rapid screening methods for detection of bioactivity.
2. Safety evaluation of food ingredients. → Evaluation of absorption, body distribution and metabolism of food ingredients.
3. Detection and control of food spoilage or pathogenic microorganisms. → Identification of biomarkers (genes, proteins, metabolites) representative for specific food spoilage and/or pathogenic microorganisms.
4. Efficacy testing of bioactive functional food ingredients. → Changes in gene expression and proteome relevant to the states or treatment of certain diseases.
5. Food allergy. →Identification of allergic proteins through sophisticated proteomics based on recognition of specific posttranslational modification and digestion-resistant peptide features.
6. Quality and authenticity of foods. → Proteome of certain food (wheat, wine, fish) can be used to authenticate food origin or food quality.
7. Production of food ingredients. → The yield of bioprocesses (production of amino acids, carbohydrates, etc.) may be controlled through metabolome/proteome of microorganisms used for such production.
8. Food processing. → Proteome and/or metabolome of starter culture of fermentation processes (beer, cheese, sausage, etc.) can be used to predict the quality of the fermented end-product.

Proteomic techniques offer a new approach to identify protein in food matrices and to study protein–protein interactions in raw and processed foods, as well as interactions between proteins and other food components. Because it provides a sensitive measure of changes in protein structure occurring at specific amino acid residues, it can be used to map covalently bound constituents that may be produced during processing events[8]. Additionally, 2-D mapping under various extraction procedures can be used to gain information on non-covalent interactions. Also, biomolecules or ratios between specific biomolecules in the raw material can be identified that are most critical to the quality of the end-product.

Bioactive compounds, ingredients for functional foods, can be screened by the proteomic techniques. Specific applications of proteomics in analysis of food quality are already reported in the literature and some examples are here reviewed.

5.2.1 Proteomics in meat science

Proteomics offers a powerful tool to identify the specific gene products that may be involved in meat quality alterations. 2D-PAGE, coupled with MS, allows for the identification and characterization of marker proteins and enzymes, as well as the specific levels, that are expressed in certain individual animals or those produced under different conditions[9]. Proteome map of muscle mouse is available on line (www.expasy.ch) and has been employed as a reference in comparing meat from different species, in studying changes in meat quality associated with post mortem aging or those induced by the interaction of muscle proteins with lipids, carbohydrates, and other meat components[10]. In a recent study, proteome analysis was applied to the identification of molecular changes that occur in muscle tissue and meat of pork during post mortem storage of the carcass[11].

5.2.2 Proteomics in cereal science

Proteomic analysis of amphiphilic proteins of wheat kernels was performed in order to increase knowledge in their physiological and technological function. This proteomic approach provided very useful information about protein components linked to bread wheat quality and particularly to kernel hardness.[12,13] The proteome of wheat was used to predict the quality of bread produced from it. Multivariate analysis has also been applied to support proteome analysis in order to simplify data handling obtained by MALDI-TOF MS. Examination of gliadin data based on MS revealed that quality among wheat varieties could be determined by means of principal component analysis.

Rice (*Oryza sativa L.*) proteome has been so well studied for its implications in the human diet that it is becoming the reference organism to study plant proteomics[14].

5.2.3 Application in milk and cheese science

Milk proteins have been studied for over 50 years; proteomics and associated technologies facilitate further advances in the knowledge of milk proteins. Proteomics allows the detection,

identification and characterization of milk proteins, facilitates the analysis and the detailed characterization of milk proteins of large numbers of milk proteins simultaneously: high-abundance proteins and low-abundance proteins, milk fat globule membrane proteins, post-translational modification analysis such as phosphorylation and glycosylation, by means of two-dimensional gel electrophoresis in combination with various mass spectrometry techniques.[15] 2DE technique was used for studying the composition and variation of proteins and peptides in bovine milk with various somatic cell counts and in tank milk exposed to various kinds of heat treatment and storage conditions.[16] Cow's milk proteins reactive to IgE in children with documented IgE-mediated cow milk allergy (CMA) were separated and identified by 2-DE and peptide mass spectrometry.[17] Proteolysis too started to be studied: a rapid screening of peptides issues of proteolysis coming from different cheese varieties was developed using (MALDI-ToF-MS)[18]; by means of SDS-page and immunoblotting, high specific anti-αS(αS1+αS2)-casein monoclonal antibody and amino acid sequence determination, it was possible to identify three main αS-casein-derived polypeptides in cheese and their production was evaluated in function of the three enzymes most involved in cheese proteolysis (pepsin, chymosin, and plasmin)[19]. A fingerprint of the proteins in cheese has been made extracting bacterial proteins from the cheese matrix, separating them using proteomic tools and thus creating a reference map. This work identified five functional groups of proteins involved in proteolysis, glycolysis, stress response, DNA and RNA repair and oxido-reduction, suggesting physiological changes such as stress and nutriment starvation of the different strains at the end of cheese ripening, which module the dynamics of populations thermophilic starters/ propionibacteria (following for example, enzymes of dying bacteria). [20]

5.2.4 Application in food technology, processing and nutritional quality

Proteomics has been used in monitoring changes in specific food protein components and/or the occurrence of interaction upon processing. Moreover, it led to the identification of markers for either specific food processing technologies or for quality of processed food. For example, in an interesting study, 2-DE provided a convenient way to identify the different proteins that are produced in tomato fruit under heat stress[21,22]. Proteomic contributes to safety evaluation of food ingredients, in the evaluation of absorption, body distribution and metabolism of food ingredients, in the detection of food spoilage or pathogenic microorganisms (identificating biomarkers representative for specific food spoilage and/or pathogenic microorganisms[23]: these biomarkers can be used for specific detection methods for these micro-organisms[24]); proteomic contributes to quality and authenticity of foods (as the proteome of certain food (wheat, wine, fish) can be used to authenticate food origin or food quality). Moreover, proteomic can contribute to production of food ingredients by controlling the yield of bioprocesses (production of amino acids, carbohydrates, etc.) through metabolome/proteome of microorganisms used for such production or the food processing, as in complex fermentations (beer, cheese, sausage,

etc.), containing different microorganisms and complex substrates, the quality of the proteome or metabolome of the starter culture can be used to predict the quality of the fermented end-product. Therefore, enormous cost savings can be achieved if 'bad' starter culture batches can be eliminated.[25]

It is expected that the application of the integrated nutrigenomics (transcriptomics, proteomics, metabolomic) approach in nutritional sciences, will allow for accelerated implementation of mechanistic knowledge in food design, in food technology it will reduce research and development times, thereby reducing costs and shortening time-to-market.

5.2.5 Proteomics in food allergy prevention

The first and simplest applications of proteomic approaches in allergy and asthma research consist in the fast identification of new allergenic molecules, that can be easily achieved by protein extraction from source, separation on 1D or 2D gels and visualized by IgE immunoblotting. [26] The greatest potential of proteomic technology derives from the ability to determine posttranslational modifications, which determine antigenic activities, which are not detectable by cDNA cloning and difficult to analyze by classical chemical methods. Moreover, the rapid progress of proteomic methods contribute to the elucidation of micro-heterogeneity of allergenic molecules deriving from post-translational modifications and from the presence of allergen isoforms in natural extracts[27]. Proteomics contribute to the identification of digestion-resistant peptides (epitope eliciting the allergic response).

A systematic proteomic analysis of rice leaf, root, and seed tissue using 2-DE followed by MS/MS and multidimensional protein identification technology allowed the detection and identification of more than 2500 proteins, which represents the most comprehensive proteome exploration to date[28]. This allowed identification of several previously characterized allergenic proteins in the seed sample, indicating the potential of proteomic approaches to survey food samples with regard to the occurrence of allergens.

5.3 Peptidomics

Peptide research has experienced considerable development during the past few decades. Peptides are a paramount example of how nature diversifies from one single gene to release multiple, regulated functionalities at the desired sites and time. To achieve this, peptides are sequentially generated by a complex network of more than 500 proteases, acting at intracellular sites, upon secretion, in extracellular environments, and, finally, serving (regulated) degradation. This cycle of maturation, activation, and degradation points out that the **peptidome** is mechanistically linked to the proteome: the distribution between both is regulated by proteases and counter-regulated by protease inhibitors.

The term peptidomics is relatively new and was first mentioned in the literature in 2001[29,30,31]. Logically, peptidomics is the specification of the complement of peptides of a cell (organelle,

tissue or organism). Peptidomics is the missing link between proteome and metabolome[32]. Peptidomics has thus been put among other established "omics" technologies, namely transcriptomics, proteomics and metabolomics, as described in Figure 5.4.

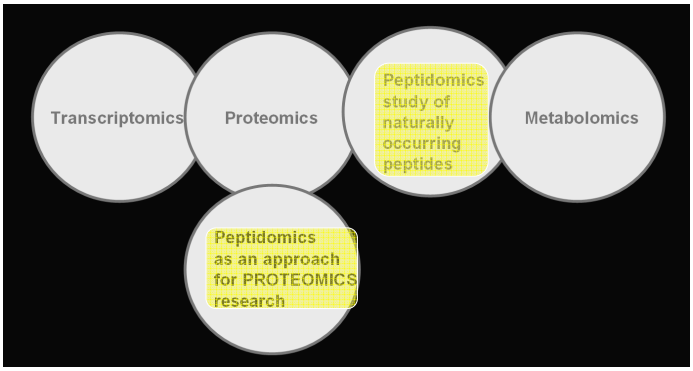


Figure 5.4. Peptidomic role in respect to the others "omics" techniques (from Mikhail Soloviev presentation at the first international workshop "Peptidomics: methods and applications" held on 6-7th September 2005 at Royal Holloway University of London).

This concept aims at the comprehensive visualization and analysis of small polypeptides, thus covering the mass range between proteomics and metabolomics. As in proteomics, the core technology in peptidomics is mass spectrometry, supported by bioinformatics. They are ideally suited for sensitive and comprehensive peptide analysis, especially in combination with the massive information content of today's genomic and transcriptomic databases. Given the high diversity of native peptides in living systems, clinical chemistry and modern medicine are the prime application areas.

The importance of small polypeptides in biological processes and the technical challenge come by working on complete or selectively depleted peptide mixtures ("affinity peptidomics" or "combinatorial peptidomics" respectively). The development of a technology which covers complex mixtures of peptides with low molecular weight and small proteins (0.5 to 15 kDa) is necessary, since peptides, amongst them families of hormones, cytokines and growth factors, play a central role in many biological and regulatory processes. The techniques involved should combine peptide sequence identification with peptide profiling. Standard proteomic approaches are not suitable for a systematic peptide analysis, since they do not cover the low molecular mass window. First of all, separation of peptide mixtures by using 2-D-gels is not performed in a satisfactory fashion as for proteomics and even 2D gel electrophoresis only shows a part of the proteome.

Second, staining short peptides gives much lower intensities than staining proteins since staining depends on the amount of peptide or molar concentration of peptide bonds and is therefore insufficient for detection. Third, MALDI mass mapping is not the method of choice for the process of identification. Short chain peptides contain a limited number of cleavage sites for proteolytic digestion. When (tryptic) digestion is performed, internal fragments are less

frequent, and peptides derived from the N- or C-terminal or non-digested peptides are present at a higher percentage. This is why MS methods for peptidomics are coupled with (multidimensional) liquid chromatography[33],[34], with the great advantages of automation, replicability and high speed of analysis. In the same direction, Top down and shotgun proteomics are gaining place in overcoming gel separation and there is a growing interest in global "gel free" proteomics[35].

5.4 References

[1] WasingerVC Cordwell SJ Cerpa-Poliak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphery-Smith I. Electrophoresis 16 (1995) 1090-1094

[2] Kvasnička F. Review Proteomics: general strategies and application to nutritionally relevant proteins *Journal of Chromatography B*, **2003,** 787: 77–89

[3] Hu L., Ye M., Jiang X., Feng S. and Zou H. Advances in hyphenated analytical techniques for shotgun proteome and peptidome analysis—A review, *Analytica Chimica Acta* **2007** 598, 2, 193-204

[4] Fournier M. L., Gilmore J. M., Martin-Brown S. A., and Washburn M. P. Multidimensional Separations-Based Shotgun Proteomics, *Chem. Rev*. **2007,** 107, 3654-3686

[5] Matthiesen R. Methods, algorithms and tools in computational proteomics: a practical point of view. *Proteomics.* **2007** 16:2815-32.

6 Robinson CV, Grob M, Eyles SJ, Ewbank JJ, Mayhew M, Hartl FU, Dobson CM e Radford SE Conformation of GroEL-bound alpha -lactalbumin probed by mass spectrometry. *Nature* **1994**, 372, 646-651

7 Caprioli RM., Farmer TB., Gile J.. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* **1997**, 69, 4751-60.

[8] Nadolski M. J. and Linder M. E. Protein lipidation *FEBS Journal* **2005** 274, 20, 5202–5210.

[9] Hollung K., Veiseth E., Jia X., Mosleth Færgestad E. and Hildrum K. I., Application of proteomics to understand the molecular mechanisms behind meat quality, *Meat Science* **2007**, 77, 97-104,

[10] Carbonaro M. Proteomics: present and future in food quality evaluation *Trends in Food Science & Technology* **2004**, 15: 209–216

[11] Lametsch, R., and Bendixen. E. Proteome analysis applied to meat science: Characterizing postmortem changes in porcine muscle. *J. Agric. Food Chem.* **2001.** 49:4531–4537.

[12] Miour N, Merlino M, Leroy P, Branlard G, Proteomic analysis of amphiphilic proteins of hexaploid wheat kernels. *Proteomics* **2002** 2:632-641.

[13] Bahrman N, Le Gouis J, Negroni L, Amilhat L, Leroy P, Lainé A-L, Jaminon O Differential protein expression assessed by two-dimensional gel eletrophoresis for two wheat varieties grown at four notrogen levels. *Proteomics* **2004** 4:709-719.

[14] Agrawal GK, Rakwal R. Rice proteomics: a cornerstone for cereal food crop proteomes. *Mass Spectrom Rev.* **2006** 1:1-53.

[15] O'Donnell R., Holland J.W., Deeth H.C., Alewood P., Milk proteomics, *Internat Dairy Journal* **2004** 14:1013–1023

[16] Lindmark-Mansson H., Timgren A., Alden G., Paulsson M., Two-dimensional gel electrophoresis of proteins and peptides in bovine milk *Internat Dairy Journal* **2005** 15 111–121,

[17] Natale M., Bisson C., Monti G., Peltran A., Perono Garoffo L., Valentini S., Fabris C., Bertino E., Coscia A. and Conti ACow's milk allergens identification by twodimensional immunoblotting and mass spectrometry .*Mol. Nutr. Food Res.* **2004**, 48, 363 – 369

[18] Piraino P., Upadhyay V. K., Rossano R., Riccio P., Parente E., Kelly A. L., McSweeney P. L.H., Use of mass spectrometry to characterize proteolysis in cheese *Food Chemistry* **2007**, 101: 964–972

[19] Gaiaschi A., Beretta B., Poiesi C., Conti A., Giuffrida M. G., Galli C. L., and Restani P. Proteolysis of αs-Casein as a Marker of Grana Padano Cheese Ripening *J Dairy Sci* **2000** 83:2733–2739

[20] Gagnaire V., Piot M., Camier B., Vissers J. P.C., Jan G., Léonil J. Survey of bacterial proteins released in cheese: a proteomic approach International *Journal of Food Microbiology* **2004** 94 185– 201

[21] Polenta G. A., Calvete J. J. , González C, B. Isolation and characterization of the main small heat shock proteins induced in tomato pericarp by thermal treatment *FEBS Journal* **2007** 274,24,6447-6455

[22] Proteomics of Microorganisms -/- Fundamental Aspects and Application *Advances in Biochemical Engineering/Biotechnology* Volume 83/**2003**

[23] Hefford, M A.; D'Aoust, S; Cyr, T D.; Austin, J W.; Sanders, G; Kheradpir, E; Kalmokoff, M L.Proteomic and microscopic analysis of biofilms formed by Listeria monocytogenes 568, *Canadian Journal of Microbiology*, **2005** 51,3 pp. 197-208

[24] Wilkins C. L., Lay J. O., Winefordner J. D. Identification of Microorganisms by Mass Spectrometry, **2006,** ed. WILEY

[25] Kvasnička F., Proteomics: general strategies and application to nutritionally relevant proteins *Journal of Chromatography B*, **2003,** 78777–89

[26] Crameri R. The potential of proteomics and peptidomics for allergy and asthma research, *Allergy* **2005**: 60: 1227–1237

[27] Blume C, Lindner B, Pecker WM, Petersen A. Microheterogeneity of the major grass group 6 allergen Phl p 6: analysis by mass spectrometry. *Proteomics* **2004**;4:1366–1371.

[28] Koller et al., 2002

[29] Verhaert P, Uttenweiler-Joseph S, de Vries M, Loboda A, Ens W, Standing KG. Matrix-assisted laser desorption/ionization quadrupole time-of-flight mass spectrometry: an elegant tool for peptidomics. *Proteomics.* **2001** Jan;1(1):118-31.

[30] E. Clynen, G. Baggerman, D. Veelaert, A. Cerstiaens, D. Van der Horst, L. Harthoorn, R. Derua, E. Waelkens, A. De Loof and L. Schoofs Peptidomics of the pars intercerebralis±corpus cardiacum complex of the migratory locust, Locusta migratoria *Eur. J. Biochem.* **2001** 268: 1929±1939

[31] Schulz-Knappe, P. et al. Peptidomics: the comprehensive analysis of peptides in complex biological mixtures. *Comb. Chem. High Throughput Screen.* **2001** 4: 207–217

[32] Soloviev M., Finch P., Peptidomics: Bridging the gap between proteome and metabolome *Proteomics* **2006**, 6, 3, 744 - 747

[33] I. Neverova, J. E. Van Eyk, Role of chromatographic techniques in proteomic analysis *Journal of Chromatography B*, **2005** 815 51–63

[34] Issaq H. J., Chan K. C., Janini G. M., Conrads T. P., Veenstra T. D. Multidimensional separation of peptides for effective proteomic analysis *Journal of Chromatography B*, **2005,** 817, 35–47

[35] Lambert J.-P., Ethier M., Smith J. C., and Figeys D., Proteomics: from Gel Based to Gel Free *Anal. Chem.* **2005**, 77, 3771-3788

Results, part 1:

PROTEINS and FOOD SAFETY: FOOD ALLERGENS

1 HIDDEN ALLERGENS IN BABY FOODS

1.1 Introduction

Cow's milk allergy (CMA) is an immunological response to milk proteins causing adverse clinical reactions. CMA is one of the most common forms of childhood food allergy, affecting 2–3% of infants and young children[1]. The major allergens of milk proteins are caseins, (whole casein fraction, Bos d 8) which are also the major protein constituents of milk, α-lactalbumin (Bos d 4) and ß-lactoglobulin (Bos d 5), the main proteinaceous component of whey. CMA is mainly an IgE-mediated hypersensitivity reaction, although other mechanisms and delayed manifestations have also been identified. CMA develops in early infancy and is considered to be transient in most cases.[2] Its frequency is linked to the high cow's milk consumption during infancy and early childhood which makes it the most important source of food allergens during these years.[3] Newly born children are exposed to bovine milk protein when they need a substitute or a supplementation for breast-feeding: in that case, cow's milk-based formulas are used.[4]

Children suffering from CMA need to follow a CM-free diet. In that way, hypoallergenic formulas have been developed, since the 1940s. These products, initially based on extensively hydrolyzed bovine casein, were tolerated by the vast majority of infants allergic to cow's milk. However, over the years, case histories of infants reacting to these extensively hydrolyzed formulas (eHF) have been reported (residual allergenicity).[5] The products based on hydrolyzed casein have a further major drawback that acceptance is difficult because of the unpleasant bitter taste. Alternatives initially were whey hydrolisates, which were less bitter but, in fact not avoidable due to their origin, whey can carry over caseins. Nowadays, hypoallergenic formulas have been developed based on soy and rice, and they were found to be nutritionally adequate for infant feeding[6], totally free from caseins, but having a protein efficiency ratio equal to caseins, and to meet patient taste preferences. In a recent study, a rice based hypoallergenic formula, was validated as a possible alternative for children with cow's milk allergy[7]. The protein source is an extensively hydrolyzed rice protein isolate, supplemented with L-lysine and L-threonine in order to bring amino acid composition closer to that of human milk.

However, very often hypoallergenic infant formulas are produced on the same lines as other milk-based products. These lines obviously undergo intensive "Cleaning in Place" treatments (based on surfactants, acid and basic detergents, disinfectants and steam). Moreover, the first product going through the lines is discarded. All these efforts cannot totally prevent cross contamination from milk proteins, so that a constant analytical control is needed. There are only a few validated methods for the detection and quantification of minute amounts of allergens in foods. Immunological methods can involve either human IgE or animal antisera. Dot-immunoblotting and SDS-PAGE/immunoblotting are sufficient for qualitative detection of food

allergens, while Rocket- immunoelectrophoresis and Enzyme-linked immunosorbent assays (ELISA) can be used for quantifying hidden food allergens[8,9]. Also DNA detection methods have been used to detect trace amounts of DNA related to hidden allergens.[10]

Currently, ELISA methodology is the method of choice for rapid and sensitive protein-based food analysis for routine laboratories in both industry and food safety control institutions. Several ELISA test kits for allergen detection and quantification are currently commercially available, although they are not always standardized and do not give always the same response [11] and suffer from false positives[12]. The reason for their widespread utilization is in their sensitivity, specificity, limit of detection, recovery and reproducibility and easy use.

It is important for researchers as well as for the food industry to find molecular markers for allergens and to develop methods of detection also in trace amounts. Proteomic techniques are especially useful for the identification of IgE-binding proteins for several reasons. First, conditions for solubilization of proteins are very aggressive, involving combinations of urea, thiourea, non-ionic detergents and strong reducing agents. These reagents increase the possibility of detecting allergens. Second, spots in two-dimensional electrophoresis provide substantial information: isoelectric point (pI), molecular weight and heterogeneity (charge and/or size variants) of the IgE-binding component. This information makes it possible to substantially narrow the search for protein identity. Mass spectrometry may finally identify or confirm the identification of the allergen.

1.2 Aim of the work

In the present work, the possibility of using the bottom-up proteomic approach to identify hidden caseins in infant hypoallergenic formulas has been investigated. In particular, we have analyzed 16 samples of infant formulas furnished by a well-known baby food industry in which contamination had been detected and quantified by ELISA methods ("Casein assay kit" Tepnel BioSystems). The aim of this work was to develop confirmatory methods based on 1D and 2D-electrophoresis in order to analyze the full protein content of the products and to identify milk-derived proteins. Putative milk-derived spots have been further investigated by serum-based immunoblotting and MALDI-TOF spectrometry

1.3 Experimental part

1.3.1 Samples

The 16 samples of infant formulas are summarized in Table 1.1. Blank samples were taken from a local market.

Table 1.1. Samples used in the present work.

PRODUCT	CODE	INGREDIENTS	PROTEIN CONC.§ (µg/ml)	CASEIN CONTAMI-NATION* (ppm)
Four cereals	FC1	Wheat, rice, barley, oat flours; mineral salts, vitamin, flavorings	346,85	22,54
Four cereals	FC2		471,66	5,08
Four cereals	FC3		623,08	3,87
Four cereals	FC4		247,23	0
Wheat semolina	WS1	Durum wheat semolina, mineral salts, vitamin, flavorings	712,88	0
Wheat semolina	WS2		804,39	>25
Wheat semolina	WS3		1113,6	2,30
Rice based product	RB1	Rice flour, mineral salts, vitamin flavorings	321,98	0
Rice based product	RB2		101,36	0
Rice protein based product, highly hydrolyzed - 2	RBHH1	Maltodextrins, glucose syrup, hydrolyzed rice protein, vegetal oils, maize starch, mineral salts, emulsifiers, L-lysine, L-threonine, L-tryptophane, taurine, L-carnitine	706,69	>30
Rice protein based product, highly hydrolyzed - 2	RBHH2		429,47	5,06
Rice protein based product, highly hydrolyzed - 2	RBHH3		589,54	3
Rice protein based product, highly hydrolyzed - 2	RBHH4		623,5	0
Rice protein based product, highly hydrolyzed - 1	RBHH5	maltodextrins,glucose syrup, hydrolyzed rice protein, vegetal oils, maize starch, mineral salts, emulsifiers, L-lysine, L-threonine, L-tryptophane,taurine, L-carnitine,nucleotides	567,8	0
Rice protein based product, highly hydrolyzed - 1	RBHH6		593,42	1,6
Rice, maize, tapioca based product	RMT1	Rice, maize, tapioca flours; mineral salts, vitamin flavorings	305,47	3,90

* determined by ELISA("Casein assay kit" Tepnel BioSystems); § determined by Bradford method, at λ 595 nm

1.3.2 Protein extraction

1.3.2.1 Reagents

- Acetone (CH_3COCH_3)
- Milli Q H_2O obtained with Millipore Alpha Q system
- Sodium Carbonate Na_2CO_3
- Sodium Bicarbonate $NaHCO_3$
- Trichloroacetic acid TCA 10% (CCl_3COOH) in acetone (CH_3COCH_3)
- Tris-HCL 50 mM pH 8.8
- Methanol CH_3OH (VWR)
- Formic acid HCOOH solution (0.1% HCOOH in H2O), (ACROS)
- PBS (Phosphate-Buffered-Saline), preparation of 1 L solution
  - 8 g NaCl (125 mM)
  - 0.2 g KCl (2.5 mM)

- 1.44 g Na_2HPO_4 (8 mM)
- 0.24 g KH_2PO_4 (1.5 mM)
- 800 ml of distilled H2O
- Adjust the pH to 7.4 with HCl
- Add Milli Q H_2O to 1 liter.

1.3.2.2 Instrumentation

- Centrifuge ALC PK110
- Homogenizer Ultraturrax T50 basic (IKA-WERKE)
- Sonicator ULTRA Sonik™ 57X NEY
- Centrifuge HETTICH UNIVERSAL 320 R
- Amicon® Ultra-4 Centrifugal Filter Units Amicon Ultra 10K device – 10,000 NMWL (Nominal Molecular Weight Limit)

1.3.2.3 Procedures

i CASEIN ASSAY KIT extraction method

2 g of sample are mixed with 30 ml of extraction buffer (aqueous buffer of (Na_2CO_3) 0,05 M and ($NaHCO_3$) 0,05 M, to a final pH of 9.6), the mixture is homogenized for 2 min and centrifuged at 2800 g for 10 min; the supernatant is kept and stored at -20°C.

ii TCA/Acetone extraction method

10 gr of sample are mixed to 100 ml Tris-HCl 50 mM buffer at pH 8.8. the mixture is homogenized for 2 min, centrifuged at 2800g for 10 min; the supernatant is collected, 3 volumes of TCA 10%/acetone at 4° are added and the protein precipitation is obtained at -20° overnight, in Eppendorf vials. To complete the precipitation, a 10 min centrifugation at 4000g is done, the precipitate is recovered after discarding the supernatant, the pellet is dissolved in TCA 10%/acetone at 4°, sonicated 1 minute, centrifuged at 4000g for 10 min, and this procedure is repeated one time with TCA 10%/acetone at 4°, and a second time with acetone at 4°. From the pellet, proteins are extracted with PBS (paragraph iii).

iii PBS extraction method

0.5 to 5 g of sample (based on product) are mixed with 15 ml of PBS buffer; after 1 h of gentle vortexing in cool (ice bath), the mixture is centrifuged at 5000g, 1 h, 4°C; the supernatant is recovered and dried under nitrogen flux.

iv Desalting

Prerinsing: the ultra-filtration membranes in Amicon Ultra devices contain trace amounts of triethylene glycol (TEG). The device must be pre-rinsed with 5 washes of a CH_3OH:H_2O 1:1 solution.

Preserving: the devices are stocked in a CH_3OH:H_2O 5:95solution, at 4 °C until use.

Use: the ultra-filtration device is loaded with 4 ml of the solution to be recovered and centrifuged. The sample is loaded and centrifuged and then 3 washes with a 0.1% HCOOH water solution are done. The filtered solution is discarded each time and the non filtered suspension (upper part of the device) is recovered after washes using an Eppendorf pipette P100. A major recovery is obtained adding 250+250 µl of the HCOOH solution, the membranes are washed with this solution using the pipette P100 and the recoveries are added to the suspension in an Eppendorf.

The membranes used in Amicon Ultra devices are characterized by a nominal molecular weight limit (NMWL), that is, their ability to retain molecules above a specified molecular weight. Solutes with molecular weights close to the NMWL may be only partially retained. Membrane retention depends on the solute molecular size and shape. The device is loaded with the liquid to be filtered and the filtration is achieved by centrifugation at 2800g.

1.3.3 Protein quantification method

1.3.3.1 Reagents

- Coomassie G
- methanol (CH_3OH)
- H_3PO_4 85%
- Bovine serum albumin (BSA)
- Dye stock of Coomassie (5x): Stir very well 100 mg Coomassie G in 50 ml of CH_3OH + 100 ml H_3PO_4 85% + 200 ml Milli Q H_2O; final concentration of Coomassie G: 0,5 mg/ml; 42,5% H_3PO_4. stock the solution at 4°C. At the moment of use, dilute dye stock 1:5 in Milli Q H_2O (assay reagent).

1.3.3.2 Instrumentation

- Spectrophotometer PERKIN ELMER UV/VIS Lambda Bio 20
- UV cuvettes (single use) 4 ml, SIGMA-ALDRICH

1.3.3.3 Procedure

First, a calibration curve was obtained with standard solutions of BSA (0, 250, 500, 1000, 1500 e 2000 µg/ml); 2 ml of assay reagent and 40 µl of protein standard solution are mixed in the cuvette; the measure at the UV spectrophotometer was done immediately after mixing. For sample measurement, 2 ml of assay reagent and 40 µl of protein extract (eventually diluted) were mixed together.

1.3.4 Proteins analysis by electrophoresis

1.3.4.1 Reagents

- 20 XT Reduction buffer (BIORAD)

- 2-Iodoacetamide (CH_2CONH_2) (FLUKA)
- 50% glycerol(BIORAD)
- acetic acid CH_3COOH, (solution 5% in Milli Q H_2O)
- agarose (BIORAD)
- Bromophenol blue (BIORAD)
- Casein mix (SIGMA-ALDRICH)
- Carrier ampholytes (BIORAD)
- CH_3CN, HPLC grade (VWR)
- CH_3OH, HPLC grade (VWR)
- CHAPS (BIORAD)
- Coomassie Brilliant Blue R-250 (BIORAD)
- Criterion Precast Gel 16,5% Tris-Tricine/peptide, 14 wells (BIORAD)
- Criterion XT Precast Gel, 12% Bis-Tris, IPG+1 comb (BIORAD)
- Criterion XT Precast gel, 4-12% Bis-Tris, 18 wells (BIORAD)
- Dithiothreitol DTT ($HSCH_2(CHOH)_2CH_2SH$) (FLUKA)
- Formic acid 0.1% (HCOOH), 99%, (ACROS)
- Methanol (CH_3OH), HPLC grade, (VWR)
- Milli Q H_2O obtained with Millipore Alpha Q system
- Mineral oil (BIORAD)
- NaOH, (0.1N solution)
- Polypeptide SDS-PAGE Molecular Weight Standards (BIORAD)
- Ready Strip IPG strips, 11 cm, pH: 3-10 (BIORAD)
- SDS, in powder and in solution 20% (SIGMA-ALDRICH)
- SDS-PAGE Molecular Weight Standards, Broad Range (BIORAD)
- Silver nitrate ($AgNO_3$)
- TFA trifluoroacetic acid (ACROS)
- Tricine sample buffer 4x (BIORAD)
- Tris/HCl (SIGMA-ALDRICH) (100mM, 50mM, 25mM, 20mM solutions in Milli Q H_2O)
- Tris/Tricine/SDS Running buffer (BIORAD)
- Tween 20 (SIGMA-ALDRICH)
- Urea (NH_2CONH_2) (SIGMA-ALDRICH)
- XT MES Running buffer: XT MES 1X (BIORAD)
- XT sample buffer (BIORAD)

1.3.4.2 Instrumentation

- Centrifuge 1-13 (SIGMA-ALDRICH)
- Densitometer GS-800 (BIORAD)

- Glass Syringe 50 µl (HAMILTON)
- Midi format vertical electrophoresis: Criterion Cell (BIORAD)
- Midi format vertical electrophoresis: Criterion Cell (BIORAD)
- PDquest software (BIORAD)
- Pipettes P5000, P1000, P100, P20, P10 (EPPENDORF)
- Power supply: Power Pac Universal (BIORAD)
- PROTEAN IEF Cell (BIORAD)
- Software PD-Quest (BIORAD)
- Thermocycler PCR Sprint (THERMOHYBAID)

1.3.4.3 Procedures for 1D Gel electrophoresis of peptides and proteins with MW < 50kD

i Polypeptide standard preparation

- 2,5 µl Polypeptide SDS-PAGE Molecular Weight Standards
- 47.5 µl of Tricine sample buffer (prepared with 25 µl Tricine sample buffer 4x, 5 µl XT reduction buffer 20x and Milli Q H_20 to final 100 µl volume)
- 5 min incubation at 95 °C

ii Sample preparation

- 25 µl Tricine sample buffer
- 2,5 µl XT Reduction buffer 20x
- 10 µg of sample protein
- Milli Q water to final 50 µl volume
- 5 min incubation at 95 °C

iii Electrophoretic run

For each wells of the Criterion Precast Gel 16,5% Tris-Tricine/peptide, 20µl of sample prepared as described or 2.5 µl of Polypeptide standard are loaded. The running buffer used is Tris/Tricine/SDS Running buffer, the voltage applied to the Criterion Cell is 120 V. The cell is placed in a water bath to avoid gel heating. The run lasts 90 min environ. 1% of Coomassie blue in the sample buffer runs as indicator.

1.3.4.4 Procedures for 1D Gel electrophoresis of peptides and proteins with MW < 150kD

i Protein standard preparation

- 2,5 µl SDS-PAGE Molecular Weight Standards, Broad Range
- 47.5 µl of XT buffer (prepared with 25 µl XT buffer 4x, 5 µl XT reduction buffer 20x and Milli Q H_20 to final 100 µl volume)
- 5 min incubation at 95 °C

ii Sample preparation

- 25 µl XT sample buffer

- 2,5 µl XT Reduction buffer 20x
- 10 µg of proteins, following protein sample quantification described in paragraph 1.3.3
- Milli Q water to final 50 µl volume
- 5 min incubation at 95 °C

iii Electrophoretic run

For each wells of the Criterion XT Precast gel (4-12% Bis-Tris), 20µl of sample prepared are loaded as described or 2.5 µl of Protein standard. The running buffer used is XT MES Running buffer, the voltage applied to the Criterion Cell is 150 V. The cell is placed in a water bath to avoid gel heating. The run lasts 60 min. In the sample buffer there is a 1% of Bromophenol blue, which runs as indicator.

1.3.4.5 Procedures for 2D- electrophoresis of peptides and proteins with MW < 150kD

i Isoelectrofocusing

In this case samples must be desalted, as described in paragraph iv, and dried under N_2. 500 µg of proteins are redissolved in 200 µl of rehydration buffer, whose composition is urea 8M (24 g), DTT 50mM (385 mg), 4% CHAPS (2g), 0.2% Carrier ampholytes (250 µl of the stock 40% w/v), 0.0002% Bromophenol blue (100 µl of 0.1%stock), Milli Q H_2O to 50ml. 180 µl of redissolved samples are put in the IEF tray, and the strip placed with the active part in contact with the sample. +/- strip direction follow +/- sense of the tray. Strips are 11 cm long, and a pH linear gradient from 3 to 10. Active rehydration of the strip is performed overnight applying low voltage (50 V) to the IEF tray. After 1 hour of rehydration, mineral oil is added to cover completely the strips. The PROTEAN IEF CELL is programmed for 8 hours of active rehydration and then to start the focusing run (Table 1.2).

Table 1.2. Focusing run condition

STEP	VOLT	TIME (hours)	VOLT/HOURS	RAMP
1	250	3	-	Slow
2	8000	2.5	-	Linear
3	8000		25000	Rapid

At the end of the focusing, strips are taken and exceeding mineral oil is discarder on filter paper.

ii Strip equilibration

Strips are placed in a tray containing DTT Equilibration buffer for 10 min and then in a second tray containing IAA Equilibration buffer for 10 min. Each tray has lines for each strips and each line can contain 2.5 ml of equilibration buffer. Equilibration base buffer is composed of 6M Urea (36g), 2%SDS (10ml of a 20%solution in water), 0.05M Tris/HCl buffer pH 8.8 (3.3ml of a 1.5M stock solution, pH 8.8), 20% glycerol (40 ml of a 50%stock solution), Milli Q water to 100 ml. 2%

DTT Equilibration buffer is prepared immediately prior to use adding 200 mg of DTT every 10 ml of Equilibration base buffer. 2.5% IAA Equilibration buffer is prepared immediately prior to use by adding 250mg of IAA every 10 ml of Equilibration base buffer.

iii Second dimension: gel electrophoresis

Equilibrated strips are applied to the top of the Precast gel Criterion XT (12% Bis-Tris, IPG+1 well). The strip is aligned so that the plastic back of the strip is against the back plate and the IPG strip is touching the top of the gel. "+" part of the strip is towards the well for the MW standards. MW standards are placed in the well (as described in paragraph 1.3.4.4). Molten agarose is added to fix the strip and when it is set, the gel can be put in the electrophoresis cell for the run, covered with the buffer.

1.3.4.6 GEL STAINING

Several staining methods are available. We chose two reversible and sensitive methods. Gels are removed from the plastic case and immediately placed in the stain.

i Colloidal Coomassie

It is prepared as follow:

- 0.5 g Brilliant blue R-250 (0.1%)
- 170 ml CH_3OH (34%)
- 15 ml H_3PO_4 (3%)
- 85 g $(NH_4)_2SO_4$ (17%)
- Milli Q water to 500 ml.
- First, Brilliant blue is redissolved in phosphoric acid, then all the other constituents are added. 4 hours of stirring can be necessary before dissolution is completed.

The gel is placed in a glass-made vessel with the stain covering it. The gel is soaked in a solution of the dye for at least 1 hours. Any dye that is not bound to protein diffuses out of the gel during the destain steps, when the gels are rinsed with a solution of 10% acetic acid, 40% CH_3OH, 50% Milli Q water, changed at least twice, to achieve the desired contrast.

ii Reversible silver staining

Reagents used in silver stain are here reported.

- Fixer: 30% ethanol,10% acetic acid in H_2O
- Reducent-sensitizer: 0.02% sodium thiosulfate in water
- Silver reagent: 0.2% silver nitrate ($AgNO_3$) in H_2O
- Developer: 0.3 % sodium carbonate (Na_2CO_3) + 0.025% formaldehyde, added just prior to use, 0.06mM sodium thiosulphate (Na_2S_2O3)
- Stop solution: 0.005% Tris, 2.5% acetic acid

The gels are fixed in the Fixer solution overnight, then are washed in 20% ethanol in H_2O for 20 min and H_2O for 10 min. The gel is treated with reducent-sensitizer solution for 1 min, then is

washed in water three times for 20 sec each. The gels are treated with the silver reagent for 45 min under shaking, then the silver reagent is washed out of the dish and wash the gel is washed for 20 sec in a large amount of water. The gel is covered with the developer, under shaking until the gel has reached the desired degree of development (changing the developer if it turns to a turbid brown color). The reaction is stopped by rapidly removing the developer and adding the stop solution for at least 5 min. Finally, the gel is washed in water.

1.3.5 In gel digestion

The band or spot of interest is cut from the gel, destained, and then treated with a protease. The enzyme penetrates the gel matrix and digests the protein to peptides, which then are eluted from the gel by washing.

*1.3.5.1 Reagents**

- 2-Iodoacetamide (CH_2CONH_2) (FLUKA)
- Dithiothreitol DTT ($HSCH_2(CHOH)_2CH_2SH$) (FLUKA)
- Ammonium bicarbonate NH_4HCO_3
- Potassium ferricyanide $K_3Fe(CN)_6$
- Sodium thiosulphate $Na_2S_2O_3$
- CH_3CN (VWR)
- HCOOH (ACROS)

1.3.5.2 Spot excision and destaining

With a P1000 tip, cut on its edge to adapt it to the spot dimension, protein bands or spots are excised from a stained polyacrylamide gel and placed into a 0.5 ml tube. Coomassie stained spots are destained as follow: 200 µL of 25 mM NH_4HCO_3/50% acetonitrile are added (enough to immerse the gel particles) and the tubes are vortexed overnight (12 h) on a low setting. The solution is removed and the wash is repeated for additional time (2 hr mixing). Resulting gel particles are clear, so it is possible to dehydrate gel pieces by adding acetonitrile (100 µL). At this point the gel pieces shrink and become an opaque-white color. Tubes are centrifuged for 2 min at 4000g and CH_3CN is removed with P100 by air-dry for 5-10 min.

The silver staining used is reversible[13], with a destaining solution of 30 mM potassium ferricyanide $K_3Fe(CN)_6$ and 100mM sodium thiosulphate $Na_2S_2O_3$, both in water. Spots are incubated in this solution in gentle shaking until the spot is destained, then the gel is washed with Milli Q water for 5 min and with 100mM NH_4HCO_3 solution in water for 20 min.

* Reagents specific to this part of the work are here listed. Some common reagents are omitted as already reported in the previous part of this chapter.

1.3.5.3 Reduction and alkylation of cysteine residues

When maximum protein coverage is required or when digesting a band from a one dimensional gel (proteins separated by two dimensional gel electrophoresis should already be reduced and alkylated so these steps can be omitted) reduction and alkylation of cysteine residues is performed: 30 µl of a freshly prepared 10 mM DTT in 100mM NH_4HCO_3 solution are added to cover the gel pieces, and reduction take place for 1hour at 56°C. DTT solution is replaced with roughly the same volume of freshly prepared 55 mM IAA in 100mM NH_4HCO_3 (30 µl). Incubation lasts 45 min at 56 °C, in a dark place. The gel spot is washed with 100 µl of 100 mM NH_4HCO_3 for 10 min, dehydrated with 100 µl of CH_3CN (10min washing + evaporation), hydrated with 100 µl of 100 mM NH_4HCO_3, again dehydrated with 100 µl of CH_3CN and finally vacuum dried.

1.3.5.4 Enzymatic digestion

Gel particles are rehydrated in 40 µL trypsin solution (or volume necessary to cover the expanding gel pieces), 12 ng/µl in 50 mM NH_4HCO_3 and 5mM $CaCl_2$ to keep them immersed throughout digestion. The incubation lasts 12 to 16 hrs at 25ºC, in mild shaking. 5% aqueous TFA is added to halt the digestion. The surnatant is recovered. Peptides are definitively extracted with 100µl of 50mM NH_4HCO_3 , then 3 times with 100µl of an aqueous solution with 5% HCOOH and 50% CH_3CN. Surnatants collected are dried under N_2 flux.

1.3.6 MALDI-TOF analysis

1.3.6.1 Reagents

- MALDI Calibrants
- α-Cyano-4-hydroxycinnamic acid (CHCA), (SIGMA-ALDRICH)

1.3.6.2 Instrumentation

- MALDI-TOF LR (MICROMASS)
- Mass Lynx 4.0 software (Waters)

1.3.6.3 Procedures

i Sample preparation

Dried mixture of peptides is redissolved with 40µl of 0.1%-1% HCOOH solution in Milli Q water, ensuring that the final sample solution has a pH<4.

ii Zip-Tip desalting

They are necessary:

- Wetting solution (CH_3CN)
- Equilibration solution (0.1% TFA in Milli Q H_2O)
- Washing solution 1: Milli Q H_2O
- Washing solution 2: 0.1% TFA/Milli Q H_2O

- Eluting solution (50% ACN in 0.1% TFA)

The ZIP TIPs C18 are suitable for an Eppendorf pipette P10. Using the maximum volume setting of 10 µL, the necessary solution is taken into the tip and dispensed to waste. This procedure is repeated 3 times with the Wetting solution and 3 times with the Equilibration solution.

The binding of peptides to the ZipTip is achieved by fully depressing the pipette plunger to a dead stop, aspirating and dispensing the sample 7–10 times from and into the tube for maximum binding of complex mixtures. The binding capacity is 3 - 5 µg of peptides.

Desalting is achieved by using the maximum volume setting of 10 µL, aspirating into the tip and dispensing to waste 3 times the Washing solution 1 and 3 times the Washing solution 2. The elution is achieved aspirating 3 times 2 to 5 µl of the Eluting solution and recovering the eluted fraction into a new tube.

iii Sample deposition on the target plate

1 µl of the desalted sample is spotted on the target plate and, immediately after, 1 µl of matrix solution (α-cyano-4-hydroxycinnamic acid, CHCA, 10 mg/ml in 50%CH_3CN 50% TFA 0.1%) is spotted over it and briefly mixed-on-target by aspiration/dispensing into the P10 tip. After evaporation has occurred, the target is ready for analysis.

iv Calibration standards

Each analysis in a MALDI-TOF needs its own calibration curve. Typical peptide mixture is:

- 1 µl of Bradykinin fragment 1-7 (MW 756.3997 Da), 10 pmol/µl
- 1 µl of Human Angiotensin II (MW 1,045.5423 Da), 10 pmol/µl
- 2 µl of P14R (synthetic peptide) (MW 1,532.8582 Da), 10 pmol/µl
- 3 µl of Human ACTH fragment 18-39 (MW 2,464.1989 Da), 10 pmol/µl
- 4 µl of Insulin oxidized B chain (bovine) (MW 3,493.6513 Da), 10 pmol/µl
- 6 µl of Insulin (bovine) 5,729.6087 (monoisotopic), 10 pmol/µl

The mixture is vortexed, and 1 µl of the mixture is spotted onto the target together with 1 µl of the CHCA matrix.

To achieve the best accuracy, an internal standard (reference mass) is spotted onto the lock mass well of the target (well at the centre of each group of 4 of the 96 wells). Typical lock mass is 1 µl of Human Angiotensin II (MW 1,045.5423 Da), 10 pmol/µl mixed with 1 µl of CHCA matrix solution.

v Instrumental conditions of analysis (MALDI-TOF MS)

- Mode: Reflectron Positive
- Pulse Voltages: 2300-2500V
- Source Voltage: 15000V
- Reflectron Voltage: 2000V
- Laser energy variable
- TLF delay: 500ns

- Sample Period: 0,5ns

Data are acquired automatically (after calibration) using Mass Lynx 4.0 software and processed with ProteinLynx parameters file to search for the most intense and significant signals and to give .txt output compatible with Mascot database.

vi Peptide mass fingerprint on Mascot database

Peptides are searched as fingerprint of their original protein in MASCOT database: (http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=PMF).

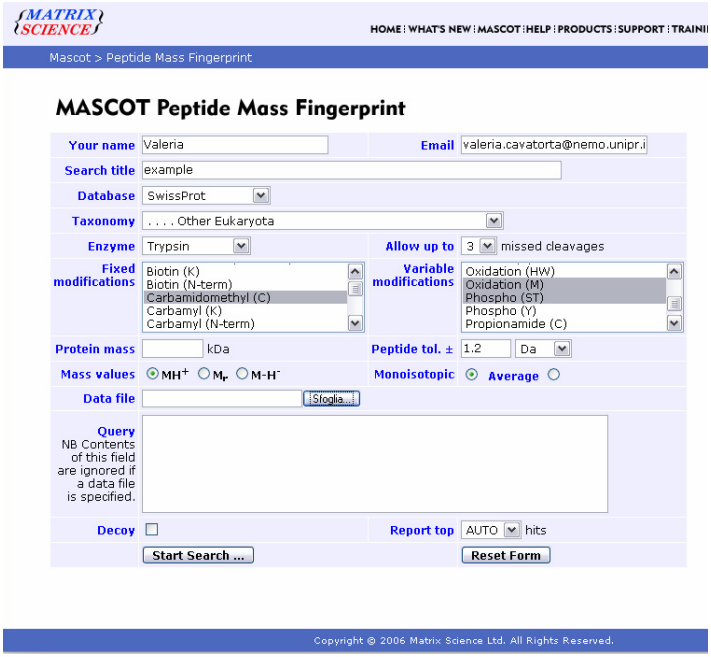


Figure 1.1. Mascot window for peptide mass fingerprint search.

The output is a ranking based on Protein score. Protein score is -10*Log(P), where P is the probability that the observed match is a random event. Protein scores greater than 67 are significant ($p<0.05$) and a sequence coverage is also available.

1.3.7 Western blotting

1.3.7.1 Reagents

- Glycine (BIORAD)
- SDS Sodium Dodecyl Sulphate, (SIGMA-ALDRICH)
- Ponceau solution: 0.1% Ponceau S in 5% acetic acid (SIGMA-ALDRICH)
- PVDF membrane (BIORAD)
- Fiber pad (BIORAD)
- Filter paper (BIORAD)
- Albumin, from hen eggs white (FLUKA)

- Anti-Human IgE from rabbit (BETHYL LABORATORIES);
- Goat anti rabbit-HRP (BIORAD)
- Opti-4-CN kit (BIORAD)
- Acetic acid (CH_3COOH), (CARLO ERBA)
- PBS buffer
- Tween 20 (SIGMA)
- PBS-T (PBS buffer + 0.1% Tween 20)
- Blocking buffer (3% albumin in PBS-T)
- 0.1 and 3 % ovalbumin solution in PBS-T
- REAGENT Opti-4-CN
- Opti-4-CN DILUENT

1.3.7.2 Instrumentation

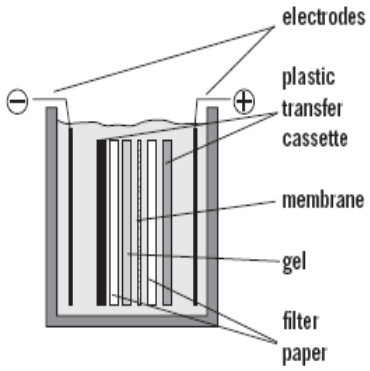
- Trans-blot electrophoretic transfer cell (BIORAD)
- Power supply: Power Pac Universal (BIORAD)
- GS-800 densitometer (BIORAD)

1.3.7.3 Procedures

After the SDS PAGE, the separated proteins are transferred on Polyvinylidene Difluoride PVDF membrane. The transfer is performed, according with the manufacturer's instructions, as follows.

Transfer buffer (25mM Tris, 192mM glycine, 5% CH_3OH, 0.01% SDS, pH 8.3) is prepared. The not-stained gel is placed into transfer buffer for 15 min, cutting fringes of the wells of the gel and the residual Bromophenol Blue run indicator. Cut pieces of the filter paper and of the PVDF membrane are wet for about 30 sec in CH_3OH on a rocker at room temp, then are wetted 10 min in Milli Q H_2O, 15 min in the transfer buffer and finally in the transfer buffer, together with fiber pad and filter papers. "Sandwich" for Bio-Rad's Transblot is assembled as follows:

Cathod (-) side of transblot
Fiber pad
Filter paper
Gel
PVDF membrane
Filter paper
Fiber pad
Anod (+) side of transblot



The closed sandwich is placed in the Transblot tank, which is filled with transfer buffer. Transfer occurs for 1 hr at 100 V in a refrigerated system (cooling coil). When it finished, the membrane is stain in Ponceau S, following manufacturer instructions, and destained in a 5 % acetic acid in Milli Q H_2O solution. The membrane is then dried and washed 2 times in PBS-T, for 5 min. The membrane is immersed in Blocking buffer and blocked for 1 hr with shaking. Then, it is washed in PBS-T. The membrane is then incubated with primary antibody (sera of allergic patients 1:50 in 0.1%ovalbumin in PBS-T) overnight at room temperature. In order to remove antibody solution the membrane is washed three times, for 5 min, in PBS-T. Then it is incubated with the secondary antibody (Rabbit anti-human IgE diluted 1:3000 in 0.1% ovalbumin in PBS-T) for 1 hour at room temperature. The antibody solution is removed washing the membrane three times for 5 min in PBS-T. Membrane is then incubated with the developing system (goat anti-rabbit antibodies, HRP-conjugated, diluted 1:3000 in 0.1% ovalbumin in PBS-T) for 1 hour at room temperature. The solution is removed and membrane is washed three times for 5 min in PBS-T. Protein detection is achieved By incubation of the membrane with the diluted Opti-4-CN reagent (prepared following manufacturer instructions) for 30 min till sufficient coloration is achieved. The membrane is washed with Milli Q Water for 15 min and the image is acquired at the densitometer.

1.4 Results and discussions

In order to evaluate the different protein extraction methods and the different kind of gels and stains to be utilized, a set of commercial samples, representative of different processing was analyzed. 1D and 2D-EF systems for casein trace analysis and detection were developed. Two detection methods were used: immunoblotting and Peptide Mass Fingerprint (PMF) by MALDI-TOF after image analysis of contaminated and casein-free product.

1.4.1 Protein Extraction methods

Three extraction methods were considered:

- Carbonate/bicarbonate buffer extraction, pH 9.6;
- TCA precipitation of proteins and extraction of interferent compounds with acetone;
- PBS buffer extraction.

Similar protein profiles were obtained but with different yields(Figure 1.2).

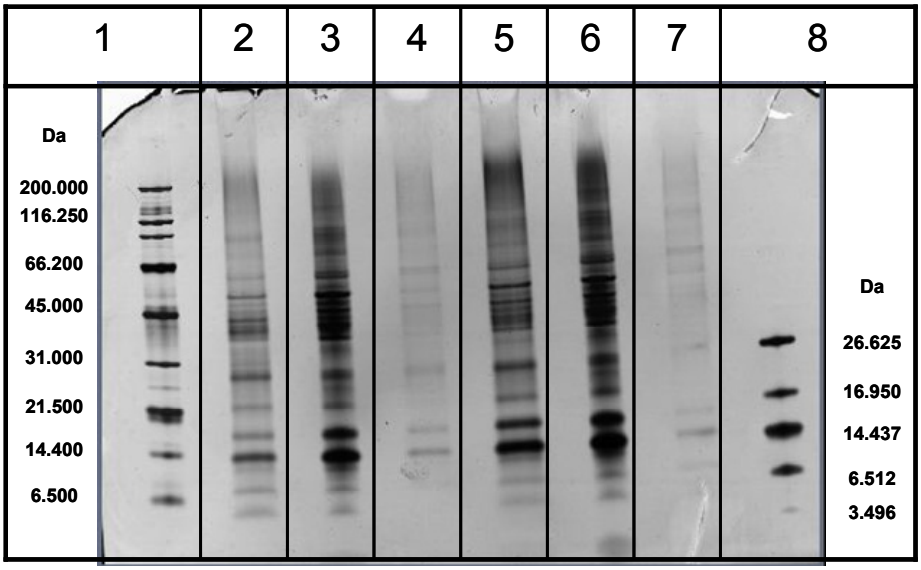


Figure 1.2. Criterion XT Precast gel 10% Bis-Tris, running voltage 120V. Line 1: HMW standards; lines 2 to 4, product WS2, wheat semolina; lines 5 and 7, product WS1, lines 2 and 5, Carbonate/bicarbonate buffer extraction; lines 3 and 6, TCA precipitation of proteins and extraction of interferences with acetone; lines 4 and 7, PBS buffer extraction; Line 8: LMW standards.

The figure clearly shows that TCA precipitation has the best yield, PBS extraction has the worst yield, and Carbonate/Bicarbonate is in between. Carbonate/Bicarbonate method has the great advantage of being simpler than TCA precipitation and, also being a rapid extraction method, was chosen for the analysis. In samples where high concentration of salts interfered with the analysis (2D-PAGE), the extract was desalted as described in paragraph in the experimental part.

1.4.2 Choice of gels and stains

Since caseins were the target of the analysis, a casein mix was used to contaminate a blank product in order to assess their detectability in the samples. . Samples were analyzed on different gels: Criterion Precast Gel 16,5% Tris-Tricine/peptide was the best gel for separating the proteins in the formula samples, as can be seen in Figure 1.3.
Moreover, samples were analyzed with different staining methods (Figure 1.4): silver staining was found to be 10 times more sensitive than Colloidal Coomassie(0.5 ppm vs 5 ppm).

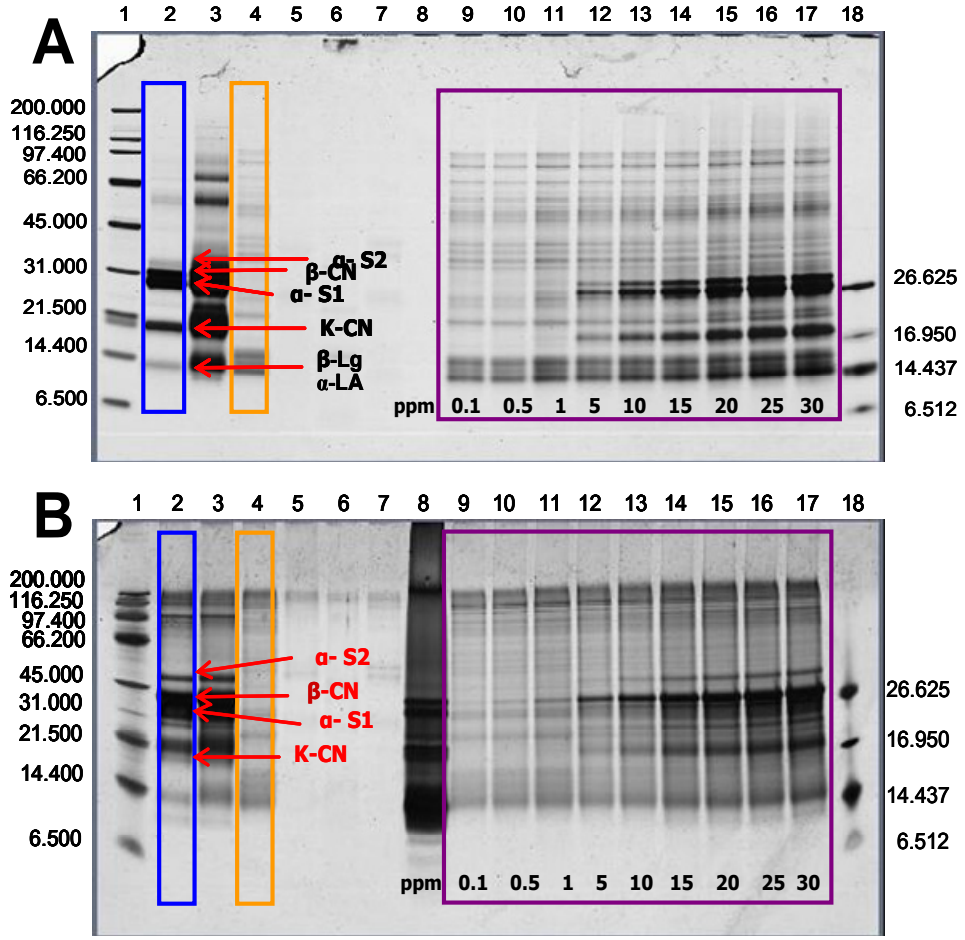


Figure 1.3. (A) Criterion XT Precast gel 10% Bis-Tris; (B) Criterion Precast Gel 16,5% Tris-Tricine/peptide. Line 1: HMW standards; Line 18: LMW standards; Line 2: casein mix; line 4: rice based product; lines 9 to 17 subsequent contamination (indicated in PPM) of the rice based product with the casein mix.

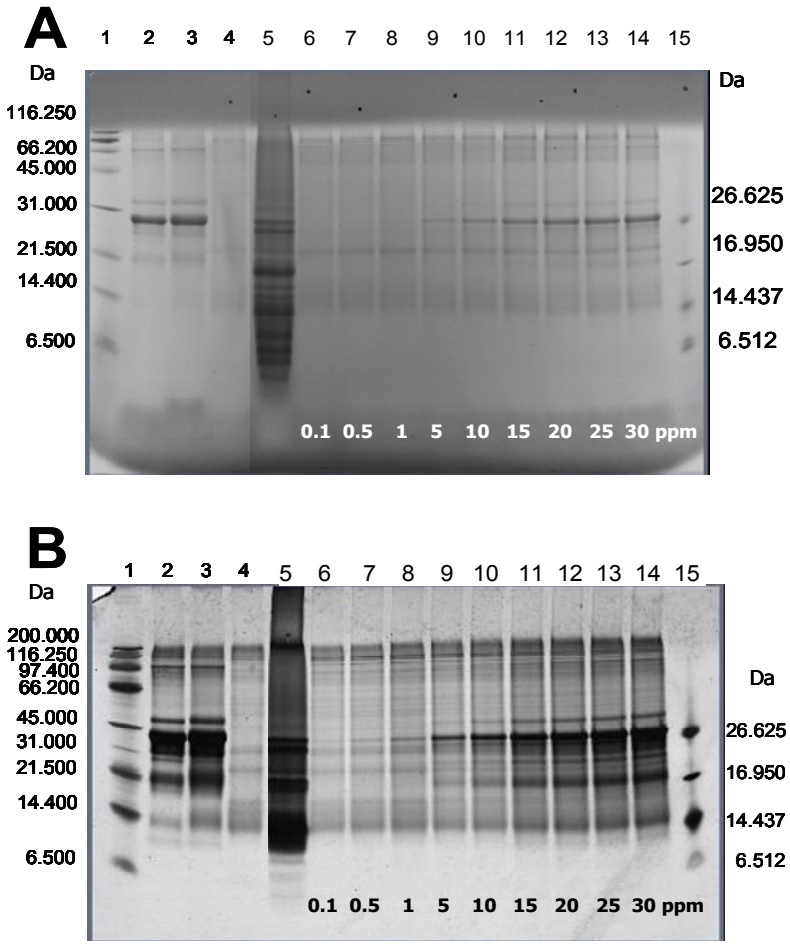


Figure 1.4. Criterion XT Precast gel 10% Bis-Tris for both gels; Line 1, HMW standards, Line 15 LMW standards; lines 2 and 3, milk extract; line 4, rice based product; line 5, casein mix, lines 6 to 14, subsequent contamination (indicated in PPM) of the rice based product with the casein mix. Gel A stained with colloidal Coomassie; gel B stained with Silver staining.

1.4.3 Analysis of the samples.

In the conditions optimized as above reported, samples (described in Table 1.1) were extracted with Carbonate/Bicarbonate buffer and analyzed on a Criterion Precast Gel 16,5% Tris-Tricine/peptide and stained with reversible silver staining. The gel is shown in Figure 1.5.

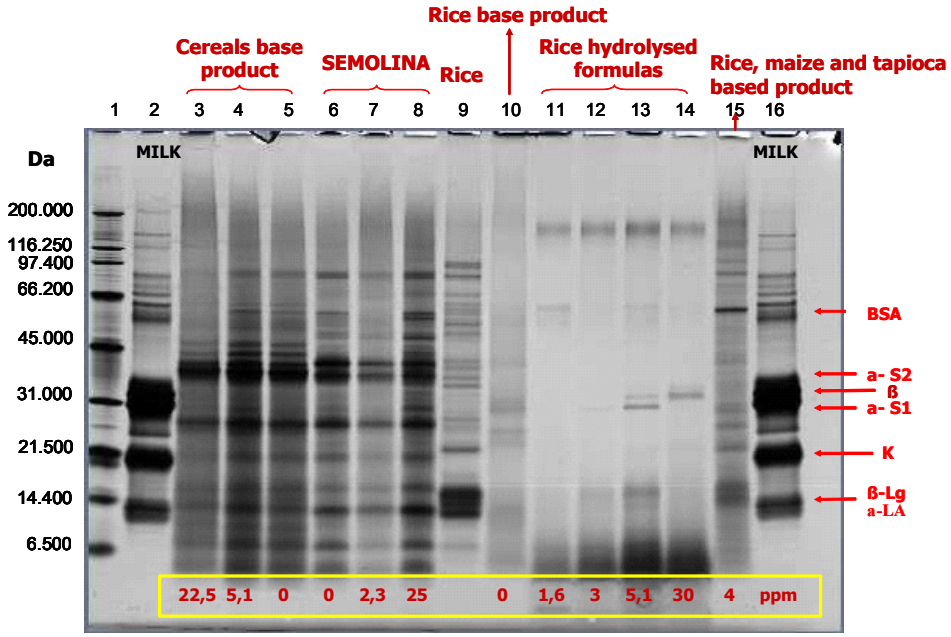


Figure 1.5. Samples analysis by 1D gel electrophoresis. Line 1, HMW standards, lines 2 and 16, milk proteins; lines 3 to 14, samples, ELISA determined contaminating protein content is indicated at the bottom.

Samples resulted very different as far as protein content and hydrolysis degree. Samples in lines 4, 8 and 12, 13 and 14 show gel marks aligned with the milk bands, which may correspond to molecular weights of αS and ß caseins. Due to the low resolution of 1D EF, these data needed to be confirmed, with a specific detection method or a more effective and resolved separation.

1.4.3.1 Western blotting

The same samples were analyzed in the same gel and proteins were transferred onto a PVDF membrane as described before. The correct transfer was monitored by Ponceau Red staining and, after destaining of the red colored membrane, a blotting with sera of milk allergic patients was performed. A pool of three sera with antibody title 1.27, 6.51 and 6.67 µg/ml (RAST determination) was used. Sera were kindly supplied by Prof. Dall'Aglio (Department of Medical Clinic, Nephrology and Prevention Sciences, University of Parma). Their specificity towards several milk proteins is evident in Figure 1.6 as the pool reacted with nearly all protein bands of milk (lines 2 and 17).

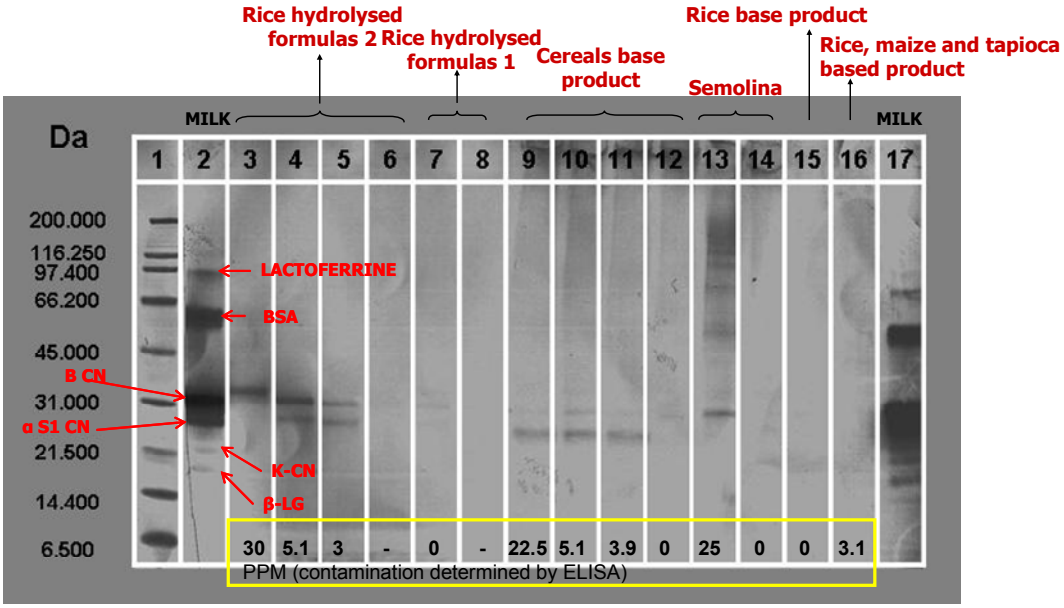


Figure 1.6. Western blotting of the most representative samples of the present work with a pool of sera of patients allergic to milk. Line 1 of HMW markers is relative to the Ponceau stain. Lines 2 and 17 relative to milk show the specificity of sera towards a broad range of protein from milk, αCN and βCN too. Lines 3 to 16 are relative to samples. Lines 6 and 8 are relative to product on sale.

The Western Blot analysis confirmed the presence of allergic proteins, in good agreement with the ELISA data. The bands reacting with sera appeared to correspond to the molecular weights of caseins and were also aligned with the milk bands reacting with sera. Almost all the reacting lines (3, 4, 5, 9, 10, 11, 13) matched with the positive data of the ELISA test. The sample in line 16 was found to be contaminated by the ELISA test, but it did not show any response by Western Blotting, suggesting a possible case of false positive.

Immunoblotting can be more specific and detailed than the ELISA tests but sera are very difficult to obtain, to handle and to store.

1.4.3.2 Bi-dimensional electrophoresis

Bi-dimensional electrophoresis was tested for its high resolving ability. The Bi-dimensional approach was applied on cereal based products, which showed in 1D-EF and Western Blotting reacting marks probably corresponding to αS1 and β caseins in three samples. Two chosen products were found to be contaminated by milk proteins by the ELISA test at 22.5 and 5.1 ppm, while a third was not contamined. The protein fraction was extracted by the Carbonate/bicarbonate buffer and desalted with Amicon devices as described. Desalting was necessary in this case as IEF doesn't work if salts are present.

The dried samples were recovered and prepared for the 1st dimension as described before. For each sample at least 2 replicates were done. After focusing, DTT reduction and IAA alkylation, the strips were loaded onto gels and run on the Criterion cell 2 by 2. After the 2nd dimension, they were stained with Silver or Coomassie colloidal blue staining. 2D maps of the proteome of the 4 cereals (blank) sample stained with the two methods are reported in Figure 1.7.

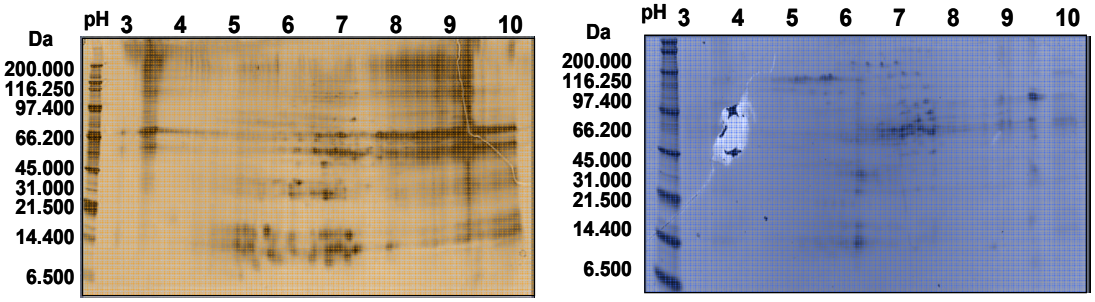


Figure 1.7. Same sample and same condition of loading, IEF and Gel EF, but two different staining. On the left, silver staining, on the right, Coomassie colloidal blue.

Silver staining resulted unambiguously more sensitive than Colloidal Blue. So, Image analysis continued on silver stained gels. Image analysis was performed with PDQuest software, showing different spots in contaminated and blank samples The spots differentiating the blank and the contaminated sample were considered, as well as the spots in common between contaminated sample and milk.

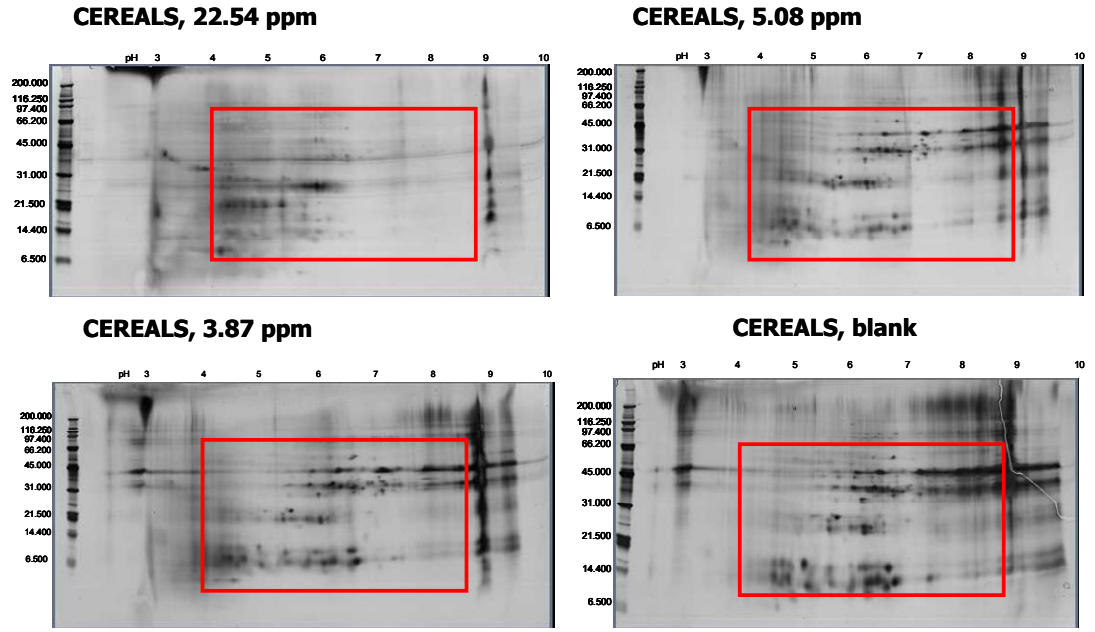


Figure 1.8. Areas of interest of the 2D maps of samples.

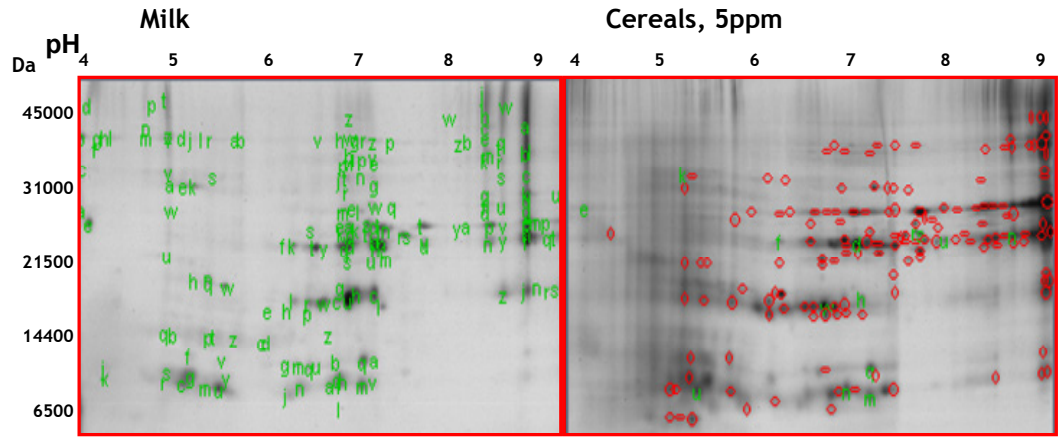


Figure 1.9. Matching of spots between gels. Green spots belong to milk, red spot to the cereals and green spots on the right mean a possible correspondence with a spot on the left (sample contamined).

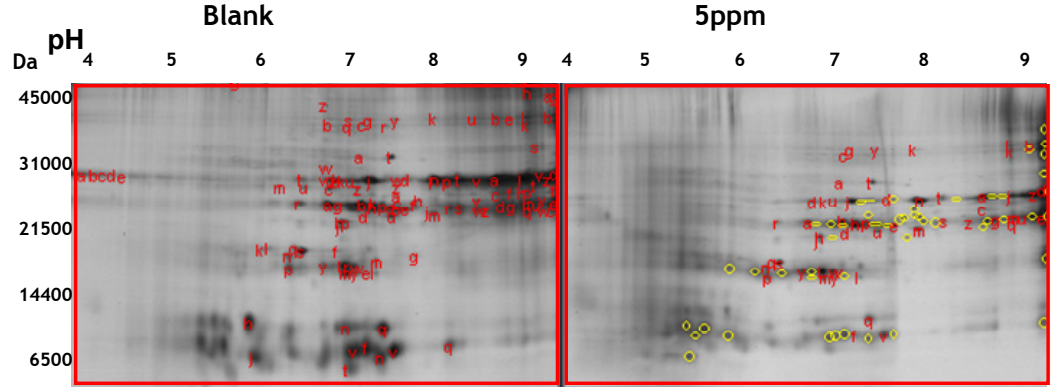


Figure 1.10. Matching of spots between gels. Red spots belong to cereals, yellow spot on the right mean no correspondence with a spot on the left (thus, probably, spots indicating a contamination).

1.4.3.3 Peptide mass fingerprints of spots by MALDI-TOF

The spots differentiating contaminated and blank samples were excised from the gel, destained and subjected to in-gel digestion, as described in the Experimental Part. Peptides were afterwards extracted from the gel and the recovered mixture was spotted on MALDI target together with the matrix. Even if the Silver staining used was described as reversible, spots were hard to destain and peptide ions in MALDI-TOF spectra were very difficult to separate from the background (useful ions, only the blue ones in Figure 1.11).

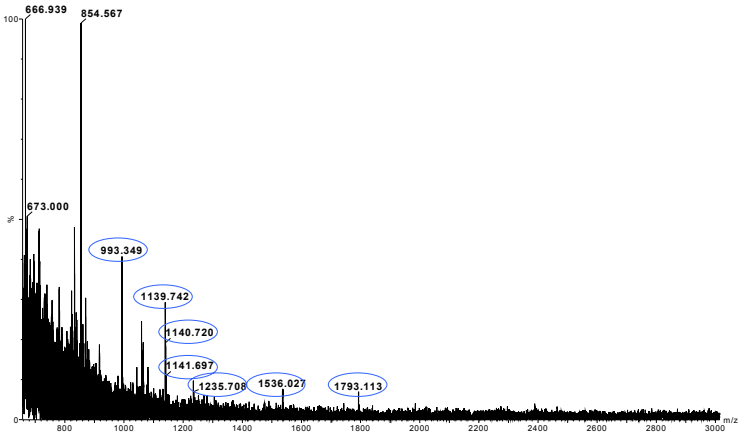
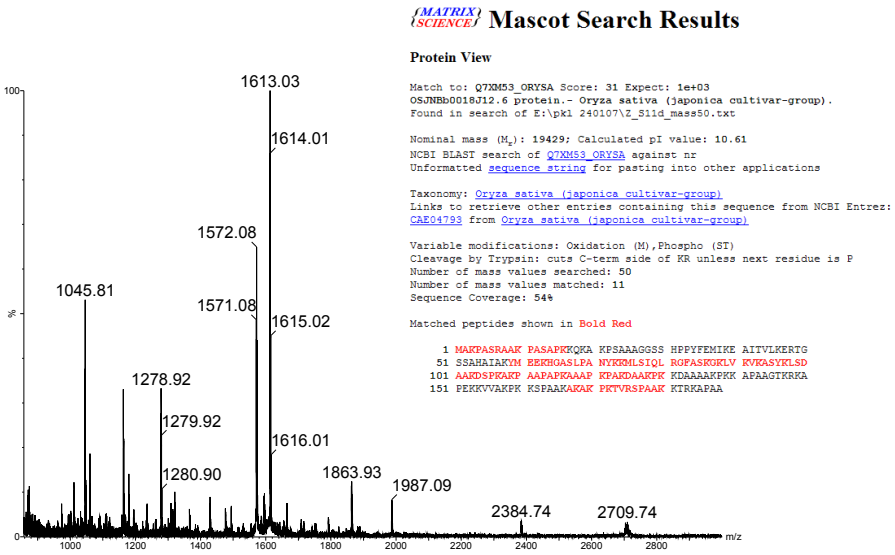


Figure 1.11. MALDI-TOF spectrum from a silver stained spot. In circles, useful ions for a peptide mass fingerprint, but they are hardly differentiated from the background.

Since it was impossible to remove the brown color from the gel, even after cycles of destaining or longer periods of incubation with the destaining solutions, the spots analyzed by MALDI analysis showed a signal suppression.

With such poor spectra, although many spots which differentiated the gels could be determined from the Image Analysis, few data in Mascot research were retrieved and they all had low scores.

The procedure of 2D-electrophoresis, image analysis and in-gel digestion was repeated with the Colloidal Coomassie blue stain. Although the stain is less sensitive spots were correctly destained and analyzed, but the spots excised from the gel resulted in cereal proteins, as in Figure 1.12, and no milk-derived protein was found in the samples.



MATRIX SCIENCE **Mascot Search Results**

Protein View

Match to: Q7XM53_ORYSA Score: 31 Expect: 1e+03
OSJNBb0018J12.6 protein.- Oryza sativa (japonica cultivar-group).
Found in search of E:\pkl 240107\Z_S11d_mass50.txt

Nominal mass (M_r): 19429; Calculated pI value: 10.61
NCBI BLAST search of Q7XM53_ORYSA against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Oryza sativa (japonica cultivar-group)
Links to retrieve other entries containing this sequence from NCBI Entrez:
CAE04793 from Oryza sativa (japonica cultivar-group)

Variable modifications: Oxidation (M),Phospho (ST)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: 50
Number of mass values matched: 11
Sequence Coverage: 54%

Matched peptides shown in **Bold Red**

1 MAKPASRAAK PASAPKKQKA KPSAAAGGSS HPPYFEMIKE AITVLKERTG
51 SSAHAIAKYM EEKHGASLPA NYKKMLSIQL RGFASKGKLV KVKASYKLSD
101 AARDSPKAKP AAPAPKAAAP KPAKDAAKPK KDAAAAKPKK APAAGTKRKA
151 PEKKVVAKPK KSPAAKAKAK PKTVRSPAAK KTRKAPAA

Figure 1.12.MALDI-TOF peptide mass fingerprint of a spot from a Colloidal Coomassie Blue stained gel.

It is very likely that Colloidal Blue Coomassie was not sensitive enough, in these conditions to detect small amounts of milk proteins associated to the low levels of contamination.
The spots stained with silver could not be destained properly, while those stained with Coomassie did not give evidence of peaks corresponding to milk proteins.

1.5 Conclusions and perspectives.

In this part of the thesis, the potentiality of gel electrophoresis, immunoblotting and mass spectrometry has been investigated, in the research of markers of contamination as compared to the common method based on ELISA tests. This approach has been used in the tricky problem of hidden allergens in food, in particular in infant formulas, where a possible contamination can dangerously undermine the health of the allergic young subjects.
A good and non time-consuming method of preparation of the samples was developed and good conditions of gel separation were also determined. By 1D-gels and immunoblotting, ELISA contamination was confirmed on all but one sample (false positive). The 2D-gel and MALDI peptide mass fingerprint (PMF) approach was evaluated on three contaminated and one uncontaminated samples. In the preliminary work here presented, the PMF approach was not able to confirm the contamination, very likely for sensitivity reasons. In the future, this can be achieved by exploring other staining methods, able to attain a very good sensitivity without compromising the MS analysis, or with separation methods targeted to restricted pH values: for example, with small range IPG strips the amount of sample loaded (previously separated by ranges of pH in Rotofor cells for example), can be more concentrated and even a contamination spot can be stained sufficiently with Coomassie Blue Staining.

1.6 References

[1] Høst A. Frequency of cow's milk allergy in childhood. *Ann Allergy Asthma Immunol* **2002**; 89:33–7.

[2] Wal J.M., Cow's milk proteins/allergens, *Ann Allergy Asthma Immunol* **2002**; 89: 3-10

[3] Ann Allergy Asthma Immunol. 2004 Nov;93(5 Suppl 3):S2-11 Bovine milk allergenicity. Wal JM.

[4] Høst, A. Development of atopy in childhood. *Allergy* **1997.** 52:695–697.

[5] Rosendal1 A. and Barkholt V. Detection of Potentially Allergenic Material in 12 Hydrolyzed Milk Formulas *J Dairy Sci* **2000** 83:2200–2210

[6] Høst A, Halkan S. Hypoallergenic formulas: when, to whom and how long: after more than 15 years we know the right indication. *Allergy* **2004**; 59:45–52.

[7] Fiocchi A., Restani P., Bernardini R., Lucarelli S., Lombardi G., Magazzùk G., Marseglia G. L., Pittschieler K., Tripodi S., Troncone R. and Ranzini C. A hydrolysed rice-based formula is tolerated by children with cow's milk allergy: a multi-centre study, *Clinical and Experimental Allergy*, **2006**36, 311–316,

[8] Besler M., Determination of allergens in foods *Trends in analytical chemistry*, **2001**, 20, 11, 662-672

[9] van Hengel A. J. Food allergen detection methods and the challenge to protect food-allergic consumers *Anal Bioanal Chem* **2007** 389:111–118,

[10] S. Rossi, E. Scaravelli, A. Germini, R. Corradini, C. Fogher, R. Marchelli, A PNA-array platform for the detection of hidden allergens in foodstuffs, *Eur Food Res Technol* **2006** 223: 1–6

[11] Koch P.; Schäppi G. F.; Poms R. E.; Wüthrich B.; Anklam E.; Battaglia R. Comparison of commercially available ELISA kits with human sera-based detection methods for peanut allergens in foods *Food Additives & Contaminants* **2003** 20, 9, 797–803

[12] Hennion M.-C., Pichon V.Immuno-based sample preparation for trace analysis , *Journal of Chromatography A,* **2003** 1000: 29–52

[13] Cohen S. L. and Chait B. T., Mass spectrometry of whole proteins eluted from sodium dodecyl sulfate–polyacrylamide gel electrophoresis gels *Anal Bioch.* **1997** 247: 257–267

2 CHARACTERIZATION AND LOCALIZATION OF ALLERGENIC PROTEINS IN ROSACEAE FRUITS BY HIGH RESOLUTION MASS SPECTROMETRY (HRMS) AND MASS SPECTROMETRY IMAGING (MSI)

2.1 Introduction: Lipid Transfer Proteins as plant protein and food allergens

A large non-specific lipid transfer protein (nsLTP) family, comprising over 100 potential members from up to 50 different species, has been described since the discovery of these proteins 30 years ago.[1,2] The 9 kDa nsLTP family is widely distributed throughout the plant kingdom. It includes basic polypeptides (pI 8.5–10) with 90–95 amino acid residues, lacking Trp. Eight conserved cysteine residues form a network of 4 disulphide bridges, which are essential to maintain the nsLTP fold and, therefore, the lipid-binding properties of these proteins. nsLTPs are encoded by a divergent multigene family, dispersed in different chromosomes, modulated by biotic and abiotic stress: studies have been carried out on *Arabidopsis thaliana*[3], *Oryza sativa*[4] and *Triticum aestivum*[5].

The possible functions of nsLTPs in plants have been extensively debated in the recent years, and their role in plant defense mechanisms against phytopathogens (bacteria and fungi) seems to be established[6,7]. The nsLTP family is ubiquitously expressed throughout the plant kingdom but the highest expression levels have been found in peripheral cell layers surrounding aerial organs, associated with cell wall and cuticle of epidermal tissues.

The 9 kDa members of this family show a broad lipid-binding specificity[8], including fatty acids, fatty acyl-CoA, phospholipids, glycolipids, hydroxylated fatty acid. The lipid-binding capacity of these proteins is closely related to the peculiar characteristics of the nsLTP fold. Several 3-D structures of 9 kDa plant nsLTPs, both free and forming complexes with different lipidic ligands, have been determined by either X-ray crystallography or nuclear magnetic resonance. All of them show an α-helical compact domain composed of 4 α-helices connected by short loops and a non-structured C-terminal tail. The fold is strongly held by a network of 4 disulphide bridges which is essential for lipid-binding. Additionally, a large number of intramolecular H-bonds contribute to the stabilization of the 3-D protein structure. Besides its unusual stability, the main feature of nsLTP folding is the presence of a large internal tunnel-like cavity following the long axis of the molecule. The surface of this cavity is coated with the side chain of hydrophobic residues included in the four helices (H1–H4) and the C-terminal tail (Figure 2.1). Relevant clinical implications can be derived from the accumulation of allergenic nsLTPs in the external tissues of plant foods, such as the skin of fruits.[9]

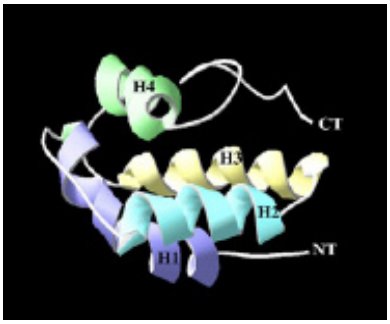


Figure 2.1. 3-D structure of the nsLTP and major allergen from peach fruit, showing the nsLTP-fold.

Moreover, an unexpected role of nsLTPs from the Rosaceae fruits in IgE-mediated allergic reactions was discovered in 1999 after studying adult patients from the Mediterranean area with allergy to these fruits. Altogether, members of the nsLTP family have been identified as allergens in plant foods[10], latex and pollens.[11] The large clinical evidence has allowed to define nsLTPs as a relevant pan-allergen family of both plant foods and pollens.[12,13]

nsLTPs trigger IgE-mediated food allergy in particular for fruits of the Rosaceae family (apple, peach) in the adult population of the Mediterranean area[14]. Allergy to these fruits was associated with birch pollinosis[15], frequent in Central and North European areas where *Fagales* trees are abundant. Patients usually show mild symptoms restricted to the oropharyngeal cavity (OAS), and seem to be primarily sensitized by pollen allergens. The subsequent reaction to plant foods, such as apple, is based on cross-reactivity. On the contrary, the population allergic to Rosaceae fruits in Mediterranean countries is rarely exposed to Fagales pollen and displays a different clinical and sensitization profile. Severe systemic manifestations are frequently found, and a relevant number of patients suffer fruit allergy without linked pollinosis. An uncharacterized IgE-binding band of approximately 10–13 kDa was initially located in crude extracts of peach and other Rosaceae fruits (apple, cherry, plum) as early as 1994, using sera from patients sensitized to fruit but not to pollen[16]. The reactive component accumulated in fruit skins and seemed to be different from those allergens associated with birch pollinosis. The identification of the 10–13 kDa IgE-binding component from peach and apple as members of the nsLTP family was independently published in 1999 by Spanish and Italian groups.[17,18,19] A major allergen of around 9 kDa was isolated from each species and designated Pru p 3 (peach, *Prunus persica*) and Mal d 3 (apple, *Malus domestica*), according to the guidelines of the I. U. I. S International Union of Immunological Societies[20]. Thus, allergenic members of the nsLTP-family have been identified as allergens not related to birch pollinosis and at least partially characterized in other Rosaceae fruits (apricot, plum, cherry, etc.), and in other families: grape, strawberry, citrus fruits (orange and lemon), tomato, vegetables (asparagus, cabbage, asparagus, lettuce, etc.), nuts (hazelnut, walnut, etc.), maize, olives. Besides these well-established and officially designated allergens, additional members of the LTP family with

potential allergenic capacity have been isolated or detected in other foods and plants, such as pomegranate, onion, carrot, peanut, rice, spelt wheat, beer and recently in marijuana[21].

Several characteristics of the nsLTP panallergen family have led to propose it as a model of true food allergens[22,23]. Its resistance to proteolytic digestion (and heat treatment) suggests that its allergenic members retain their immunogenic and allergenic motifs after passing through the gastrointestinal tract, and can interact with the associated epithelial immune system to induce both sensitization and systemic symptoms.[24]

2.1.1 Peach (Prunus persica) LTP, Pru p 3

Peach allergy is the most common form of IgE-mediated hypersensitivity to fresh fruits in the Mediterranean area.[25] Its prevalence can be estimated to 10-40% (Spain and Italy) in pollen allergic patients or even up to 75% (in Israel) in fruit and/or vegetable allergic individuals. Peach allergy is rarely observed as an isolated allergy, and most patients present with some other food or inhalant (mainly pollen) allergies. The foods most frequently associated are other members of the Rosaceae family, such as apple and pear (Pomoideae subfamily), and apricot, cherry and plum (Prunoideae subfamily). According to clinical observations, apple allergy is the most frequent food allergy associated to peach allergy.

As regards sensitization to peach, it is linked to birch pollinosis in northern Europe and here allergy to peach is mainly due to cross- reactive IgE induced by Bet v 1 (the major birch pollen allergen). In contrast, peach allergy in southern Europe is a "true food allergy"[26], in which sensitization and reactions are induced by stable fruit allergens (LTP). Pru p 3 has been proven to be resistant against pepsin digestion[27]. This LTP has been characterized in recent years: for example, the 3D structure has been elucidated[28] and some of the possible allergic epitopes of the protein were designed and their effective reactivity with allergic sera was evaluated[29].

There are still some discrepancies in its primary structure sequence. In 1992 an IgE-binding protein isolated from peach with a 8-10kD MW was characterized immunologically and named Pru p 1.[30] In 1999 the same protein was characterized as an LTP, reporting the sequence of its first 21 amino acids (obtained by Edman sequencing) and its MW (9138 Da)[18], obtained by MALDI-TOF. It was also suggested to rename this LTP as Pru p 3[18]. In the same year the complete amino acid sequence of the protein (91 amino acids) was published,[17] although the calculated MW (9170 Da by considering also disulfide bridges) was not consistent with that previously measured by MALDI-TOF. Later on,[31] two c-DNA clones obtained from different tissues of the peach plant (named Pp-LTP1 and Pp-LTP2) were also isolated and sequenced: Pp-LTP1 was found to encode a 9 kDa allergenic protein whose deduced amino acidic sequence was consistent with Pru p 1, although with several differences. Quite interestingly, the calculated MW of the protein reported in the database (9127 Da by considering also disulfide bridges) was also different from that previously measured by MALDI-TOF[18]. The same group, four years later[32], reported another sequence for the same clone (Pp-LTP1) obtained from different peach varieties. Surprisingly, the

calculated MW (9108 Da by considering also disulfide bridges) of the protein (now named Pru p 3) turned out to be different from all the previous ones. Meanwhile another group[33] (Diaz Perales et al.) reported the sequence of the cDNA enconding for the same protein. The deduced protein MW, inclusive of disulfide bridges, was consistent, within the experimental error, with that of the natural protein measured by MALDI-TOF (9136 Da). This sequence also differs from that obtained by direct amino acid sequencing in positions 9 (Ser/Ala) and 76 (Ser/His) and accordingly also the MW is different. The protein sequences present in the database are reported in Table 2.1.

Table 2.1. Sequences for peach LTP reported in UniProt database.

Reference	Protein sequence
17	ITCGQVSSAL APCIPYVRGG GAVPPACCNG IRNVNNLART TPDRQAACNC LKQLSASVPG VNPNNAAALP GKCGVHIPYK ISASTNCATV K
18	ITCGQVSSSL APCIPYVRGG GA (by Edman N-terminal sequencing) and MW 9138Da
31	ITCGQVSSNL APCIPYVRVG GAVPPACCNG IRNVNNLART TPDRQAACNC LKQLSASVPG VNPNNAAALP GKCGVHIPYK ISASTNCATV
32	ITCGQASSSL APCIPYVRGG GAVPPACCNG IRNVNNLART TPDRQAACNC LKQLSASVPG VNPNNAAALP GKCGVSIPYK ISASTNCATV K
33	ITCGQVSSSL APCIPYVRGG GAVPPACCNG IRNVNNLART TPDRQAACNC LKQLSASVPG VNPNNAAALP GKCGVSIPYK ISASTNCATV K

2.1.2 Garden plum (Prunus domestica) LTP, Pru d 3

Like many other allergies to fresh fruits and vegetables, plum allergy can take two different forms. In the North of Europe, people with birch-pollen allergy can develop a plum allergy due to the similarity between a protein in birch that causes birch-pollen allergy and a plum protein. In Mediterranean countries, people with plum allergy do not have birch-pollen allergy. Instead, they often have allergy to peach. Symptoms include generalized urticaria, abdominal pain, vomiting and life-threatening symptoms, sometimes in addition to the Oral Allergy Syndrome. These individuals tend to have more frequent and severe reactions when fruits are eaten with the skin. They also tend to develop adverse reactions to other fruits including apple, peach, apricot, cherry and nuts (such as hazelnut and walnut). The allergen that causes this kind of allergy is a LTP[34]. Little information is available about the primary structure of plum LTP. In 2001 when the plum allergen of the Mediterranean area was characterized as an LTP, the sequence of its first 25 amino acids (obtained by Edman sequencing) was reported.

2.2 Aim of the work

In order to overcome the discrepancies existing in the literature for the primary structure of peach LTP and to overcome to the lack of information on the primary structure of plum LTP, we decided to extract and purify the proteins from different fruit varieties and to perform a full structural characterization by exploiting the potential of mass spectrometric techniques, in

particular ESI-Q and ESI-LTQ-Orbitrap. With HMoreover, MALDI Imaging was used to localize the proteins in the fruit.

2.3 Experimental part

2.3.1 Reagents

- Acetone, (CH_3COCH_3) RS per HPLC;
- Ammonium bicarbonate, (NH_4HCO_3), (FLUKA);
- Acetonitrile (CH_3CN), (CHROMANORM);
- Methanol (CH_3OH), (CHROMANORM);
- Chloridric acid (HCl) solutions, from HCl 37% (CARLO ERBA);
- Dithiothreitol, [DTT], (FLUKA);
- Ethylenediaminetetraacetic acid, [EDTA], (CARLO ERBA);
- Formic acid (HCOOH), 99%, (ACROS ORGANICS);
- Iodoacetamide, [IAA], 97%, (SIGMA-ALDRICH).
- Liquid nitrogen
- Milli Q H_2O obtained with Millipore Alpha Q system;
- Polyvinyl polypirrolidone, [PVPP], (SIGMA-ALDRICH);
- Potassium chloride (KCl), (CARLO ERBA);
- Potassium hydrogen phosphate (KH_2PO_4), (FLUKA);
- Sinapinic acid, SA, ($C_{11}H_{12}O_5$), (FLUKA);
- Sodium chloride (NaCl), (BDH);
- Sodium hydrogen phosphate (Na_2HPO_4), (RPE);
- Sodium hydroxide (NaOH), dry, (CARLO ERBA);
- Trifluoroacetic acid, [TFA], ($CF_3CO_2H.$), 99%, (ACROS ORGANICS);
- Trypsin from bovine pancreas, Proteomics Grade, Dimethylated (T6567, SIGMA-ALDRICH);
- Standard of calibration for MALDI;
- PBS buffer, prepared as follows:
  - 8 g NaCl (137 mM),
  - 0.2 g KCl (2.7 mM),
  - 1.44 g Na_2HPO_4 (10 mM) weighed from Na_2HPO_4 *$12H_2O$,
  - 0.24 g KH_2PO_4 (1.5 mM),
  - 800 ml of Milli Q H_2O,
  - Adjust the pH to 7.4 with HCl,
  - Milli Q H_2O to 1 liter.

2.3.2 Instrumentation

- Amicon Ultra-4-Centrifugal filter devices, nominal cut off 5000 Da, Millipore;
- Centrifuge, ALC PK110;
- Centrifuge, refrigerated, ALC 4237R;
- Centrifuge, refrigerated, UNIVERSAL 32OR Hettich;
- Cryostat, LEITZ 1720, Ernst Leitz Wetzlar GmbH, Germany;
- Homogenizer, UltraTurrax T50 basic, IKA-WERKE;
- Magnetic agitator;
- Immersion mixer, BRAUN;
- Table mixer, Moulinex
- pH-meter 691, Metrohm;
- Rotavapor, BÜCHI 461 WATER BATH;
- Sonicator, ULTRAsonikTM 57X NEY;
- Termocycler, PCR SPRINT, THERMO HYBAID;
- Semipreparative HPLC
  - C8 semipreparative column, XTerra Prep RP8 column (7.8x300 mm, particle size 10 µm, pore size 120 Å),
  - Automated gradient controller, (Waters);
  - Pumps model 510, (Waters);
  - Tunable absorbance detector 486 (Waters);
  - Printer data module 745 (Waters);
- Analytical HPLC
  - C18 analytical column, PHENOMENEX GEMINI C18, 5 µm 110 Å (250 × 4,6 mm);
  - PHENOMENEX JUPITER column, C18, 5 µm, 300 Å (250x4.6 mm);
  - PHENOMENEX JUPITER column, C4, 5 µm, 300 Å (250x4.6 mm);
  - HPLC ALLIANCE WATERS 2695;
  - Photodiode Array Detector WATERS 996;
  - Mass Spectrometer WATERS ACQUITY SQ Detector with ESI interface;
- Software Mass Lynx 4.1
- Ultimate 3000 nano HPLC (Dionex)
- Mass Spectrometer ESI- LTQ Orbitrap (Thermo-Scientific);
- Mass Spectrometer MALDI-TOF/TOF UltraflexTM (Bruker);
- Software FlexImaging 2.0TM

2.3.3 Samples

2.3.3.1 Peach

Peach LTP was extracted from three different peach varieties: "Italia K2" (white flesh), "Toscana" (yellow flesh) and "Rita star" (Nectarine yellow flesh), all available on the Italian market.

2.3.3.2 Plum

Plum LTP was extracted from six varieties of European plum: RED BEAUTY (red-purple color), GOCCIA D'ORO (yellow-amber), REGINA CLAUDIA (yellow-green), BLACK (blue-black), AMOLA (red) and FORTUNA (green), all available on the Italian market.

2.3.4 Total protein extraction

2.3.4.1 Peach

Skins were obtained by cutting 2 to 3 mm from the external part of fruits. 50 g of ground skins and pulps of peaches were homogenized in 150 ml of a phosphate buffer solution (PBS, pH=6.8), at a 25% w/v ratio and extracted for 2 h at 4°C under continuous stirring after addition of polyvinyl-polypirrolidone (PVPP) (4% w/w). After centrifugation (3400 g[†] for 45 min at 4 °C), the surnatant was filtered through filter paper, then dialyzed and concentrated on Amicon-ultra-4 centrifugal filter devices with a nominal molecular cut-off of 5 kD. The surnatant was recovered, filtered through 0.45µm filters and concentrated under a nitrogen flux (1 ml from).

2.3.4.2 Plum

The LTP extraction used for peach was slightly modified as plum has a higher content of phenols which may interfere with the analysis. Plum skins were obtained by cutting 1 to 2 mm from the external part and immediately placed in liquid nitrogen to freeze up. Frozen skins were ground on a Moulinex mixer and 40 g were homogenized by Ultraturrax in 16 ml of phosphate buffer solution (PBS, pH=6.8), at a 2:5 w/v (skins/PBS) ratio. EDTA (2mmol/l final concentration in 16 ml) and 2% PVPP were added and the mixture was kept for 1 hour at 4°C under continuous stirring. After centrifugation (3400g for 45 min at 4 °C), the surnatant was filtered through filter paper, then dialyzed and concentrated on Amicon-ultra-4 centrifugal filter devices with a nominal molecular cut-off of 5 kD. The surnatant was recovered, filtered through 0.45µm filters and concentrated under a nitrogen flux.

2.3.5 LTP purification

2.3.5.1 Peach

The total protein extract from the skins or the pulps was dissolved in 0.1% formic acid in H_2O and first purified by a semipreparative HPLC system by using a Waters XTerra Prep RP8 column,

[†] g, acceleration gravity

at a 4 ml/min flow rate; eluent A: 0.1% HCOOH + 0.2% CH_3CN in Milli Q H_2O; eluent B: 0.1% HCOOH + 0.2% Milli Q H_2O in CH_3CN. Gradient: 10 min isocratic elution with 100% A, 30 min of linear gradient from 0% to 60% of B, 10 min isocratic elution with 40% A and 60% B, and reconditioning. 150 µl of total protein extract were injected each time.

The fraction eluting from 20 to 45 min was collected, dried, redissolved in water and purified again with an analytical HPLC system by using a Jupiter Phenomenex column, C18 (250x4.6 mm, particle size 5 µm, pore size 300 Å). Flow rate: 1 ml/min, eluent A 0.1% HCOOH + 0.2% CH_3CN in Milli Q H_2O; eluent B 0.1% HCOOH + 0.2% Milli Q H_2O in CH_3CN. Gradient: 5 min isocratic elution with 100% A, 35 min of linear gradient from 0% to 50% of B, 10 min of isocratic elution with 50% A and 50% B, and reconditioning. 70 µl of total protein extract were injected each time. Purification was efficiently achieved by using a T-split before the ESI interface single quadrupole analyzer. The pure protein was collected at its elution time by using the MS detector for the identification of its characteristic MS ions. ESI-MS spectrometer conditions were the following: positive ion mode, capillary voltage 3 kV, cone voltage 30 V, source temperature 100°C, desolvation temperature 150 °C, spraying gas (N_2) 100 l/h, desolvation gas (N_2) 400 l/h, full scan acquisition from 150 to 1500 m/z in continuum mode and 2.9 sec of scan time.

2.3.5.2 Plum

The total protein extract from the skins was dissolved in 0.1% HCOOH in H_2O and immediately purified on an analytical HPLC system by using a GEMINI PHENOMENEX column, as the quantities required were lower than the ones used for peach LTP analysis. Later, a Phenomenex Jupiter C4 column was used with better results. Flow rate, gradient and purification by T-split were the same as in paragraph 2.3.5.1, except for the MS spectrometer, which this time was a Waters SQ single quadrupole detector. ESI-MS spectrometer conditions were the following: positive ion mode, capillary voltage 3.2 kV, cone voltage 30 V, source temperature 100°C, desolvation temperature 150 °C, spraying gas (N_2) 100 l/h, desolvation gas (N_2) 400 l/h, full scan acquisition from 100 to 1900 m/z in continuum mode and 3 sec of scan time.

2.3.6 Tryptic digestion of purified LTP

The purified LTPs were quantified by weight. LTP digestion was achieved with the incubation of 55 µg of the protein for 24 hours with 2 µg of trypsin (enzyme:substrate ratio 1:27.5) in 50 mM ammonium hydrogen carbonate (NH_4HCO_3) buffer, pH 7.8; the digestion was followed by disulfide bridge reduction with 2.8 µl of 20 mM dithiotreitol (DTT) solution for 1 h at 54 °C and alkylation of free thiols with 11.2 µl of a 20 mM iodoacetamide (IAA) solution for 1 h at room temperature in the dark with occasional vortexing. Exceeding IAA was neutralized by adding 11.2 µl of a DTT 20 mM solution.

As far as peach LTP is concerned, peptide analysis was performed with an ALLIANCE WATERS 2695 HPLC system, interfaced with an ESI-MICROMASS ZMD (single quadrupole) by using a Jupiter

Phenomenex Column C18 (300 Å, 250 x 4,6 mm, 5 µm), applying the conditions described above. Flow rate: 1 ml/min; eluent A: 0.1% HCOOH + 0.2% CH_3CN in Milli Q H_2O; eluent B: 0.1% HCOOH + 0.2% Milli Q H_2O in CH_3CN. Gradient: 5 min isocratic elution with 100% A, 35 min linear gradient from 0% to 40% of B, 10 min isocratic elution with 60% A and 40% B, and reconditioning. Injection volume: 50 µl.

As far as plum LTP is concerned, peptide analysis was directly performed on a Dionex Ultimate 3000 nano HPLC coupled with the LTQ-Orbitrap mass spectrometer equipped with a conventional ESI source. Spray voltage was set at 2.8 kV, capillary voltage at 100 V and tube lens at 175 V; flow rate was 3 µl/min. The MS/MS spectra were registered at a nominal resolution of 60000 (at m/z 400) and mass range 1200-2000 m/z. (measures by CISM-Florence, Italy).

2.3.7 Exact mass determination

The lyophilized protein was resuspended in 0.1% TFA:CH_3CN 1:1 (v/v) at a final concentration of about 50 pmol/µl. Direct injection spectra were acquired on a LTQ-Orbitrap (ThermoFischer, San José, CA, USA) equipped with a conventional ESI source. Voltages and acquisitions were set as just described, for accurate mass determination.

The peach LTP sample was reduced by TBP (Tributylphosphine) and analyzed by direct infusion on the MS Orbitrap spectrometer. The plum LTP sample was reduced with DTT and alkylated with IAA on the Single Quadrupole mass spectrometer.

2.3.8 LTP characterization by the TOP-DOWN approach

The peach LTP sample used for accurate mass determination was treated with 5 mM tributylphosphine for 30 min at 37 °C in order to reduce the disulfide bridges.

The mass spectrum of the reduced LTP was registered with the same parameters described above. Top Down MS/MS experiments were performed on 7+, 8+ and 6+ ions in order to obtain sequence information. A collision energy of 20 (instrumental arbitrary units) and an activation time of 35 msec were used. Data were analyzed with the ProSightPTM software (prosightptm.scs.uiuc.edu).

2.3.9 MALDI IMAGING

A few thin slices (about 250 µm thick) were cut from the frozen fruit (peach or plum) by means of a Leitz 1720 cryostate and mounted on conductive ITO microscope slides, then sprayed with a 20 mg/ml solution of sinapinic acid in ACN 50% and TFA 0.2%. Multiple spraying cycles were performed, allowing solvent evaporation between cycles, until complete matrix coverage of the tissue was obtained. Microscope slides were placed on a Bruker MALDI target and submitted to analysis in a Ultraflex™ MALDI-TOF/TOF mass spectrometer. Spectra were acquired from tissue slices with a 400 µm spatial resolution, both on x- and y-axis; 500 shots were acquired from each position. Collected data were analyzed with a FlexImaging 2.0™ software.

2.4 Results and discussions

2.4.1 Peach LTP

2.4.1.1 LTP purification

The first aim of the work was to develop an easy and non time-consuming extraction method for LTP. Beginning with peach, the samples were treated with the chelating agent polyvinylpyrrolidone (PVP) in PBS buffer in order to remove polar impurities (such as phenols), then were desalted and concentrated on ultrafilters with a molecular cut off of 5kD. A first concentration step of the extracts was obtained by HPLC by using a semipreparative C8 HPLC column. The fraction recovered by eluting from 20 to 45 min was further purified by using an analytical RP C18 HPLC column with an ESI-MS (single quadrupole) analyzer.

The Pru p 3 protein was extracted from three different peach varieties: white flesh "Italia K2", yellow flesh "Toscana" and Nectarine yellow flesh "Rita star" variety. For each variety the content was investigated both in pulps and skins. The HPLC-UV chromatogram obtained after the first semipreparative preconcentration of the total protein extracted from pulp and skin of the "Italia K2" variety is reported in Figure 2.2. It clearly shows the almost exclusive presence of LTP in the skin of the peaches. The white flesh "Italia K2" variety appeared to be the richest in LTP as compared to the other varieties. (Figure 2.3)

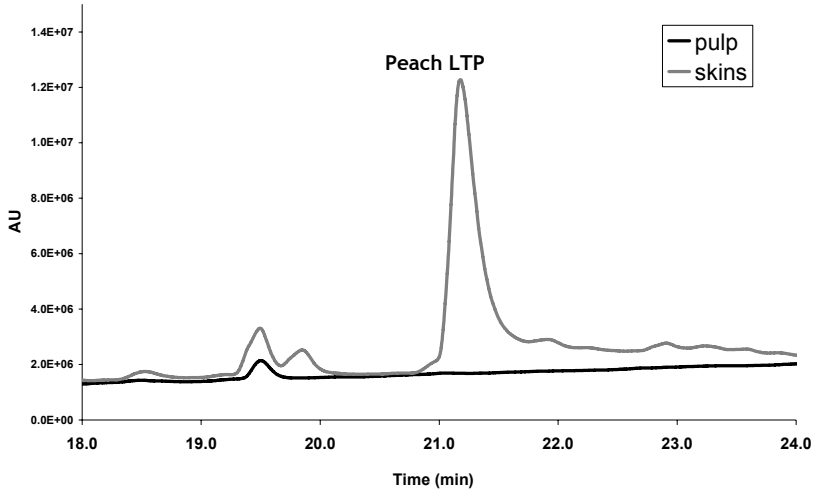


Figure 2.2. HPLC chromatogram (UV detection at 240 nm) of the crude extracts from the peach skins and pulps of white flesh "Italia K2" variety.

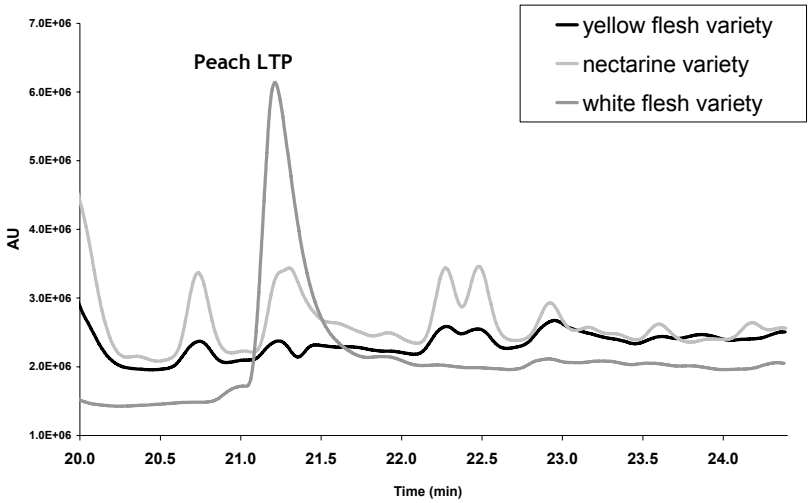


Figure 2.3. HPLC chromatogram (UV detection at 240 nm) of the crude extracts from peach skins of different varieties.

The ESI-MS spectrum of the protein is shown in Figure 2.4: ions with m/z 1306.1, 1142.9 and 1016.0 are clearly evident, corresponding respectively to the 7, 8 and 9 charged ions, with an average MW of 9135±1 Da.

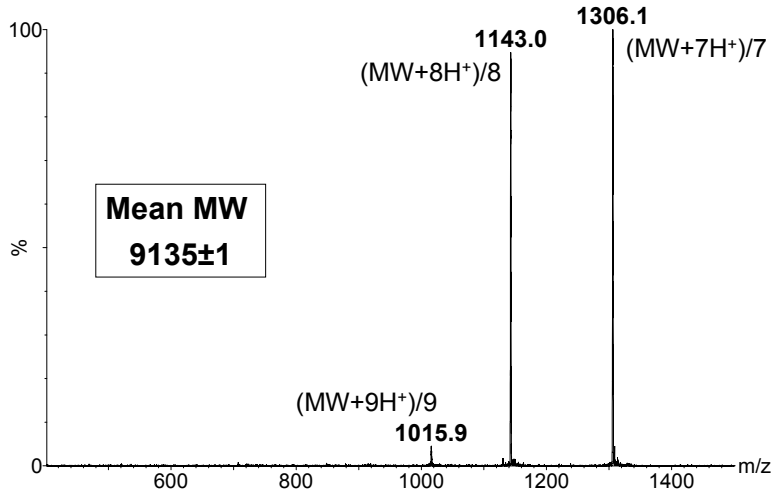


Figure 2.4. Annotated ESI-MS (single quadrupole) for the purified peach LTP protein.

The MW of the protein perfectly matched the sequences Q9LED1 reported by Diaz-Perales et al.[33] and it was consistent with the MW observed by Sanchez-Monge et al.[18] by using MALDI-TOF. It should be underlined that all varieties showed the presence of the same protein with the same molecular weight, indicating that the same LTP was present in the three peach varieties. The ions corresponding to the MW of the other sequences reported in the literature were also searched, but were not found in any variety. No isoforms were detected in this work.

The protein was further purified by using the same analytical RP C18 HPLC column interfaced with the ESI-MS (single quadrupole) analyzer. The column eluate was allowed to flow through an on-line PDA detector followed by a single quadrupole mass spectrometer, but only 20% (200 µl/min) was allowed to enter the MS analyzer by a T-split. Ions at 1306 m/z (corresponding to the 7H+ charged protein) and 1143 m/z (corresponding to the 8H+ charged protein) were monitored by the XIC (eXtract Ion Chromatogram) technique for the definitive purification of the protein: by using the XIC for determining the effective elution time, the protein was recovered at a very high degree of purity. A UV chromatogram obtained after the purification is shown in Figure 2.5.

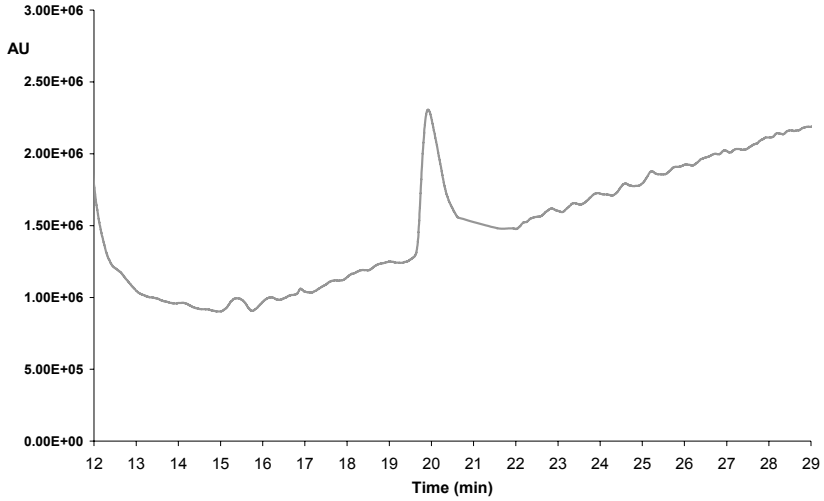


Figure 2.5. HPLC Chromatogram (UV detection at 240 nm) of the purified peach LTP protein.

Thus, a new method for LTP purification was developed, essentially based on two subsequent purification steps: preparative and analytical HPLC of the extracted peach fraction. In one-two days of work, this method easily allows to obtain enough LTP at a high degree of purity for full characterization, avoiding the use of low-resolution gel permeation chromatography.

2.4.1.2 Peach LTP characterization

i Bottom-up approach: trypsin digestion

The tryptic peptides deriving from purified peach LTP were analyzed by HPLC/ESI-MS (Figure 2.6), and characterized according to their MW and to the fragments generated by the in source fragmentation (Table 2.2). This was possible for peach LTP because of the amount of information available for its primary structure.

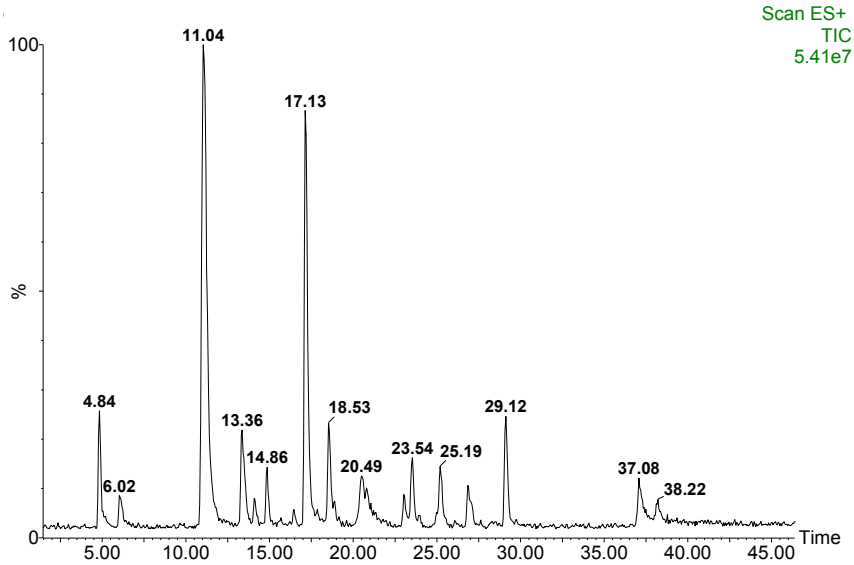


Figure 2.6. HPLC-MS chromatogram of peptide mixture obtained by tryptic digestion, reduction and alkylation of peach LTP.

Table 2.2. Identification of the tryptic peptides generated by peach LTP digestion

r_t (min)	Assigned sequence[a]	Position	Calculated MW (Da)[b]	Found MW (Da)	Diagnostic fragments (m/z)
6.2	VR	17-18	273.2	274	275 (MH^+); 158 (z_1); 175 (y_1)
13.4	TTPDR	40-44	588.3	588	589 (MH^+); 488 (y_4); 387 (y_3); 290 (y_2); 175 (y_1)
14.9	NLAR	36-39	472.9	473	359 (y_3); 246 (y_2); 175 (y_1); 299 (b_3)
18.5	ISASTNC*ATVK	81-91	1152.3	1152	1152 (MH^+); 1039 (y_{10}); 881 (y_8); 794 (y_7); 398 (y_7)[a]; 577 (MH_2^{2+})[c]
18.6	NAAALPGK	65-72	740.4	741	742 (MH^+); 596 (b_7); 539 (b_6); 444 (b_5)
20.6	GGGAVPPAC*C*NGIR	19-32	1387.6	1386	1046 (b_{11}); 695 (MH_2^{2+})[c]; 524 (y_9)[c]
23.6	C*GVSIPYK	73-80	923.4	924	925 (MH^+); 463 (MH_2^{2+})[c]; 707 (y_6); 608 (y_5)
24.0	QLSASVPGVNPN	53-64	1182.3	1182	592 (MH_2^{2+})[c]; 1083 (MH^+)
25.3	QLSASVPGVNPNNAAALPGK	53-72	1905.1	1906	954 (MH_2^{2+})[c]; 662 (y_{14})[c]; 747 (b_{16})[c]; 804 (b_{17})[c]; 890(y_{19})[c]
29.2	ITC*GQVSSSLAPC*IPY	1-16	1755.0	1753	877 (MH_2^{2+})[c]; 650 (y_5); 553 (y_4); 738(b_{14})[c]

[a] C* indicates a carboxamidomethylcysteine
[b] monoisotopic MW when less than 1000Da
[c] doubly charged ions

Peptides eluting at 23.6 and 29.2 min were particularly diagnostic, since they include the amino acids which differed in the different sequences published: position 6, 9 and 76. By considering the alkylation of cysteines, our data showed that these peptides perfectly matched the sequences ITCGQVSSSLAPCIPY and CGVSIPYK, corresponding respectively to sequence 1-16 and 73-80 of Diaz-Perales et al.[33]. The analogous peptides corresponding to the other published sequences (Table 2.1) were searched in the chromatogram by using the XIC (eXtract Ion Chromatogram) technique, but were not found (data not shown). The tryptic peptides gave an 88% coverage of the entire sequence, thus providing a reliable indication on the protein identity. These data, together with the molecular weight determined, strongly suggested that the natural protein corresponds unequivocally to the sequence proposed by Diaz-Perales et al. and therefore has two serine residues in positions 9 and 76, rather than Ala and His as reported in another paper[17].

ii Exact mass determination

In order to measure the accurate mass of peach Pru p 3 and putative plum LTP, the pure protein was dissolved in CH_3CN:TFA 0.1 % 1:1 and injected in an Orbitrap high resolution mass spectrometer. The monoisotopic mass measured for Pru p 3 was 9129.46, perfectly consistent with the theoretical mass from the sequence Q9LED1 (in Expasy database) of Diaz-Perales[33] and with the presence of 4 disulphide bridges. Figure 2.7 shows the theoretical and the measured isotopic pattern of the hexaprotonated ion of peach Pru p 3. The theoretical and observed spectra matched perfectly. Mean MW obtained by the weighted average of the isotopic peaks is 9136 Da, consistent with previous data.

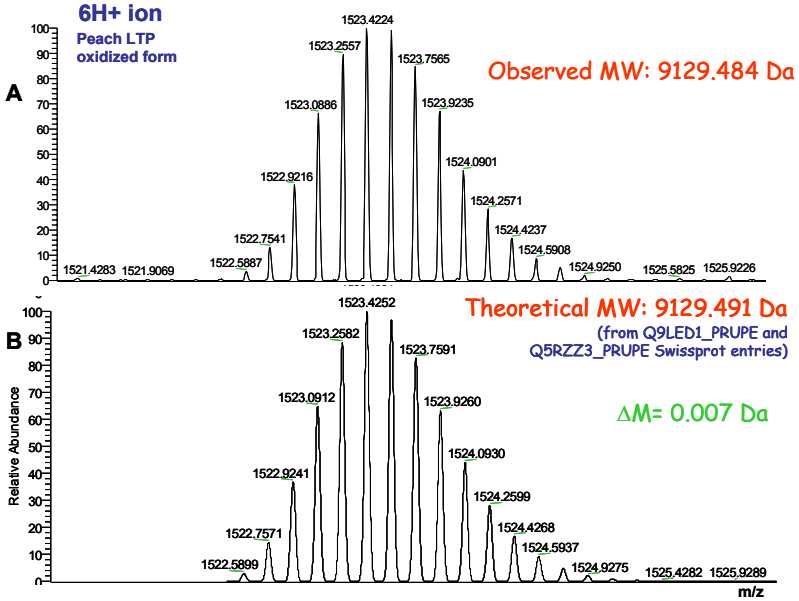


Figure 2.7. (A) Isotopic pattern of purified LTP 6+ ion (B) theoretical isotopic pattern of 6+ ion from Q9LED1 sequence (oxidized form).

iii Number of cysteine residues determination

The presence of the 4 disulphide bridges in peach LTP was confirmed after reduction of the purified protein with 5 mM tributylphosphine, as shown by a mass increment at LTQ-ORBITRAP analyzer of 8.064 Da (or 1.344 m/z) (Figure 2.8), i.e. 8 cysteine residues.

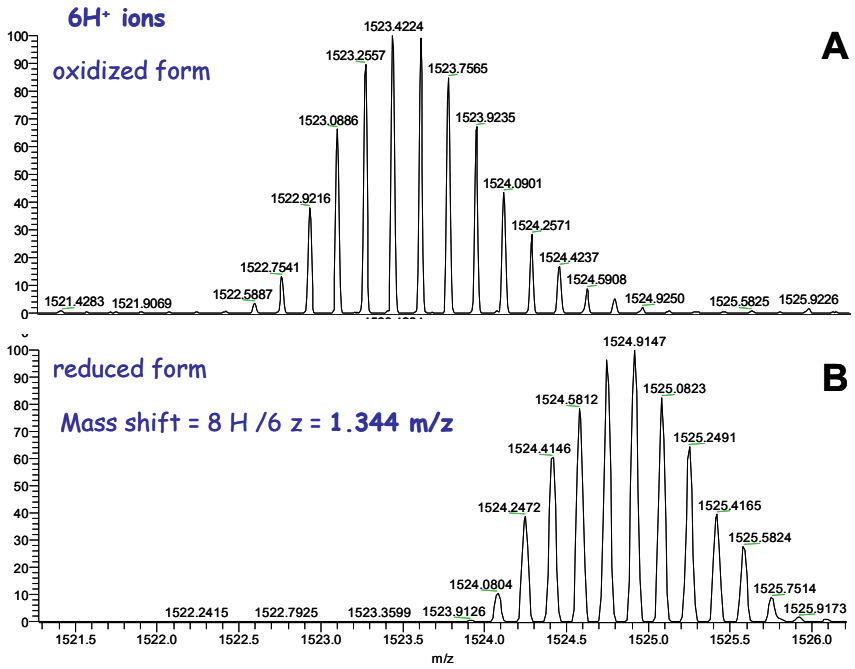
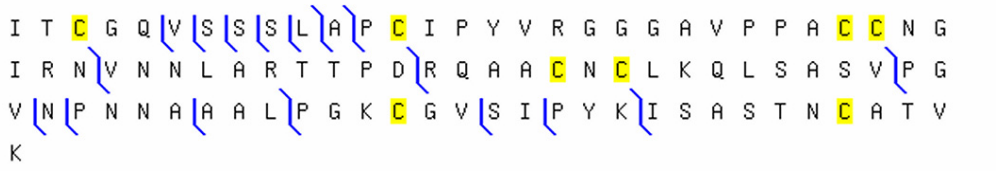


Figure 2.8. Isotopic pattern of purified LTP 6H+ ion: (A) oxidized form; (B) reduced form.

iv Top down approach: MS/MS of the intact protein

The Top Down approach was tried directly on the natural peach LTP but the MS/MS spectrum was very poor, probably due to the very compact structure of the protein due to disulfide bridges, very resistant to fragmentation in the conditions employed. Instead, the MS/MS spectrum of the reduced protein gave enough information to elucidate the identity of amino acids in position 6, 9, 19, and 76, which differ in the various sequences reported in the database. As can be seen in Figure 2.9, ions observed in the MS/MS experiment are consistent with the Diaz-Perales et al.[33] sequence.



Ion	Observed Mass (Da)	Theoretical Mass (Da)	Mass Error (Da)
B10	975.451	975.470	-0.018
B11	1046.488	1046.507	-0.019
B33	3241.604	3241.562	0.042
B43	4323.043	4323.113	-0.070
B58	5895.802	5895.900	-0.098
B69	6914.577	6914.419	0.158
B80	8043.885	8044.015	-0.130
Y11	1093.524	1093.546	-0.021
Y14	1481.730	1481.757	-0.027
Y16	1681.844	1681.873	-0.029
Y22	2223.102	2223.141	-0.039
Y25	2478.254	2478.299	-0.046
Y29	2874.428	2874.475	-0.047
Y30	2988.463	2988.518	-0.055
Y33	3241.604	3241.661	-0.057
Y48	4814.339	4814.447	-0.108
Y58	5895.802	5895.999	-0.197
Y80	8090.899	8091.054	-0.154
Y81	8161.940	8162.091	-0.151
Y82	8275.026	8275.175	-0.149
Y83	8362.049	8362.207	-0.158
Y84	8449.084	8449.239	-0.155
Y85	8536.117	8536.271	-0.154
Y86	8635.181	8635.339	-0.158

Figure 2.9. Graphical and textual representation of fragments observed in Top-Down experiment on 6^+ ion of fully reduced LTP.

The TOP/DOWN approach confirmed the unequivocal sequence of peach LTP with 100% of protein sequence coverage, even better then the tryptic digest.

2.4.2 Plum LTP

2.4.2.1 LTP purification

The purification technique presented for peach LTP was appreciably faster than all the other methods reported in the literature so far, and it was applied also for plum LTP in this thesis work. In this case, the first purification step on a semipreparative column was omitted and purification was carried out only on an analytical column C18. This allowed a faster purification but with lower yields. A good compromise resulted to be the use of an analytical C4 (data not shown). The extraction method was, in this case, slightly modified: fundamentally, to prevent cross linking between phenols and proteins, already observed for phenol-rich fruits[35], skins were

immediately placed under liquid nitrogen and crushed before homogenization. Moreover, the extraction buffer this time contained also EDTA as complexating agent, together with PVPP. The TIC chromatogram this time was more complex (Figure 2.10) than the ones obtained for peach, but this effect could be due to the fact that the semipreparative HPLC was omitted.

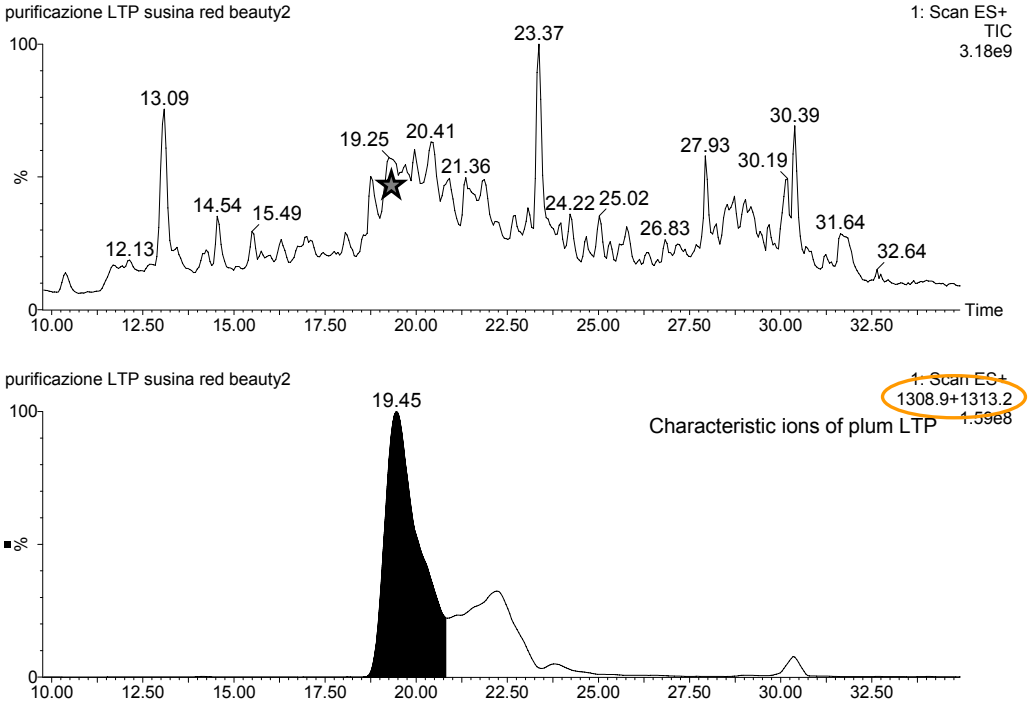


Figure 2.10. TIC chromatogram of plum total protein extract (upper part), and extracted ion chromatogram XIC of putative plum LTP characteristic ions: in black, the shape used to follow the elution of the protein and, consequently, for its purification.

The information found in the literature on Plum LTP MW and sequence were scarce. A chromatographic peak of about 9 kDa, at about the same elution time as peach LTP (with this analytical column, data not shown), was supposed to be the plum LTP. The spectrum is reported Figure 2.11.

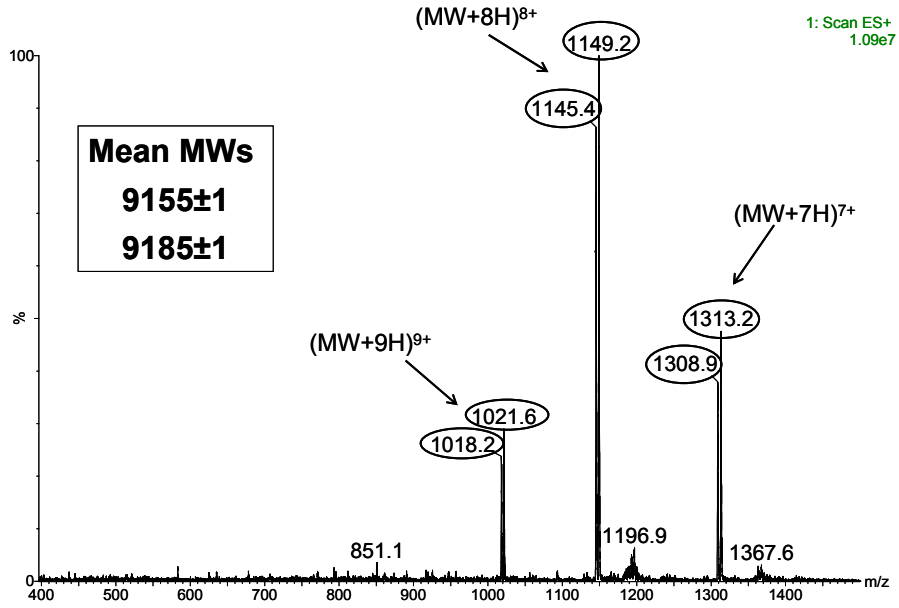


Figure 2.11. MS spectrum of the presumed Plum LTPs.

An interesting thing was observed: two signals were present for every multicharged ion, corresponding to two MWs, 9155 and 9185 Da; these two proteins eluted at the same retention time, as can be seen when extracting the two XIC (in Figure 2.12).

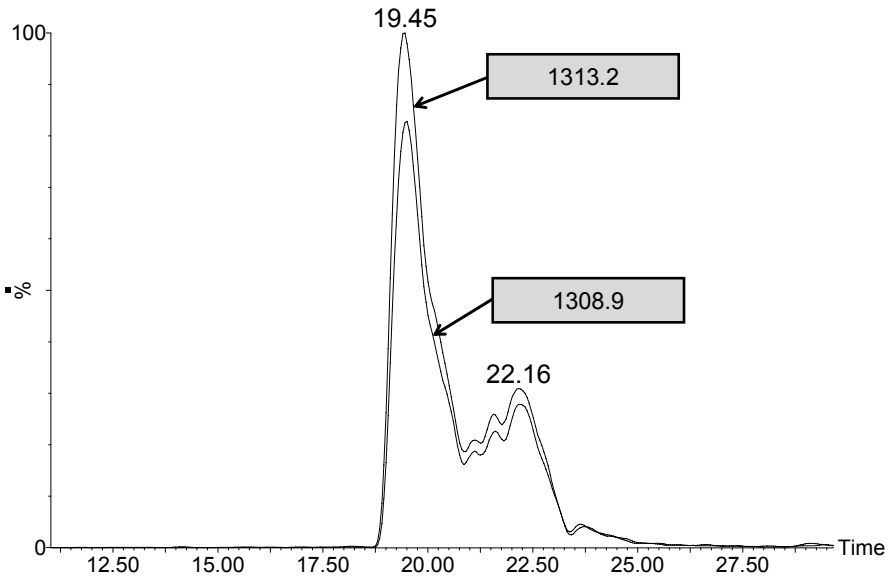


Figure 2.12. Super imposed eXtracted Ions Chromatograms (XIC) of characteristic ions belonging to the two different MWs.

The 30 Da shift of mass was not compatible with expected mass shifts due to common post translational modifications, so we presumed that these two signals were two proteins, presumably LTPs, and that they could be two isoforms, that is, two versions of the same protein with only few amino acid different in the sequence. This variability has already been observed for other LTPs, but never in fruits[2].

Considering different plum varieties, the LTP distribution turned out to be quite complicated, unlike as in peach LTP. In fact, three different varieties of peach always showed the same MW, whereas in plum, by analyzing six varieties, it was found that each variety had two signals at that elution time, variously distributed (Table 2.3).

Table 2.3. summary of ions and determined MW of putative LTP of plum in the several varieties studied.

Plum varieties	Characteristic ions at 19.5 min of retention time and charge state	MW
 **RED BEAUTY**	$(MW+7H)^{7+}$: 1308.9, 1313.2 $(MW+8H)^{8+}$: 1145.4, 1149.2 $(MW+9H)^{9+}$: 1018.2, 1021.6	**MW: 9155, 9184**
 **BLACK**	$(MW+7H)^{7+}$: 1308.95, 1313.36 $(MW+8H)^{8+}$: 1145.5, 1149.21, $(MW+9H)^{9+}$: 1018.34, 1021.68	**MW: 9156, 9186**
 **GOCCIA D'ORO**	$(MW+7H)^{7+}$: 1312.98, 1315.63 $(MW+8H)^{8+}$: 1149.02, 1151.29 $(MW+9H)^{9+}$: 1021.37, 1023.42	**MW: 9184, 9202**
 **REGINA CLAUDIA**	$(MW+7H)^{7+}$: 1305.29, 1306.74 $(MW+8H)^{8+}$: 1142.35, 1143,29 $(MW+9H)^{9+}$: 1016.26, 1016.58	**MW: 9130, 9140**
 **FORTUNA**	$(MW+7H)^{7+}$: 1305.42, 1306.74 $(MW+8H)^{8+}$: 1142.09, 1143.29 $(MW+9H)^{9+}$: 1015.44, 1016.58	**MW: 9130, 9140**
 **AMOLA**	$(MW+7H)^{7+}$: 1309.07, 1315.75 $(MW+8H)^{8+}$: 1145.50, 1151.36 $(MW+9H)^{9+}$: 1018.40, 1023.63	**MW: 9156, 9202**

As can be seen, the different putative LTP isoforms were found several times with a peculiar distribution among varieties, graphically visualized (considering the MW of the different proteins) in Figure 2.13.

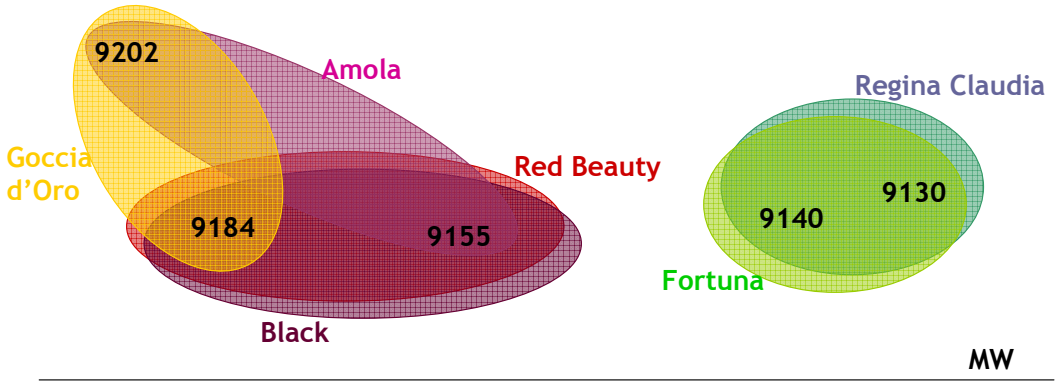


Figure 2.13. MW repetitions between plum varieties studied.

This variability could be due to the hetero-pollinating character of plum[36] and may be defined by the presence of several alleles of the same gene or as families of related genes.

Moreover, these results may have strong implications for the allergenicity of the fruit, since allergenic potential is usually strictly related to the primary sequence. Other allergenic proteins have been discovered with polymorphisms, i.e., a mixture of similar molecules with minor variations in their amino acid sequences. These are called isoallergens or allergenic variants depending on the degree of similarity. Polymorphism is considered to have important effect on the epitopes recognized by T lymphocytes, monoclonal antibodies and IgE of allergic patients. Individual polymorphisms can affect the basal level of allergenicity as well as the cross-reactivity with other allergens. This is the same presumption which has guided different groups to produce hypoallergenic mutants by multiple mutation in the primary structure.[37,38,39,40,41] The use of isoforms with low or total absence of IgE binding capacity but with high capacity to stimulate T cell response has been suggested as an alternative to the conventional immunotherapy for allergic diseases. Standardization of allergenic compounds can be affected by the differing proportions of isoforms in allergenic sources from different regions.[42]

In order to confirm that the observed proteins can be identified as LTPs, it was decided to determine, at least partially, their sequence, by means of purification, exact mass determination, sequence determination. This part of the work was done on the Red Beauty variety.

2.4.2.2 Plum proteins characterization as LTPs

In the Red Beauty variety the two putative isoforms were purified together since they could not be separated by reverse phase HPLC. Purified LTP isoforms were digested with trypsin, reduced with dithiotreitol (DTT) and alkylated with iodoacetamide (IAA) as described in the Experimental part.

i Trypsin digestion

The tryptic peptides deriving from purified plum proteins were reduced and alkylated, and gave a good HPLC-MS (single quadrupole) chromatogram, but very little information was obtained about the peptide sequences, also on account of the fact that no complete sequence of plum LTP is known.

More detailed structural information of tryptic peptides was obtained by a MS/MS experiment performed at CISM (Firenze, Italy). The same digestion mixture was separated in a nano-LC Dionex system and analyzed on line with the LTQ-ORBITRAP spectrometer. The system carried out the selection of multi-charged ions with m/z > 800 (in order to avoid interfering signals) and performed a peptides fragmentation. The ion selection discarded many of the low molecular weight peptides released by the digestion. Daughter ions were analyzed in the high-resolution ORBITRAP analyzer. In Figure 2.14 an example of peptide fragmentation obtained in LTQ-ORBITRAP is reported.

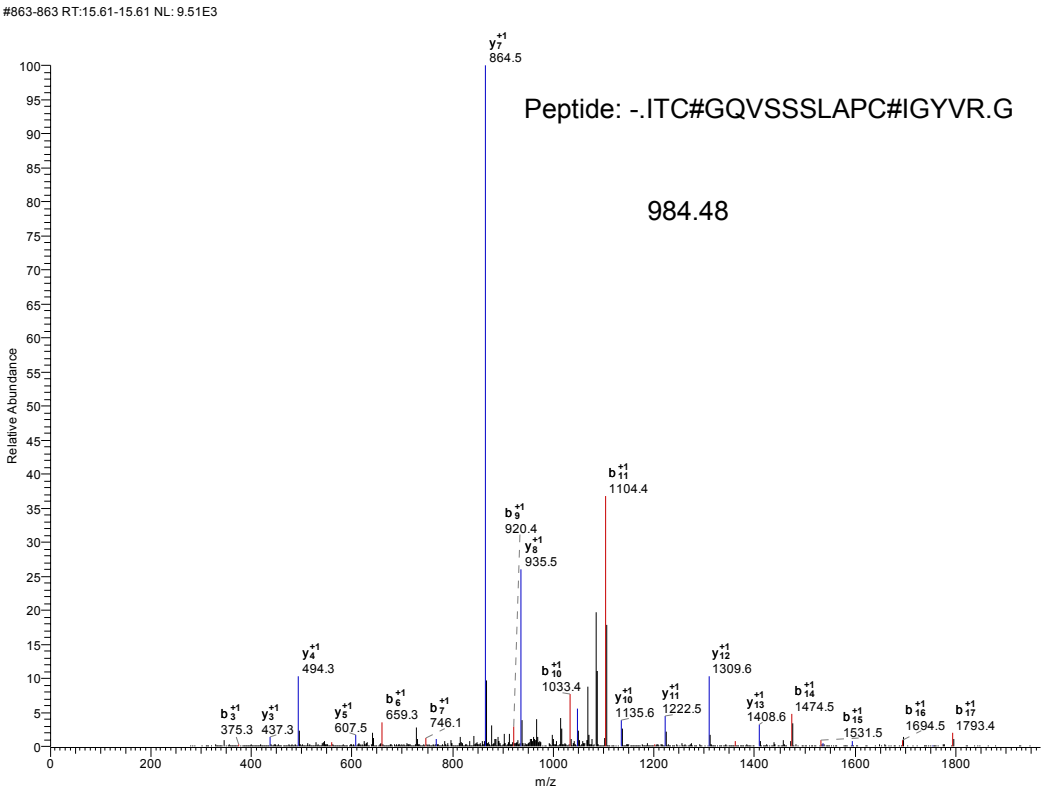


Figure 2.14. Fragmentation of ion 984.48 ($2H^+$) produced from Plum LTP digestion, reduction and alkylation. The sequence can be determined by the b and y ions.

Few large peptides were sequenced and 78% of the protein was covered, also by comparison with known LTP sequences of other fruits. :

ITCGQVSSSL APCIGYVR GGGAVPPACCNGIR XXXXXXXXXXXXXXXXXXX QLSGSISGVNPNNAAALPGK CGVNIPYK ISATTNCSNVK

The sequence resolved by means of LTQ-ORBITRAP showed a high homology degree with other LTPs but noteworthy differences from the available N-terminal sequence of a plum LTP: ITCGQVSSNLAPCINYVKGGGAVP. The different amino acids are marked in red.

No peptides containing the different amino acids were found; so that we think that they might be located in the central part of the protein, still not characterized by this analysis. The work is currently in progress to elucidate the central part of the protein. This part of the sequence should include ,as deduced by other LTPs , some Lys and Arg, typical cleavage sites of trypsin, and so, probably, this part of the protein is extensively proteolysed. Lowering the m/z value of ions selected to be fragmented could result in a too high background and in a decrease in the efficacy of the LTQ-Orbitrap. The next approach will therefore make use of enzymes with different specificities.

ii Exact mass determination

In order to measure the accurate mass of putative plum LTPs, the pure protein was dissolved in CH_3CN:TFA 0.1 % 1:1 and injected in a Orbitrap high resolution mass spectrometer.

The monoisotopic masses measured for plum LTP were determined from the $7H^+$ ions of the two proteins (Figure 2.15). The ions were so resolved that Δ m/z between ions corresponded to $7H^+$ ions (A), and that permitted the deconvolution of the spectrum to isotope pattern of MW (B). With the isotopic distribution of MW, it was possible to calculate the mean MW as a mean of these values and the exact MW, relative to the first isotope (C). When the complete sequence of the protein by the bottom-up approach is elucidated, exact MW would be the perfect check for the sequence.

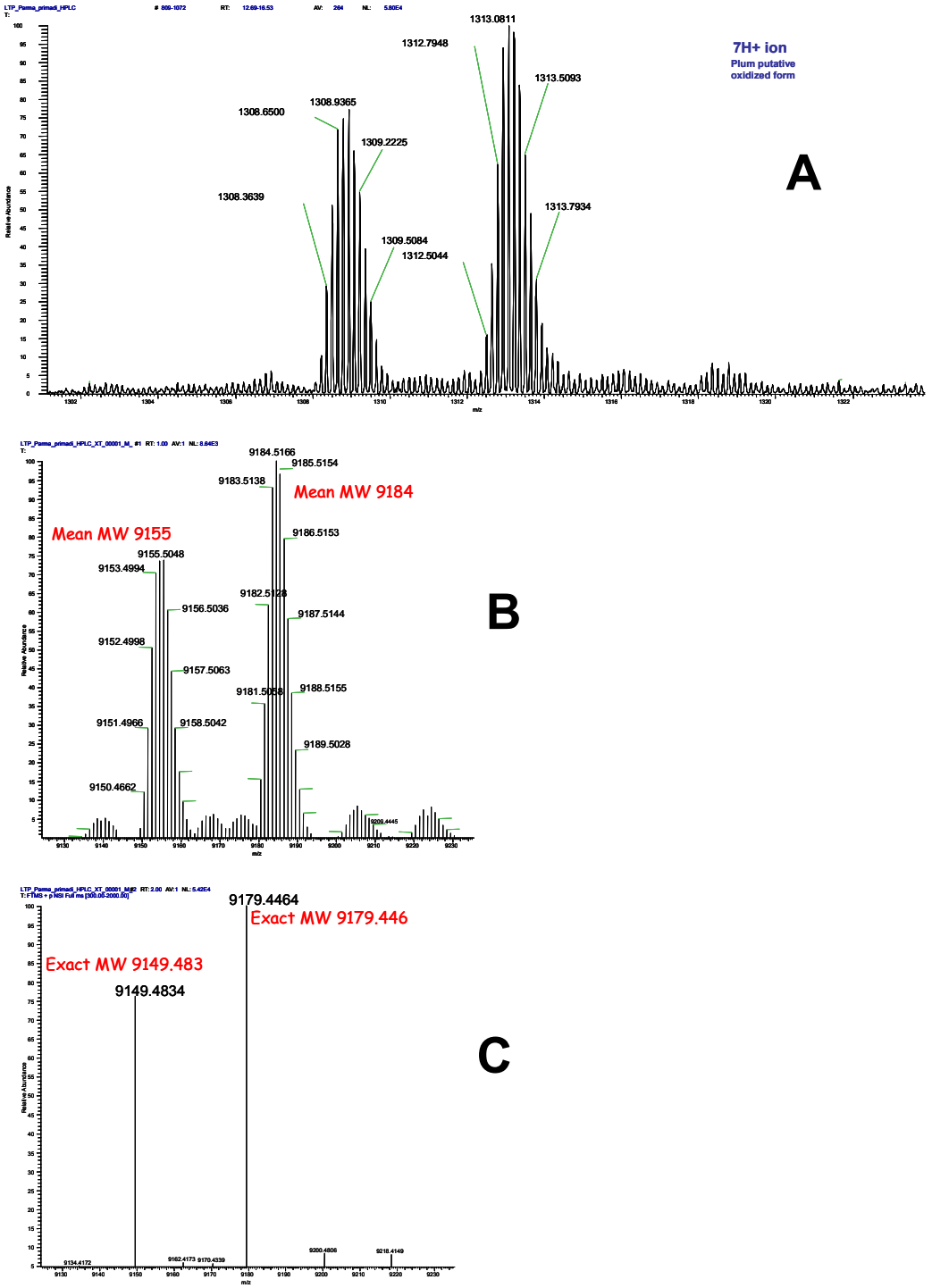


Figure 2.15. (A) Isotopic pattern of purified plum LTP 7+ ion; (B) Deconvoluted spectrum with isotopic pattern of purified plum LTP proteins, with calculated MW indicated; (C) Exact MWs of purified plum LTP proteins, calculable from spectrum (B).

iii Number of cysteine residue determination

The presence of the 4 disulphide bridges in plum LTP was confirmed after reduction of the purified protein. As far as plum LTP is concerned, the treatment of the purified protein with TBP to reduce disulphide bridges increased too much the background signal, thus preventing a direct measurement of the mass shift. Reduction of S-S bridges and carboxamidomethylation of free thiols were therefore performed on the purified protein with DTT and IAA. This time, the spectrum was acquired on the Waters SQ mass spectrometer (Figure 2.16). Since reduction and alkylation did not have 100% of yield, as expected, in the spectrum there was the co-presence of the native and of the carboxamidomethylated forms of LTP. The mass shift confirmed 8 carboxamidomethylation, so 8 cysteines forming 4 disulphide bridges. This evidence, together with the protein (partial) sequence, strengthened the assumption that these two signals were both LTPs.

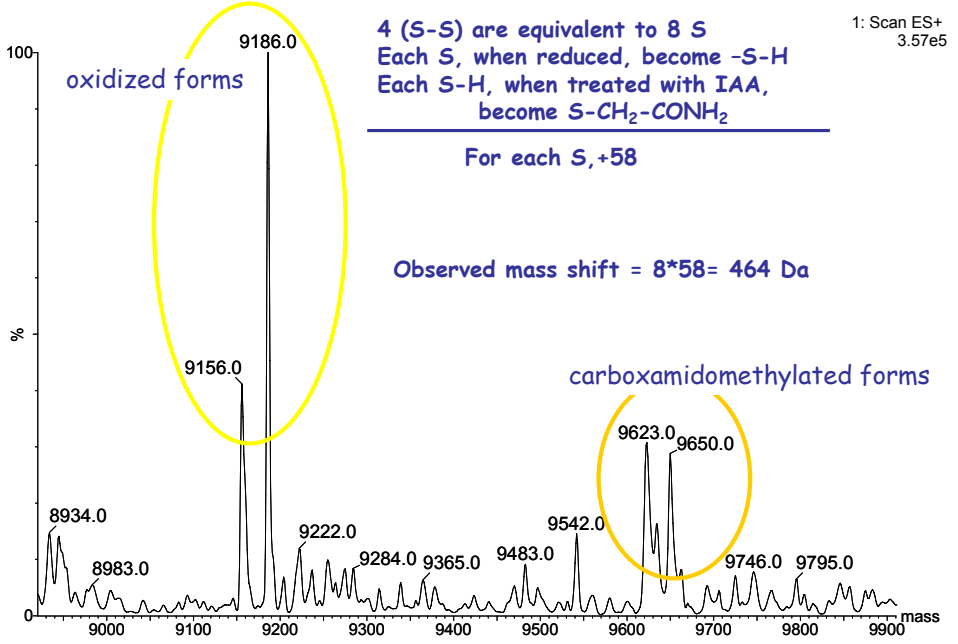


Figure 2.16. Co-presence of oxidized and carboxamidomethylated forms of plum putative LTP in the deconvoluted spectrum acquired with the SQ spectrometer.

2.4.3 Distribution of the LTP proteins in pulp and skin by MALDI Imaging

In order to monitor the LTP distribution in fruit pulp and skin, 250 µm slices were cut from the frozen fruits. After matrix deposition by spraying, spectra were acquired in a MALDI-TOF spectrometer, and for each slice several spectra from the whole slice surface were collected with a spatial resolution of 400 µm. From all the spectra acquired (MALDI Profiling), the signals of interest can be *extracted*, thus ion images of the selected m/z are available, superposed on the fruit section (MALDI Imaging). In Figure 2.17 it is shown how, from a thin section of a peach,

it was possible to obtain the MALDI-TOF profile of the principal protein components in the 6000-10000 Da molecular weight range and to extract signals of ions of interest which could be localized in the section of the fruit.

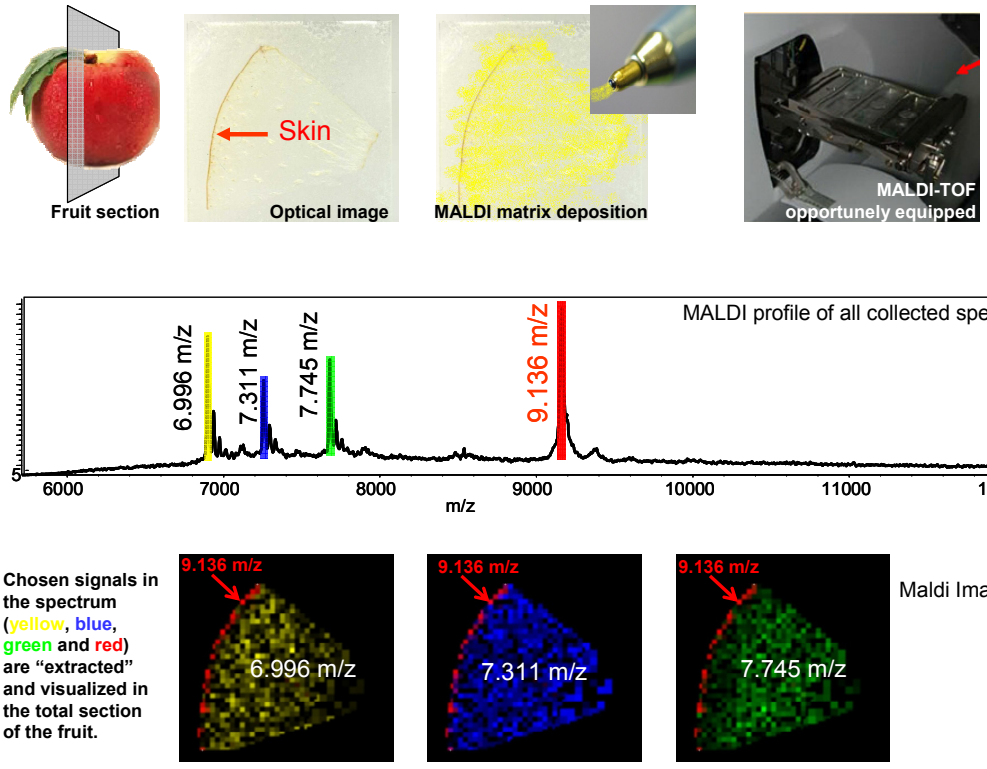


Figure 2.17. Procedure and data of MALDI IMAGING on peach.

The LTP distribution in the peach was easily assessed by MALDI imaging, using the purified protein as standard, since a little de-calibration of the Time-Of-Flight analyzer occurs when a sensible thickness is added before the tube length.

The analysis of acquired data showed that the Pru p 3 signal was present exclusively in the more external layer of the fruit, corresponding to the skin (Figure 2.17). On the other hand, in the pulp no signal belonging to peach LTP was observed, in perfect agreement with the previous HPLC-PDA measurement on pulp and skin extracts and with the literature.

The analysis was repeated also on a thin section of plum. Acquired data for plum LTP showed that both signals were present, each one of them, exclusively, in the more external layer of the fruit, corresponding to the skin. Ions extracted and monitored were relative to a single protein at a time (in Figure 2.18 the signal 9188 m/z is monitored). On the other hand, in the pulp, no signal belonging to plum 9156 or 9188 m/z signals was observed.

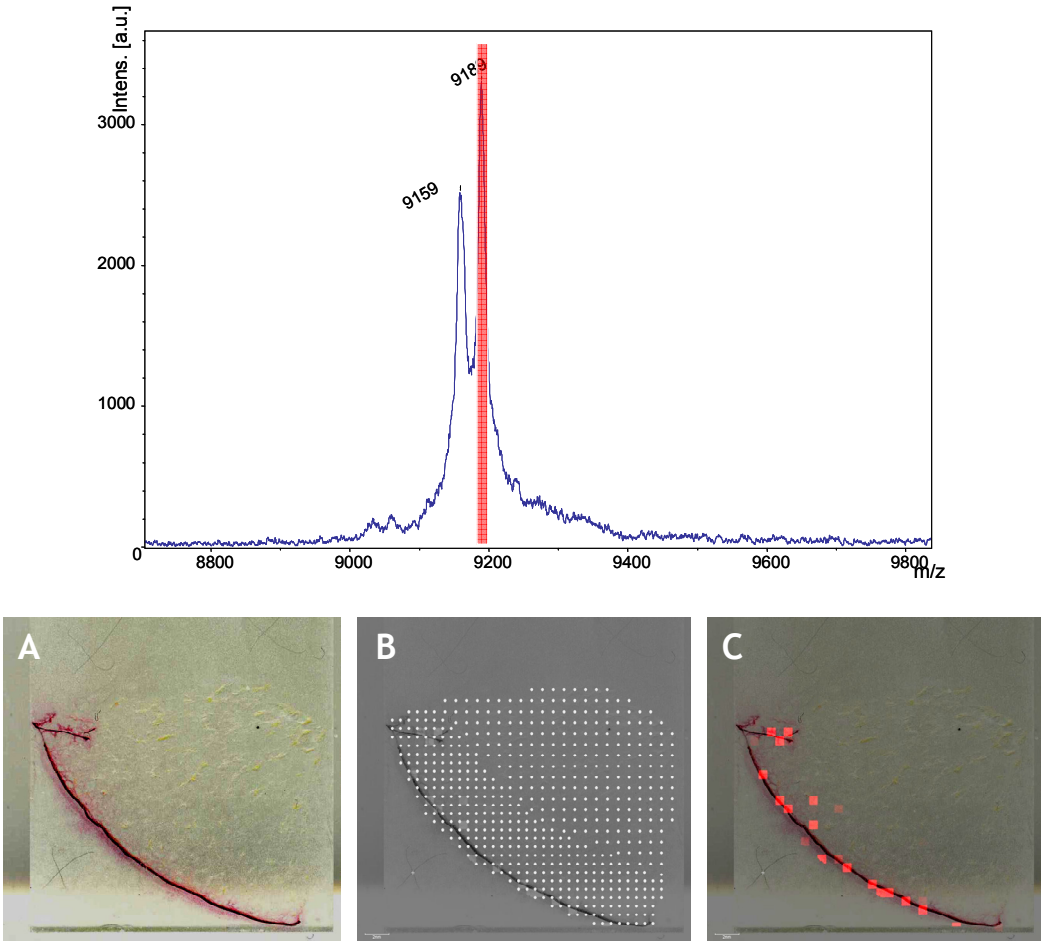


Figure 2.18. Maldi Imaging of plum LTP (second isoform).

Imaging analysis confirmed that LTP is concentrated in the outer part of the fruit, and this definitely confirms also the identity of the two proteins of plum as LTPs.

2.5 Conclusions

The results here presented definitely confirmed the unequivocal sequence of peach LTP, demonstrating at the same time the enormous potential of advanced MS techniques for obtaining rapidly high quality structural and functional data of food-relevant proteins. A preliminar characterization of plum LTPs was also obtained and the co-presence of two isoforms in the skins of the fruit was shown.

In our case, a link could be hypothesized involving the wide diffusion of peach allergy and the stability of the peach LTP primary structure; in addition, the lower frequency of allergy to plums can be correlated to the variability of the corresponding LTP sequences.

2.6 Acknowledgements

LTQ-Orbitrap and MALDI MS Imaging measurements were performed at C.I.S.M., Centro Interdipartimentale di Spettrometria di Massa, Università degli Studi di Firenze, Sesto Fiorentino, Italy. Prof. Gloriano Moneti, Dr. Giuseppe Pieraccini and Dr. Guido Mastrobuoni are gratefully acknowledged for their kind assistance.

2.7 References

1 Kader J.C., .Stimulation of phospholipid exchange between mitochondria and microsomal fractions by proteins isolated from potato tuber, *Biochim. Biophys. Acta* **1977** 380 1975, 31–44.

2 G. Salcedo, R. Sánchez-Monge, D. Barber, A. Diaz-Perales Plant non-specific lipid transfer proteins: An interface between plant defence and human allergy *Biochimica et Biophysica Acta* **2007** 1771, 781–791,

3 Vincent Arondel , Chantal Vergnolle, Catherine Cantrel, Jean-Claude Kader, Lipid transfer proteins are encoded by a small multigene family in Arabidopsis thaliana *Plant Science* **2000** 1571–12

4 F. Vignols, M. Wigger, J.M. Garcia-Garrido, F. Grellet, J.C. Kader, M. Delseny, Rice lipid transfer protein (LTP) genes belong to a complex multigene family and are differentially regulated,*Gene* **1997** 195: 177–186.

5 F. Boutrot, A. Guirao, R. Alary, P. Joudrier, M.F. Gautier, Wheat nonspecific lipid transfer protein genes display a complex pattern of expression in developing seeds, *Biochim. Biophys. Acta* **2005** 1730: 114–125.

[6] J.C. Kader, Lipid-transfer proteins in plants, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1996** 47: 627–654.

7 Andre´ de Oliveira Carvalho, Valdirene Moreira Gomes. Role of plant lipid transfer proteins in plant cell physiology—A concise review *Peptides* **2007** 28: 1144–1153

[8] Françoise Guerbette, Michèle Grosbois, Alain Jolliot-Croquin, Jean-Claude Kader and Alain Zachowski Lipid-transfer proteins from plants: Structure and binding properties *Molecular and Cellular Biochemistry* **1999** 192: 157–161,.

[9] M. Fernândez-Rivas And M. Cuevas Peels of Rosaceae fruits have a higher allergenicity than pulps, *Clinical and Experimental Allergy,* **1999**, 29: 1239-1247,

[10] Rosa Sánchez-Monge, Manuel Lombardero, Francisco Javier García-Sellés, MD,c Domingo Barber, PhD,b and Gabriel Salcedo, Lipid-transfer proteins are relevant allergens in fruit allergy *J Allergy Clin Immunol* **1999**;103:514-9.

[11] Heimo Breiteneder and Clare Mills Nonspecific lipid-transfer proteins in plant foods and pollens: an important allergen class, *Current Opinion in Allergy and Clinical Immunology* **2005**, 5, 3, 275-279

[12] G. Salcedo A. Diâz-Perales R. Sânchez-Monge, Fruit allergy: plant defence proteins as novel potential panallergens, *Clinical and Experimental Allergy,* **1999,** 29: 1158-1160,

[13] A. Diâz-Perales, M. Lombardero, R. Sânchez-Monge, F. J. Garciâ-Selles, M. Pernas, M. Fernândez-Rivas, D. Barber G. Salcedo: Lipid-transfer proteins as potential plant panallergens: cross-reactivity among proteins of Artemisia pollen, Castanea nut and Rosaceae fruits, with different IgE-binding capacities *Clinical and Experimental Allergy*, **2000**, 30: 1403-1410

[14] J. F. Crespo, J. Rodriguez Food allergy in adulthood, *Allergy* **2003:** 58: 98–113,

[15]. Ebner C, Hirschwehr R, Bauer L,Breiteneder H, Valenta R, Ebner H et al. Identification of allergens in fruits and vegetables. IgE-cross reactivity with important birch pollen allergens Bet v1 and Bet v 2 (birch profillin). *J Allergy Clin Immunol* **1995**;95:962–969.

[16] Montserrat Fernandez-Rivas, Ronald van Ree, Manuela Cuevas, J Allergy to Rosaceae fruits without related poliinosis, *Allergy Clin Immunol*, **1997**;100:728-3

[17] E.A. Pastorello, L. Farioli, V. Pravettoni, C. Ortolani, M. Ispano, M.Monza, C. Baroglio, E. Scibola, R. Ansaloni, C. Incorvaia, A. Conti,The major allergen of peach (Prunus persica) is a lipid transfer protein,*J. Allergy Clin. Immunol* **1999**. 103: 520–526.

[18] R. Sanchez-Monge, M. Lombardero, F.J. Garcia-Selles, D. Barber, G. Salcedo, Lipid-transfer proteins are relevant allergens in fruit allergy,*J. Allergy Clin. Immunol*. **1999** 103: 514–519.

[19] E.A. Pastorello, V. Pravettoni, L. Farioli, M. Ispano, D. Fortunato, M. Monza, M.G. Giuffrida, F. Rivolta, E. Scibola, R. Ansaloni, C. Incorvaia, A. Conti, C. Ortolani, Clinical role of a lipid transfer protein that acts as a new apple-specific allergen, *J. Allergy Clin. Immunol*. **1999**104: 1099–1106.

[20] http://www.allergen.org/Allergen.aspx

[21] Pedro Gamboa, MD, Rosa Sanchez-Monge, M. Luisa Sanz, MD, Arantxa Palacı´n, Gabriel Salcedo, Araceli Diaz-Perales, Sensitization to Cannabis sativa caused by a novel allergenic lipid transfer protein Can s 3,*J Allergy Clin Immunol* **2007**, 120, 6, 1459-1460,

[22] G. Salcedo, R. Sanchez-Monge, A. Diaz-Perales, G. Garcia-Casado and D. Barber Plant non-specific lipid transfer proteins as food and pollen allergens, *Clin Exp Allergy* 2004; 34:1336–1341,

[23] Laurian Zuidmeer and Ronald van Ree, Lipid transfer protein allergy: primary food allergy or pollen/food syndrome in some cases, *Curr Opin Allergy Clin Immunol* **2007** 7:269–273.

[24] L. K. Poulsen, In search of a new paradigm: mechanisms of sensitization and elicitation of food allergy, *Allergy* **2005**: 60: 549–558

25 Matthias Besler, Javier Cuesta Herranz, Montserrat Fernández-Rivas. Peach (Prunus persica) *Allergen Data Collection from Internet Symposium on Food Allergens* **2000** 2(4): 185-201

[26] Ferna´ ndez-Rivas M, Gonza´ lez-Mancebo E, Rodrı´ guez-Pe´ rez R, Benito C, Sa´ nchez-Monge R, Salcedo G et al. Clinically relevant peach allergy is related to peach lipid transfer protein, Pru p 3, in the Spanish population. *J Allergy Clin Immunol* **2003**;112:789–795.

[27] R. Asero, G. Mistrello, D. Roncarolo, S.C. de Vries, M.F. Gautier, C.L. Ciurana, E. Verbeek, T. Mohammadi, V. Knul-Brettlova, J.H. Akkerdaas, I. Bulder, R.C. Aalberse, R. van Ree, Lipid transfer protein: a panallergen in plant-derived foods that is highly resistant to pepsin digestion, *Int. Arch. Allergy Immunol*. **2000,** 122 20–32.

[28] Pasquato N, Berni R, Folli C, et al. Crystal structure of peach Pru p 3, the prototypic member of the family of plant nonspecific lipid transfer protein panallergens. *J Mol Biol* **2006;** 356:684–694.

[29] Identification of IgE-binding epitopes of the major peach allergen Pru p 3 Gloria García-Casado, Luis F. Pacios, Araceli Díaz-Perales, Rosa Sánchez-Monge, Manuel Lombardero, Francisco J. García-Selles, Florentino Polo, Domingo Barber, and Gabriel Salcedo, *J Allergy Clin Immunol* **2003**;112:599-605

[30] Llconart R.; Cistero A.; Carreira J.; Batista A.; Moscoso del Prado J.; Food allergy: identification of the major IgE-binding component of peach (Prunus persica). *Annals Of Allergy*, **1992,** 69, 128-130.

[31] Botton A.; Begheldo M.; Rasori A.; Bonghi C.; Tonutti P. Differential Expression of Two Lipid Transfer Protein Genes in Reproductive Organs of Peach (Prunus persica L. Batsch). *Plant Sci.,* **2002**, 163, 993-1000.

[32] Botton A.; Vegro M.; De Franceschi F.; Ramina A.; Gemignani C.; Marcer G.; Pasini G.; Tonutti P. Different expression of Pp-LTP1 and accumulation of Pru p 3 in fruits of two Prunus persica L. Batsch genotypes. *Plant Science.* **2006**, 171, 106-113

[33] Diaz-Perales A.; Garcia-Casado G.; Sanchez-Monge R.; Garcia-Selles F.J.; Barber D.; Salcedo G.; cDNA cloning and heterologous expression of the major allergens from peach and apple belonging to the lipid-transfer protein family. *Clin. Exp. Allergy,* 2002, 32, 87–92.

[34] E.A. Pastorello, L. Farioli, V. Pravettoni, M.G. Giuffrida, C. Ortolani, D.Fortunato, C. Trambaioli, E. Scibola, A.M. Calamari, A.M. Robino, A.Conti, Characterization of the major allergen of plum as a lipid transfer protein, *J. Chromatogr., B*, **2001** 756: 95–103.

[35] Emilia V. Vassilopoulou , Laurian Zuidmeer , Jaap Akkerdaas , Neil Rigby , F. Javier Moreno , Nikolaos G. Papadopoulos , Photini Saxoni-Papageorgiou , Clare Mills , Ronald van Ree Optimized techniques for the extraction of grape allergens appropriate for in vivo and in vitro testing and diagnosis *Molecular Nutrition & Food Research* **2007**, 51, 3 , 360 – 366,

[36] *Functional Plant Genomics,* edited by J. F. Morot-Gaudry, P. Lea &J. F. Briat, Science Publishers: **2007**. 699 pp.

[37] Gerald Reese, Barbara K. Ballmer-Weber, Andrea Wangorsch, Stefanie Randow and Stefan Vieths Allergenicity and antigenicity of wild-type and mutant, monomeric, and dimeric carrot major allergen Dau c 1: Destruction of conformation, not oligomerization, is the roadmap to save allergen vaccines *Journal of Allergy and Clinical Immunology,* **2007** 119, 4, 944-951

[38] Takuya Nakazawa , Toshiro Takai , Hideki Hatanaka , Eri Mizuuchi , Teruyuki Nagamune , Ko Okumura , Hideoki Ogawa Multiple-mutation at a potential ligand-binding region decreased allergenicity of a mite allergen Der f 2 without disrupting global structure. *FEBS Lett.* **2005** 28;579:1988-94

[39] Api m 1 by multiple mutations is concomitant with low IgE reactivity Alteration of the tertiary structure of the major bee venom allergen, Cécile Buhot, Alexandre Chenal, Alain Sanson, Sandra Pouvelle-Moratille, Michael H. Gelb, André Ménez,, Daniel Gillet and Bernard Maillère, *Protein Sci.* **2004** 13: 2970-2978

[40] Mohan B. Singh, Nicole de Weerd, Prem L. Bhalla Genetically Engineered Plant Allergens with Reduced Anaphylactic Activity *Int Arch Allergy Immunol* **1999**;119:75-85,

[41] Mark Larché, Cezmi A. Akdis and Rudolf Valenta Immunological mechanisms of allergen-specific immunotherapy, *Nature Reviews | Immunology*, **2006,**6: 761-771

[42] JD Alché, AJ Castro, JC Jiménez-López, S Morales, A Zafra, AM Hamman-Khalifa, MI Rodríguez-García, Differential Characteristics of Olive Pollen From Different Cultivars: Biological and Clinical Implications, *J Investig Allergol Clin Immunol* **2007**; 17, 1: 69-75

3 SIMULATED IN VITRO GASTROINTESTINAL DIGESTION OF PRU P 3 PEACH ALLERGEN: EVALUATION OF PROTEIN RESISTANCE, IDENTIFICATION OF THE GENERATED PEPTIDES AND ASSESSMENT OF THEIR ALLERGENICITY.

3.1 Introduction

LTP[1] protein structures are compact[2], since there are four disulphide bridges and therefore they are extremely resistant to heat treatments and proteolysis.[3] These characteristics make them true food allergens, since they can sensitize the patient directly in the gastro-intestinal tract[4,1] In particular, ns-LTPs have been shown to survive the acidic environment of the stomach as fully IgE-reactive structures that can trigger mast cells in the vicinity of the epithelial monolayer. In addition, adsorption of proteins is more efficient at this site than at the oral mucosa. These factors explain why allergens such as nsLTPs can cause severe symptoms and reach multiple target organs, giving rise to several effects: OAS, generalized urticaria, asthma, vomiting, diarrhoea, hypotension and even anaphylactic shock.[5]

Resistance to digestion in the gastrointestinal tract is thought to be one of the factors that may contribute to the allergenic potential of food proteins because it allows sufficient intact (or a large fragment of) protein to be taken up by the gut and sensitize the mucosal immune system. [6] In particular, the resistance to pepsin digestion[7] would allow the allergens to maintain their immunogenic and allergenic motifs and thus to interact with the immune system associated with the gastrointestinal epithelia, thereby inducing both sensitization and systemic symptoms observed in allergic patients[8]. Although the pathways of LTP sensitization via the gastrointestinal tract are not yet fully understood[9], resistance to pepsinolysis has been incorporated into the decision tree assessment for potential allergenic risk presented by novel foods.[10]

Simulated gastrointestinal digestion has been carried out for LTPs of grape and barley, but still no data are available on the actual resistance of the peach LTP to gastrointestinal enzymes, or on the peptides formed during the digestion or their potential allergenicity.

3.2 Aim of the work

In the present study, we evaluated in a model system the effects of the gastro-intestinal digestion of the peach LTP Pru p 3 by means of HPLC-MS. We wanted to quantify the extension of the proteolysis in artificially simulated gastro-intestinal conditions, to characterize the peptides produced and to study their IgE-binding capacity.

3.3 Experimental part

3.3.1 Reagents

- PBS buffer
- PVDF membrane (BIO-RAD)
- OPTI-4-CN kit (BIO-RAD)
- TBS buffer (20 mM Tris-HCl, 150 mM NaCl, pH 7.5)
- TBS-T buffer: 0.05% Tween20 in TBS
- Blocking buffer: 1% Ovalbumin in TBS-T

3.3.2 Instruments

- Jupiter Phenomenex Column C18 (300 Å, 250 x 4,6 mm, 5 µm)
- ALLIANCE 2695 HPLC system (WATERS)
- ZMD single quadrupole mass spectrometer (MICROMASS)

3.3.3 Protein extraction and purification

LTP was extracted and purified from the peach (*Prunus Persica*) variety Italia K2, as described in paragraph 2.3 of the previous chapter. Protein concentration was determined according to the Bradford method[11] by using a Perkin Elmer Lambda 20 Bio UV absorbance detector.

3.3.4 Simulated gastrointestinal digestion of peach LTP

11 µl of an LTP solution (5,5 µg/µl) were mixed to 50 µl of 0.1 N HCl solution at pH 2, 2 µl of a pepsin solution (0,5 µg/µl in water), and 11,6 µl of a PNA solution (2,5 µg/µl in water). We used as internal standard a 15-mer Peptide Nucleic Acid (MW 4238 Da) which is resistant to proteases. The solution was maintained for 3 h at 37 °C at pH 2 to simulate gastric transit. Then, 40 µl of a 50 mM ammonium bicarbonate buffer, 1µl of trypsin solution (1 µg/µl in 50 mM acetic acid) and 2 µl of a α-chymotrypsin solution (0,5 µg/µl in 50 mM acetic acid) were added. The mixture was maintained for 4 h at 37°C and pH 7.8 to simulate intestinal transit. A control was also performed, substituting enzymes with equal volumes of buffers.

The peptide mixture was analyzed by HPLC-ESI-MS. The ESI-MS spectrometer conditions were the following: positive ion mode, capillary voltage 3 kV, cone voltage 30 V, source temperature 100°C, desolvation temperature 150 °C, spraying gas (N_2) 100 l/h, desolvation gas (N_2) 400 l/h, full scan acquisition from 150 to 1500 m/z in continuum mode and 2.9 sec of scan time.). The flow rate used was 1ml/min, eluent A was constituted by 0.1 % HCOOH + 0.2 % CH_3CN in Milli Q H_2O, eluent B by 0.1 % HCOOH + 0.2 % Milli Q H_2O in CH_3CN. Gradient: 5 min of eluent A isocratic, 35 min of linear gradient to 40 % of eluent B, 10 min of isocratic 40 % B, 10 min of non linear gradient to 100 % B and reconditioning. The injection volume was 30 µl.

Peptides were characterized by means of "Expasy Peptide Cutter" software simulations, molecular weights and spectral fragments obtained by in source collision.

The percentage of undigested protein was calculated correlating the signal of the protein to the signal of PNA in the deconvoluted spectrum obtained by heights proportional to the entire signal (one height representing the protein, the other representing the PNA) by means of Mass Lynx 4.0 software.

3.3.5 Purification of peptides obtained by simulated GI digestion

Peptides formed during the simulated gastro-intestinal digestion were purified exploiting the same XIC technique described for the purification of the intact protein (paragraph 2.3 of this session of Results). By monitoring the characteristic ions of peptides formed in the simulated gastro intestinal digestion, it was possible to recover them from the T-split of the flux just before the MS spectrometer. The peptides purified were dried under an N_2 flux and stored frozen.

3.3.6 Dot blotting

Peptides were redissolved in 5 µl of PBS buffer and spotted on activated PVDF membrane for dot-blotting analysis.

Each peptide was spotted on the pre-activated membrane (see paragraph 1.3.7.3); the membrane was washed in TBS-T for 5 minutes, incubated with blocking buffer for two hours, washed twice in TBS-T for 5 min, incubated with sera of allergic patients (1:50 in blocking buffer) for 2 hours, washed twice in TBS-T for 5 min, incubated with antibodies Rabbit-anti-human IgE (1:3000 in TBS-T) for 1 hour, washed twice in TBS-T for 5 min, incubated with antibodies Goat-anti-rabbit-HRP (1:3000 in TBS-T) for 2 hours. The complex formed was detected by the OPTI-4-CN detection kit, incubated for 30 min. The reaction was stopped by washing with Milli-Q water for 15 min.

3.4 Results and discussions

3.4.1 Identification of peptides resulting from digestion

The major proteolytic events taking place during the gastro-intestinal digestion were taken into consideration in this work: pepsin action during gastric digestion and trypsin and chymotrypsin action during intestinal transit. Times of incubation were chosen according to the model of Vermeirssen[12].

Different peptides were released by digestion of ns-LTP Pru p 3 and were analyzed by RP-HPLC-ESI-MS. The chromatogram is reported in Figure 3.1. Peptides were completely identified by database search[13] and are listed in Table 3.1. A Peptide Nucleic Acid (PNA), an oligonucleotide analogue resistant to the protease action[14], was added to the mixture and used as internal standard in order to determine the amount of undigested protein.

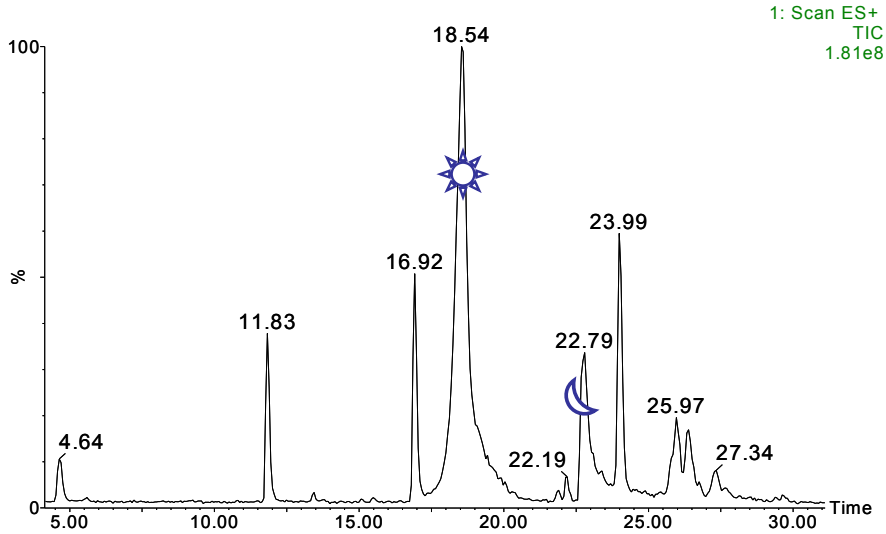


Figure 3.1. HPLC-MS chromatogram of peptides produced by simulated gastro-intestinal digestion of purified peach LTP. ☼ is the PNA signal and ☾ is the LTP.

Table 3.1. Identified peptides generated from peach LTP Pru p 3

Rt (min)	MW (Da)	Attributed fragments	sequence	MS spectra characteristic ions
4,70	273	**VR**	17-18	274.2 (MH^+) 175.1 (y_1)
11,83	588	**TTPDR**	40-44	589.4 MH^+ , 488.4 y_4 , 387.3 y_3 , 290.2y_2 , 175.1y_1
13,46	472	**NLAR**	36-39	473.4 MH^+ , 359.4y_3 , 246.2 y_2
16,92	799	**NVNNLAR**	33-39	800.5 MH^+ , 400.9 MH_2^{2+} 587.4 y_5 , 473.3 y_4 , 246.2 y_2
22,49	1663	**SASVPGVNPNNAAALPGK**	55-72	832.7 MH_2^{2+} , 555.0 MH_3^{3+} 660.5 y_{14} $2H^{2+}$
22,79	9135	**Entire protein(*)**	1-91	1306.0 $MH7^{7+}$, 1142.9 MH_8^{8+} 1016.3 MH_9^{9+}
23,99	1904	**QLSASVPGVNPNNAAAPGK**	53-72	953.1 MH_2^{2+} , 636.1MH_3^{3+}
25,97	5712	**HMW peptide 1**	(discussed in text)	1143.4 MH_5^{5+} , 953.0 MH_6^{6+} , 817.1 MH_7^{7+}
26,33	5585	**HMW peptide 2** **5712-K_{91}**	(discussed in text)	1117.9 MH_5^{5+} , 931.8 MH_6^{6+} , 798.8 MH_7^{7+}
27,34	5386	**HMW peptide 3** **5585-$T_{89}V_{90}$**	(discussed in text)	1078.1 MH_5^{5+} , 898.8 MH_6^{6+} 770.5 MH_7^{7+}
(*) MW corresponding to ExPASy entry P81402				

The low MW peptides generated during the digestion are shown in Figure 3.2.

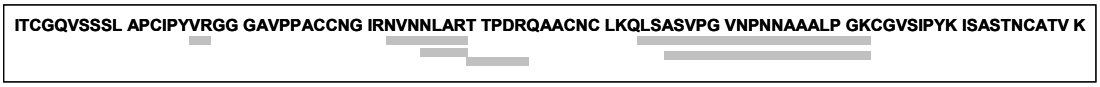


Figure 3.2. In grey, LMW peptides released after simulated GI digestion from peach LTP Pru p 3.

Beside the small peptides released by proteolysis and the intact undigested LTP protein, three high molecular weight peptides were found and, on the basis of MS analysis, were identified as the core of the LTP protein after release of the small peptides: the MW 5712 Da peptide exactly corresponded to the full protein without the released peptides (grey residues in Figure 3.2) also cleaved between K_{80} and I_{81}, still linked together by the 4 disulphide bridges. Its primary structure is represented in Figure 3.3.

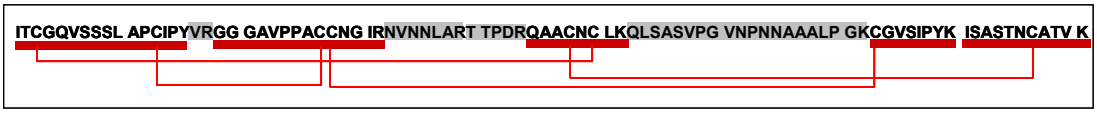


Figure 3.3. In grey, released LMW peptides. In red, the HMW peptide 5712.

The MW 5712 implies that the cleavage between Lys_{80} and Ile_{81} takes place even if the disulfide bridges do not allow the release of the corresponding peptides. The MW 5584 Da peptide is consistent with peptide 5712 without Lys_{91} and the MW 5384 Da corresponded to peptide 5584 without Tyr_{89} and Val_{90}.

In general it appears that LTP can be degraded by trypsin, since almost all its cleavage sites have been, at least partially, hydrolyzed. As far as the other enzymes are concerned, the potential cleavage sites defined by the Peptide Cutter tool of ExPASy are reported in Figure 3.4.

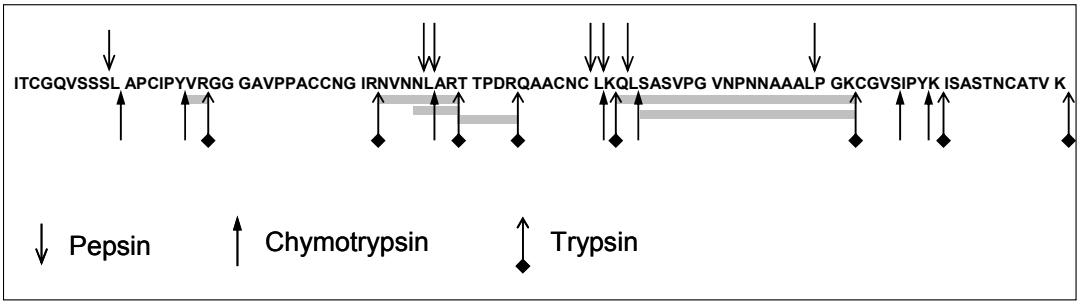


Figure 3.4. Observed peptides released by simulated GI digestion (in grey) compared to possible cleavage sites of the three enzymes used (different arrows).

These data confirmed, as already known, that Pru p 3 is resistant to pepsin[7,15], since none of its specific cleavages was observed, even after a prolonged treatment (3 hours). By performing a simulated digestion experiment only with pepsin no significant fragmentation took place. (Data not shown).

Peach LTP proved to be very resistant also to α-chymotrypsin, since only one, possibly two, cleavages of the 7 possible for this enzyme were observed: a specific cleavage between Leu_{54}-

Ser_{55} and a second cleavage between Tyr_{16} and Val_{17}, although the latter had already been noticed when peach LTP was digested only with trypsin.

On the contrary, Pru p 3 was highly digested by trypsin, in all its specific cleavage sites. This was confirmed by all LMW HMW peptides. Moreover, together with these specific peptides, the aspecific cleavage site Lys_{16}-Val_{17}, also observed in this case, had already been described and another two (between Ala_{88}-Thr_{89} and between Val_{90}-Lys_{91}) were also attributed to aspecific activity of trypsin.

We also evaluated the possibility to evaluate the relative percentage of the protein which was not digested. As appears in Figure 3.1, part of the protein (☾) was still intact. In order to evaluate the amount of protein not hydrolyzed, a comparison with the standard PNA, which is resistant to proteolytic enzymes[14], was made by measuring the corresponding areas before and after the digestion: in this way we could calculate that 35% of the peach allergen Pru p 3 remained intact. In Figure 3.5 the two chromatograms of the resulting mixture with or without enzymes are reported.

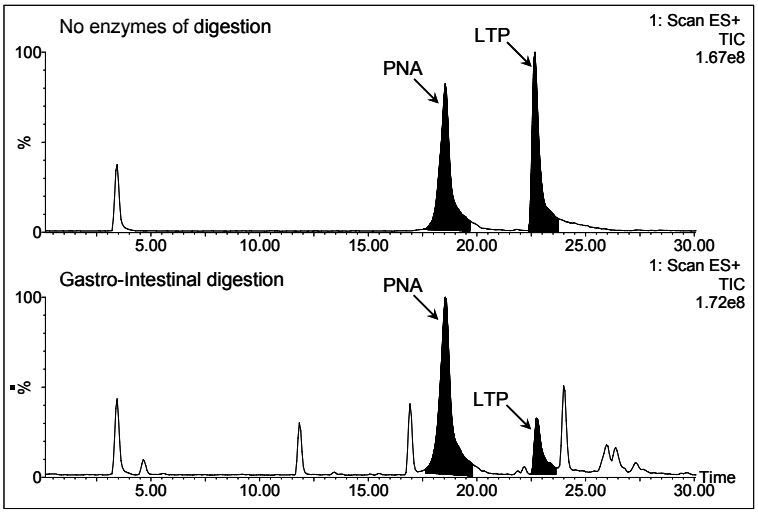


Figure 3.5. Chromatograms obtained in the absence and in the presence of enzymes. PNA is not recognized by the enzymes so its area can be used as internal standard to quantify the changes of the LTP area.

This observation confirmed that a considerable percentage of the intact protein can reach the intestinal mucosa, in a probably immunologically active structure. Moreover, even after trypsin digestion and the release of small peptides, the core of the protein still held together by the disuphide bridges probably conserved many secondary structures of the intact protein, and it could consequently be an immunologically active species. We also studied the IgE reactivity of small peptides.

3.4.2 Allergenicity of the purified peptides from simulated gastrointestinal digestion (Immunodot experiments)

Among the methods which can help in the determination of epitopes[16], we chose a Dot-Blotting protocol, and we spotted the purified peptides on a PVDF membrane, blotting them specifically with sera of allergic patients and recognizing the IgE-peptide complex (if present) by specific staining with a secondary antibody and enzymatic reaction. First, the peptides were purified by HPLC coupled to a single quadrupole mass spectrometer and recovered by collecting the flux into vials with high accuracy when the characteristic ions were identified by the MS detector. For this purpose, a T-split, necessary to adapt the HPLC flux of 1 ml/min to a flux compatible with the spectrometer (0.1 – 0.2 ml/min), was used. In this way all the main products of the digestion from 55 µg of protein were recovered singularly, lyophilized and redissolved in the opportune TBS (Tris Buffer Saline) buffer, suitable for spotting on the PVDF membrane. The three HMW peptides were collected all together due to their very close elution times. The PVDF membrane has a high specificity towards proteins, so proteins and peptides adhere to it very strongly. In order to prevent the aspecific binding of serum protein to the membrane, the still free active sites of PVDF were protected with Ovalbumin before incubation with sera. During incubation with sera of allergic patients, IgEs bound specifically to their epitopes. Residual serum was washed away and antibodies Anti-human IgE were added. These antibodies bound with high affinity to IgE. These complexes were successively recognized by specific antibodies, which were in turn linked to an enzymatic system, catalyzing a colorimetric reaction. Therefore, where the recognition peptide-IgE had occurred, a final colored spot was observed, suggesting that peptide could be considered as allergenic. In Figure 3.6 the resulting strips of the PVDF membrane is presented.

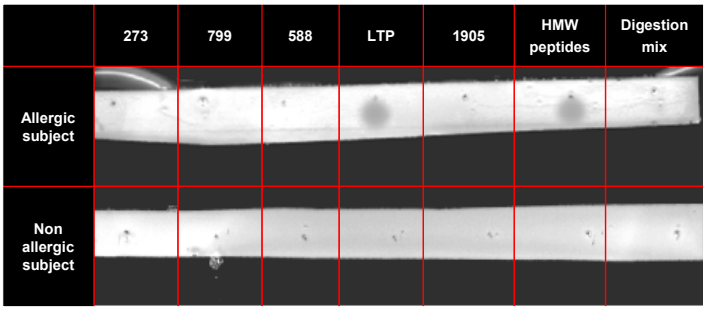


Figure 3.6. IgE immunodetection with a serum pool from patients allergic to peach of a SPOT with purified peptides.

It is immediately evident that only the entire protein and the HMW peptides were recognized and revealed by the system.

The major reactive IgE-binding epitopes characterized up to now in the Pru p 3 allergen[1729] are located in sequences 11-20, 31-40, 71-80. according to these results, sequence 11-20 is cleaved by gastro-intestinal digestion on position 16 and 18, and sequence 31-40 is also extensively

hydrolyzed in the present conditions. Thus, these two sequences can be supposed to be effective epitopes only in the intact protein: they do not appear to be recognized by the IgEs after released by digestion. On the other hand, full LTP and HMW peptides (probably maintaining the 3-dimensional structure) were easily recognized by the patient sera, These results might suggest that LTP allergenicity could be decreased if more extensive proteolysis could take place.

3.5 Conclusions and perspectives

Up to now, no simulation of gastro-intestinal digestion on peach LTP has been reported in the literature. The resistance of this protein to pepsin was evaluated only by considering its stability in electrophoresis. [18]. Our results confirmed this resistance and estimated the percentage of the undigested protein to be around 35 %.

A rapid on line purification of the digestion products and a dot-blotting detection of allergenic peptides were performed, allowing the identification for the first time of the peptides and their IgE-binding characteristics.

According to our results, two of the epitopes reported in literature for peach LTP are located in highly digestible regions of the protein: this probably implies interferences during sensitization or elicitation phases of the allergenic response. These preliminary results are quite stimulating and deserve further studies.

3.6 References

[1] Elide A. Pastorello , Anna M. Robino Clinical role of lipid transfer proteins in food allergy *Molecular Nutrition & Food Research*, **2004**, 48, 5, 356 - 362

2 Salcedo G, Sanchez-Monge R, Diaz-Perales A, Garcia-Casado G, Barber D, Plant non specific lipid transfer protein as food and pollen allergens *Clinical And Experimental Allergy* **2004** 34, 9: 1336-1341

3 G. R. Burnett, M. Wickham, A. Fillery-Travis, J. A. Robertson, P. S. Belton, S. M. Gilbert, A. S. Tatham, P. R. Shewry and E. N. C. Mills , Interaction between protein allergens and model gastric emulsions *Biochemical Society Transactions* **2002** 30, 6,

4 L. K. Poulsen In search of a new paradigm: mechanisms of sensitization and elicitation of food allergy, *Allergy* 2005: 60: 549–558,

5 P. Brandtzaeg, Current Understanding of Gastrointestinal Immunoregulation and Its Relation to Food Allergy *Ann. N.Y. Acad. Sci.* **2002** 964: 13–45

6 F. Javier Moreno Gastrointestinal digestion of food allergens: Effect on their allergenicity, *Biomedicine & Pharmacotherapy* **2007** 61 50-60,

7 Asero R, Mistrello G, Roncarolo D, de Vries SC, Gautier MF, Ciurana CL, et al. Lipid transfer protein: a pan-allergen in plant-derived foods that is highly resistant to pepsin digestion. *Int Arch Allergy Immunol* **2000**;122:20-32.

8 Scheurer S, Lauer I, Foetisch K, Moncin MSM, Retzek M, Hartz C, et al. Strong allergenicity of Pru av 3, the lipid transfer protein from cherry, is related to high stability against thermal processing and digestion. *J Allergy Clin Immunol* **2004**;114:900.

[9] E. N. Clare Mills And John A. Jenkins, Marcos J. C. Alcocer , Peter R. Shewry Structural, Biological, and Evolutionary Relationships of Plant Food Allergens Sensitizing via the Gastrointestinal Tract, *Critical Reviews in Food Science and Nutrition*, **2004,** 44:379–407

10 FAO/WHO, 2001, Evaluation of Allergenicity of genetically Modified Foods. *Report of a Joint FAO/WHO Expert Consultation of Allergenicity of Foods derived from Biotechnology,* **2001**,

[11] Bradford M.M.; A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem., **1976**, 72, 248-254.

12 Vermeirssen V. John Van Camp, Liesbet Devos, Willy Verstraete, Release of Angiotensin I Converting Enzyme (ACE) Inhibitory Activity during in Vitro Gastrointestinal Digestion: from Batch Experiment to Semicontinuous *Model, J. Agric. Food Chem*., **2003,** 51, 5680-5687

[13] http://us.expasy.org/tools/peptidecutter/

[14] Nielsen, P. E., *Peptide Nucleic Acids: Protocols and Applications /2nd Edn*, New Books Only. US Edition.2004

[15] Riccardo Asero Gianni Mistrello Daniela Roncarolo Sacco C. de Vries, Marie-Frangoise Gautier, Caroline L.F. Ciurana, EddyVerbeek, Tamimount Mohammadi, Vlasta Knul-Brettlova Jaap H. Akkerdaas Ingrid Bulder, Rob C. Aalberse, Ronald van Ree, Lipid Transfer Protein: a pan-allergen in plant-derived foods that is highly resistant to pepsin digestion *Int Arch Allergy Immunol* **2001**,124:67-69

[16] Gerald Reese, Rosalia Ayuso, Susan M. Leong-Kee, Matthew J. Plante, Samuel B. Lehrer, Characterization and identification of allergen epitopes: recombinant peptide libraries and synthetic, overlapping peptides *Journal of Chromatography B*, **2001,** 756: 157–163

[17] Gloria García-Casado, Luis F. Pacios, Araceli Díaz-Perales, Rosa Sánchez-Monge, Manuel Lombardero, Francisco J. García-Selles, Florentino Polo, Domingo Barber, and Gabriel Salcedo, Identification of IgE-binding epitopes of the major peach allergen Pru p 3, *J Allergy Clin Immunol* **2003**;112:599-605

[18] R. Asero, G. Mistrello, D. Roncarolo, S.C. de Vries, M.F. Gautier, C.L. Ciurana, E. Verbeek, T. Mohammadi, V. Knul-Brettlova, J.H. Akkerdaas, I. Bulder, R.C. Aalberse, R. van Ree, Lipid transfer protein: a panallergen in plant-derived foods that is highly resistant to pepsin digestion*, Int. Arch. Allergy Immunol*. **2000** 122 20–32.

Results, part 2:

Proteolytic peptides for cheese quality assessment

4 OLIGOPEPTIDES IN PARMIGIANO-REGGIANO CHEESE: MOLECULAR MARKERS OF TIPICALITY, TECHNOLOGY AND AGEING

4.1 Introduction

The conversion of fresh into mature cheese is largely influenced by proteolysis, although other factors, such as lipolysis and lactic and propionic acid fermentation are especially important for particular types of cheeses. Proteolysis is the most complex event in cheese ripening as it contributes to textural changes (breaking down the αS 1, αS 2, ß and para-k casein network), it decreases the water activity through water binding by setting free carboxyl and amino groups, it increases the pH by producing NH_3 from deamination of free amino acids. Moreover, proteolysis directly contribute to flavour (release of peptides and amino acids) and off-flavours (bitter hydrophobic peptides), also liberating substrates for others reactions. Thus, for the development of an acceptable cheese flavor, a well-balanced breakdown of the protein (i.e., casein) into small peptides and amino acids is necessary.[1]

Proteolysis in cheese during ripening is catalyzed by the residual activity of the rennet enzymes (chymosin, pepsin or acid proteases), natural milk proteases (plasmin, cathepsins and somatic cells proteases), and proteolytic enzymes from starter lactic acid bacteria, secondary inoculum and non-starter microflora.[2] Different enzymes with different specificities belong to the general class of "proteases": enzymes specific towards proteins are called *proteinases*, enzymes specific towards oligopeptides are so called *peptidases*. A common feature of peptidases is their inability to hydrolyze intact casein but they have the ability to hydrolyze internal peptide bonds of casein-derived peptides. Peptidases are distinguished between *endo-* or *exopeptidases* if they degrade internal bonds of the peptide or only terminal amino acids. Enzymes of rennet and milk are essentially proteinases and they are known to be highly specific towards caseins. LABs have a more complex enzymatic system which involves: a cell wall-associated caseinolytic proteinases which can hydrolyse intact caseins of high MW peptides formed by, for example, chymosin; extracellular peptidases, specific towards oligopeptides; amino acid and peptide transport systems; intracellular peptidases. All these enzymes together form the complex proteolytic system of lactococci (Figure 4.1). The progress of proteolysis in many ripened cheeses is first catalyzed by residual coagulant (chymosin, which has an hydrophilic character, has more affinity for the whey fraction, but it can also be partially trapped in the caseins network) and, to a lesser extent, by plasmin and milk proteinases, resulting in formation of large and intermediate-sized peptides which are subsequently degraded by the coagulant and enzymes from the starter and non starter microflora of the cheese, resulting in the production of small peptides and free amino acids.

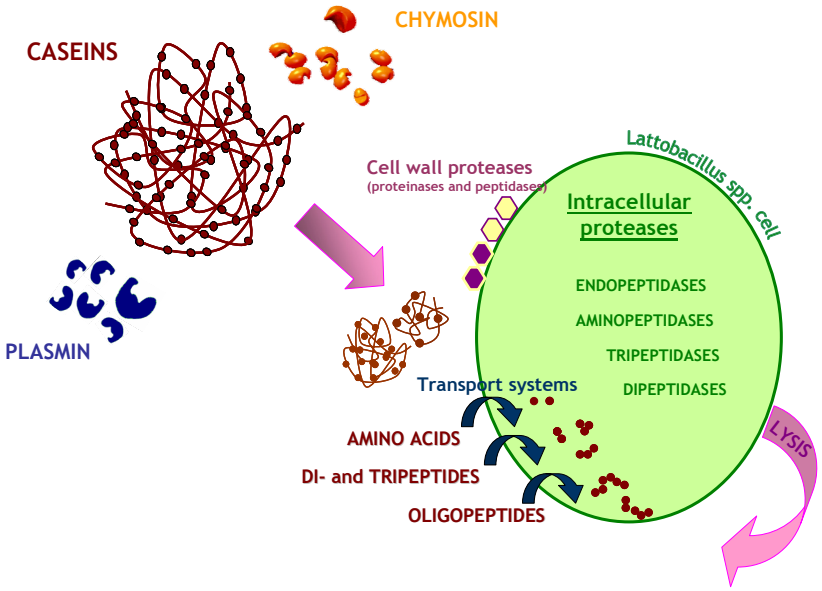


Figure 4.1. The main proteolytic enzymes involved in casein degradation during cheese ripening (adapted from ref.2).

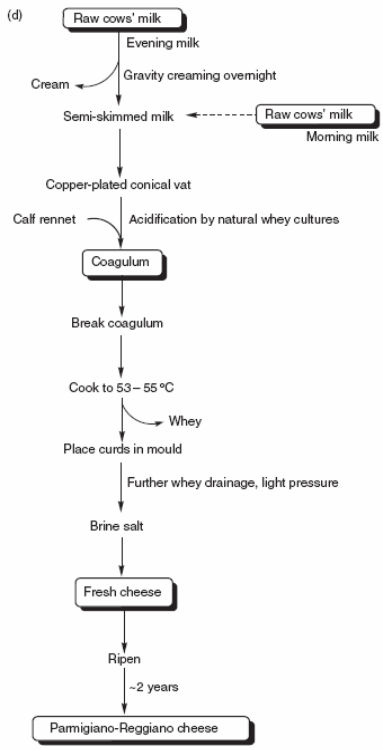


Figure 4.2. Flow diagram for the manufacture of Parmigiano-Reggiano cheese.

Parmigiano-Reggiano is a well-known Italian hard cheese, long ripened, made from raw and partly skimmed cow's milk. It is included in the list of Italian cheeses bearing the Protected Designation of Origin (PDO, EU regulation 2081/92). This definition includes technological characteristics and geographic restrictions. For example, the production of milk and its transformation into cheese can take place only in the provinces of Parma, Reggio Emilia, Modena (and in some parts of those of Bologna and Mantova). Cows have to be fed only on grass or hay and are reared following specific breeding techniques.

Obviously, also the production of cheese from raw milk must follow severe rules. The main stages of the production of Parmigiano-Reggiano cheese are as follows: the milk collected in the evening is partly skimmed by gravity all night long and it is then mixed to the whole milk of the morning after. Together with calf rennet, only a natural whey culture is allowed to be used as starter and it must be obtained from the

previous day cheesemaking whey incubated at a gradually decreasing temperature. The mixture of the raw partly skimmed cow's milk and natural whey starter is heated up to 54-56°C in copper vats and rennet is added to start curding. The curd is then poured into wheel-shaped forms, salted in brine for up to one month and aged for at least one year. The entire caseification process is reported in Figure 1.2. It takes about 550 liters of milk to make one wheel of Parmigiano-Reggiano cheese, which weighs from 32 to 38 kg.

Several works have been published characterizing proteolytic peptides in cheeses (see ref. 2 for a recent review).

Moreover, several data about the cleavage specificities of the main proteolytic enzymes are available (a summary is reported in Table 4.1), most of them a good part of them determined on model systems. Peptides are the primary results of these enzymatic activities, thus studying their presence in cheese, information on the enzymes involved can also be determined.

Table 4.1. The principal dairy proteases cleavage sites in cows' milk caseins.

	β casein	αS 1 casein	αS 2 casein	K casein
Milk Plasmin	Lys28-Lys29, Lys105-His106, Lys107-Glu108 secondary cleavages due to plasmin specificity to Lys-X and in a lesser extent Arg-X peptide bond.	Lys102-Lys103, Lys103-Tyr104, Lys105-Val106, Arg151-Gln152	Lys21-Gln22, Lys24-Asn25, Arg114-Asn115, Lys149-Lys150, Lys150-Thr151, Lys181-Thr182, Lys188-Ala189, Lys197-Thr198	
Curd Chymosin	Leu163-Ser164, Leu127-Thr128	Phe23-Phe24, Phe24-Val25, Phe32-Gly33, Leu40-Ser41, Leu149-Phe150, Phe153-Tyr154, Leu156-Asp157, Ala158-Tyr159, Phe179-Ser180, Trp164-Tyr165, Leu101-Lys102	Tyr89-Gln90, Thr120-Pro121, Asn159-Arg160, Phe88-Tyr89, Tyr95-Leu96, Gln97-Tyr98, Tyr98-Leu99, Phe163-Leu164, Phe174-Ala175Tyr179-Leu180	Phe105-Met106
LAB Proteinases CE	many possible cleavages depending on the strains			
Lab Endo Peptidases	PepO, PepF, PepE, PepG, acting expecially on α S 1 CN f(1-23), α S 1 CN f(91-100), α S 1 CN f(165-199) and β CN f(193-209)			
LAB Esopeptidases*	Aminopeptidases (PepN, PepC, PepS, PepA, PepL) Proline-specific peptidases (PepQ, PepI, PepR, PepX, PepP)			
Milk Elastase	Ile26-Asn27, Gln40-Thr41, Ile49-His50, Phe52-Ala53, Gln56-Ser57, Leu58-Val59, Asn68-Ser69, Val82-Val83, Val95-Ser96, Ser96-Lys97, Lys97-Val98, Ala101-Met102, Glu108-Met109, Phe119-Thr120, Glu131-Asn132, Leu163-Ser164, Ala189-Phe190, Phe190-Leu191 And Pro204-Phe205	Tyr159-Pro160, Phe179-Ser180, Phe24-Val25, Glu148-Leu149, Phe150-Arg151, Tyr154-Gln155, Tyr165-Tyr166, Ala163-Trp164, Tyr166-Val167, Val15-Leu16, Glu30-Val31, Tyr104-Lys105 And Ala143-Tyr144, Phe24-Val25, Tyr104-Lys105, Phe150-Arg151, Tyr154-Gln155, Val15-Leu16, Phe24-Val25, Glu148-Leu149, Ala163-Trp164, Thr171-Gln172,Pro197-Leu198		
Cathepsin D	Phe52-Ala53, Leu58-Val59, Pro81-Val82, Ser96-Lys97, Leu125-Thr126, Leu127-Thr128, Trp143-Met144, Phe157-Pro158, Ser161-Val162,Leu165-Ser166, Leu191-Leu192, Leu192-Tyr193, Phe205-Pro206	Phe23-Phe24, Phe24-Val25, Leu98-Leu99, And Leu149-Phe150	Leu99-Tyr100, Leu123-Asn124, Leu180-Lys181 And Thr182-Val183	

* Kirsi Savijoki . Hanne Ingmer . Pekka Varmanen, Proteolytic systems of lactic acid bacteria, Appl Microbiol Biotechnol (2006) 71: 394-406
§ the peptidase database http://merops.sanger.ac.uk/index.htm

Quite interestingly, LAB peptidases and proteinases specificities strongly depend on the bacterial species and strains. For example, *Lactobacillus helveticus* enzymatic pool includes aminopeptidase N (PepN), aminopeptidase C (PepC), X-prolyl dipeptidyl aminopeptidase (PepX), proline iminopeptidase (PepI), prolinase (PepR), dipeptidase (PepD), cell envelope-associated proteinase (PrtH) (which hydrolyzes Leu16-Asn17 and Asn17-Glu18 bonds of αs1-CN), oligopeptidases (PepO and PepE). *Streptococcus thermophilus* enzymatic pool includes Cell Envelope Proteinase (PrtS), glutamyl aminopeptidase (PepA), aminopeptidases (PepC and PepN), Oligoendopeptidase (PepO), aromatic and arginine aminopeptidase (PepS) and X-prolyl dipeptidyl aminopeptidase (PepX) and in its genes Oligoendopeptidase (PepF), Aminopeptidase

(PepM, PepP), XAA-Pro aminopeptidase, Prolidase (PepQ), Tripeptidase (PepT) and Dipeptidase (PepV) are encoded[3,4]. Information on carboxypeptidases from LAB is scarce, although carboxypeptidase activity of *L. helveticus* and *S.s thermophilus* has been observed on β-CN.[5]

4.2 Aim of the work

In order to improve the knowledge of the action of the different proteolytic systems involved in Parmigiano-Reggiano cheese maturation and to identify oligopeptides to be used as markers of quality and ageing, we evaluated the oligopeptide fraction of samples of Parmigiano-Reggiano. These samples were taken from wheels of Parmigiano-Reggiano cheese, produced in the same day, in the same factory, from the same milk, during the same caseification process, aged from 0 to 24 months. Also caseification milk was available. These homogeneous wheels offered a unique sampling plan to study and model the enzymatic events happening in Parmigiano-Reggiano during ageing.

4.3 Experimental part

4.3.1 Solvents and reagents

- HCl 0.1N (NORMEX)
- (L,L)-phenylalanylphenylalanine (Phe-Phe) (SIGMA-ALDRICH)
- Ethyl-ether (RIEDEL DE HAËN)
- Paper filters
- 0.45µm HVLP filters (MILLIPORE)
- PLGC filters, nominal molecular cut-off 10 000 Da (MILLIPORE)
- Milli Q H_2O
- Acetonitrile CH_3CN, HPLC grade (VWR)
- Methanol CH_3OH, HPLC grade (VWR)
- HCOOH, 99% (ACROS)

4.3.2 Instrumentation

- Ultra Turrax T50 basic (IKA-WERKE)
- ALC 4237R centrifuge
- 46 mm Sterifil Aseptic system (MILLIPORE)
- Alliance 2695 separation module (WATERS)
- Jupiter Phenomenex C18 column (5µm, 300 Å, 250x4.6 mm)
- Millipore PLGC filters (nominal molecular cut-off 10 000 Da)
- Amicon Micropartition system MPS-1 (Millipore).
- SQ mass spectrometer (Waters)

- Rotavapor Water bath 461, (BÜCHI)
- pH meter, 691 (METRÖHM)
- Bench scales (ORMA)

4.3.3 Samples

Cheese samples were obtained from the "Consorzio del Parmigiano-Reggiano". They were taken from the curd, 12 and 48 hours after production, immediately after brining (1 month of ageing) and after 2, 3, 4, 6, 8, 10, 12, 16, 20 and 24 months, from November 2005 to November 2007. Each time, the wheels were sampled from a horizontal section in three different zones: internal (I) and external (E). At each sampling time the cheese was grated and kept at -20°C until the analyses were performed (usually few days).

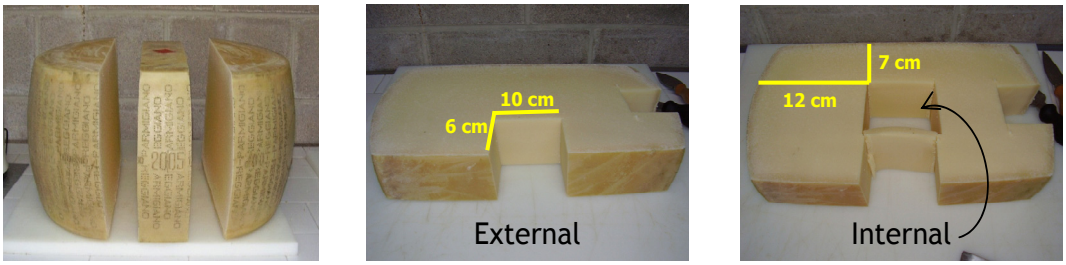


Figure 4.3. Sampling of the Parmigiano-Reggiano wheel.

4.3.4 Extraction and concentration of the oligopeptide fraction

10 g of finely grated cheese were suspended in 45 ml of 0.1 N HCl. (L,L)-phenylalanylphenylalanine (Phe-Phe) was added as an internal standard (2.5ml of a 1mM solution). The suspension was homogenized for 1 min and then centrifuged at 3400 *g* for 40 min at 4°C. The solution was filtered through paper filters and then extracted three times with 40 ml of ethyl ether. The aqueous solution was filtered through 0.45 µm filter. A total of 3 ml of the resulting solution were dried and redissolved in 900µl of a 0.1% formic acid solution (pH 3). The solution was diafiltered through filters with nominal molecular cut-off 10 000 Da at 2800 *g*. The filtrate was dried under nitrogen, redissolved in 500 µl of H_2O (0.1% HCOOH) and analyzed by HPLC/MS.Each sample was extracted at least twice.

4.3.5 LC/MS analysis of the oligopeptide fraction

HPLC-ESI-MS was carried on the same HPLC system but two elution conditions were used. First, a long HPLC was utilized to allow a perfect separation and attribution of ions to each peptide. Then, a more efficient and shorter run was developed, perfectly fitting all kind of samples. Eluents were in both cases as follows. Eluent A: H_2O (0.2 % CH_3CN and 0.1 % HCOOH); eluent B: CH_3CN (0.2 % H_2O and 0.1 % HCOOH). For the first type of analysis ("Long" running) elution was achieved in in the following conditions: 0-15 min isocratic 100% A, 15-60 min linear gradient from 100% A to 67% A, 60-69 min isocratic 67% A, 69-70 min from 67% to 60% A, 70-80 min from

60%A to 0%A, 80-85 min isocratic 0% A, and reconditioning. In the second type of HPLC analysis ("Short" running), elution was achieved by: 0-10 min isocratic 95 % A; 10-50 min linear gradient from 95 % A to 60 % A; 50-56 min linear gradient from 60 % A to 0 % A; 56-62 isocratic 0 % A (100 % B); 62-64 min linear gradient from 0 % A to 95 %A, and reconditioning. Flow rate was always 1 ml/min. MS conditions: ESI interface with 90% splitting of the column flow, positive ions, single quadrupole SQ Waters analyzer. Capillary voltage 3.2kV, cone voltage 30 V, source temperature 100 °C, desolvation temperature 150 °C, cone gas (N_2) 100 l/h, desolvation gas (N_2) 450 l/h. Acquisition in total ion mode (TIC, 100-1900 m/z); a scan time of 3 sec was used in the first type of analysis and of 1 sec in the second case.

4.3.6 Data analysis

LC/MS analysis of the extracted fractions were performed (full scan acquisition). Each HPLC-MS chromatogram was elaborated determining characteristic ions, molecular weights, retention times and in source CID fragments (if present) of the main peaks. MW determination was confirmed by Max Ent application of Mass Lynx Software in case of MWs superior to 1000 Da. Identification of peptides was possible by applying a procedure previously developed in our research group[6], exploiting in source CID fragments (as in the example reported in Figure 4.4).

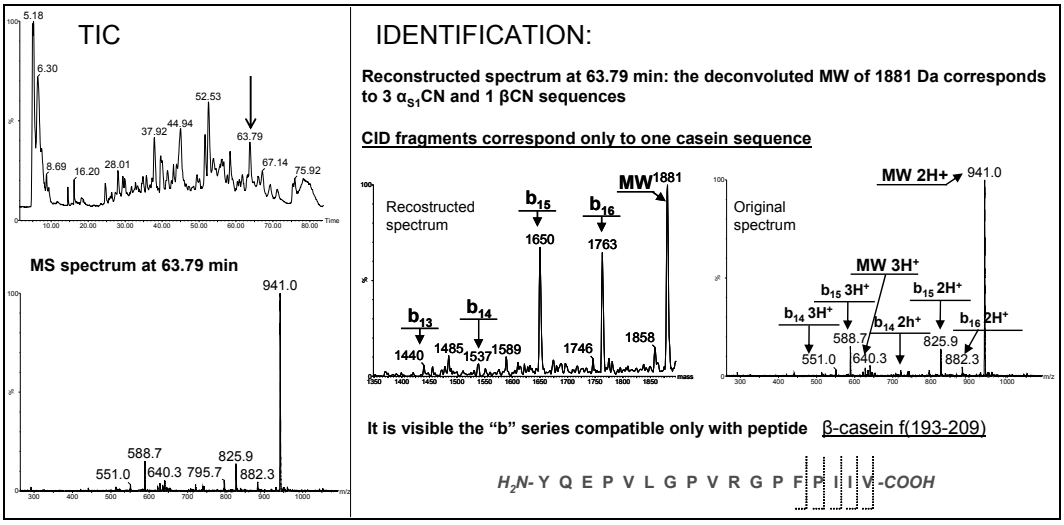


Figure 4.4. HPLC study and sequence identification.

Integration of the area of each peptide was performed in an automatic way by QuanLynx Software, after extraction of the peptide characteristic ions current (XIC technique). Integration parameters include: automatic ApexTrack Peak Integration and Noise measurement; smooth parameters: ± 1 scans and 1 smooth by "mean" method.

The integrated area of each peptide in each sample were collected, and semiquantified by dividing for the Phe-Phe (Internal Standard) area in the same sample. A moisture correction is

performed [corrected value = semiquantified value * (100/(100-W)], based on moisture data of each sample (Table 4.2).

Table 4.2. Moisture (W) values for the cheese samples.

% W	extracted curd	curd, 12h	curd, 48h	1 month	2 months	3 months	4 months
external	46.20	42.60	40.10	37.14	36.45	35.72	35.43
internal	46.20	41.20	41.20	40.08	39.30	38.83	38.68

% W	6 months	8 months	10 months	12 months	16 months	20 months	24 months
external	35.32	34.84	34.31	34.06	32.93	31.45	30.14
internal	37.86	36.84	36.70	35.67	34.98	34.24	33.35

4.3.7 Statistical analysis

Semiquantitative data were entered into a Microsoft Excel spreadsheet and analyzed by using the SPSS 15.0 software. Principal Component Analysis was performed by factorial data reduction extracting 2 factors, with an unrotated factor solution in the space of the observed variables. The first and second PCA components were calculated as independent variables on data by linear regression analysis.

4.4 Results and discussion

First of all, a detailed analysis of the main peptides present in each LC-MS chromatogram was performed. Chromatograms were very different according to the different ageing times of the cheese, as it is evident in Figure 4.5.

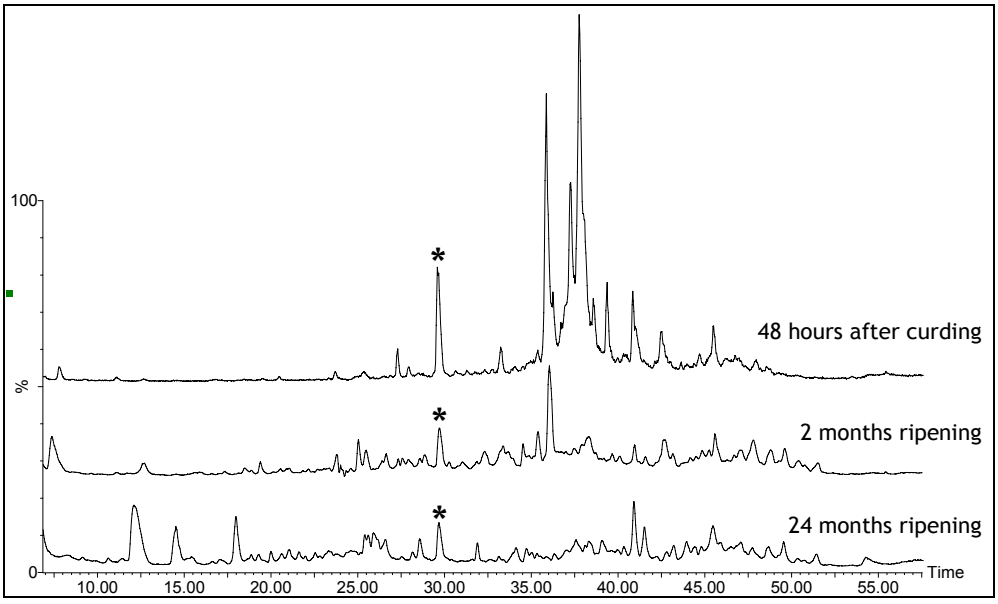


Figure 4.5. Overlaid TIC chromatograms (step = 25%) of three different cheese samples: 48 h after curding, after 2 months and 24 months of ripening. (*) Internal standard, Phe-Phe. "Short" running conditions.

MWs and sequence information from the fragment ions combined with database searching on casein sequences, led to the identifications of 82 peptides, among the most present in the various samples (reported in Table 4.3).

Table 4.3. Identified peptides in Parmigiano-Reggiano cheese.

CODE	identification	MW	Characteristic ions
p1	Lac-Val ●	**189**	144.1 (i_1) + 190.2 (MH^+)
p2	Lac-Ile ●	**203**	158.2 (i_1) + 204.2 (MH^+)
p3	Lac-Leu ●	**203**	158.2 (i_1) + 204.2 (MH^+)
p4	Lac-Met ●	**221**	222.1 (MH^+) + 176.1 (i_1)
p5	Lac-Phe ●	**237**	238.1 (MH^+) + 192.1 (i_1) + 120 (i_2)
p6	Pyr-Ile ●	**242**	243.2 (MH^+) + 197.1 (i_1) + 132 (y_1)
p7	Pyr-Leu ●	**242**	243.2 (MH^+) + 197.1 (i_1) + 132 (y_1)
p8	γ Glu-Val ●	**246**	247.2 (MH^+) + 118.1 (y_1)
p9	Lac-Tyr ●	**253**	254.2 (MH^+) + 208.2 (i_1) + 136.1 (i_2)
p10	γ-Glu-Ile ●	**260**	261.2 (MH^+) + 132.1 (y_1)
p11	γ-Glu-Leu ●	**260**	261.2 (MH^+) + 132.1 (y_1)
p12	Pyr-Phe ●	**276**	277.2 (MH^+) + 231.1 (i_1) + 166(y_1) + 120 (i_2)
p13	γ-Glu-Met ●	**278**	279.3 (MH^+) + 150.1 (y_1)
p14	γ-Glu-Phe ●	**294**	295.2 (MH^+) + 166 (y_1)+ 120 (i_2)
p15	γ-Glu-Tyr ●	**310**	311.0 (MH^+) + 182.0 (y_1)
p16	αS1 CN f (10-13)	**414**	414.3 (MH^+) + 244.1 (y_2)
p17	ß CN f (1-3)	**416**	417.3 (MH^+)
p18	αS1 CN f (1-4)	**536**	537.3 (MH^+) + 383.3 (b_3)
p19	αS1 CN f (10-14)	**542**	543.3 (MH^+) + 373.3 (y_3)
p20	ß CN f (1-4)	**545**	546.3 (MH^+)
p21	αS1 CN f (16-20)	**601**	602.4 (MH^+)
p22	αS1 CN f (30-35)	**707**	708.4 (MH^+)
p23	αS1 CN f (186-192)	**734**	734.4 (MH^+)
p24	αS1 CN f (1-6)	**745**	746.4 (MH^+) + 373.8 (MH_2^{2+})
p25	αS1 CN f (10-16)	**754**	755.3 (MH^+)
p26	ß CN f (47-52)	**755**	756.3 (MH^+) + 378.7 (MH_2^{2+})
p27	ß CN f (1-6)	**787**	788.3 (MH^+) + 657.3 (b_5) + 528.8 (b_4)
p28	αS1 CN f (18-23)	**791**	791.3 (MH^+) + 322.3 (y_2)
p29	αS1 CN f (24-30)	**805**	806.2 (MH^+)
p30	αS1 CN f (1-7)	**874**	438.3 (MH_2^{2+}) + 875.4 (MH^+)
p31	αS1 CN f (17-23)	**905**	905.4 (MH^+) + 453.3 (MH_2^{2+}) +791.5 (y_6)
p32	ß CN f (60-68)	**1001**	501.4 (MH_2^{2+}) + 334.8 (MH_3^{3+})
p34	ß CN f (111-119)	**1122**	1123.2 (MH^+) + 562.3 (MH_2^{2+})
p35	α S2 CN f (154-162)	**1132**	566.8 (MH_2^{2+})

CODE	identification	MW	Characteristic ions
p36	αS1 CN f (115-123)1P	**1138**	1139.1 (MH^+) + 570.3 (MH_2^{2+})
p37	αS1 CN f (1-9)	**1140**	570.8 (MH_2^{2+}) + 381.0 (MH_3^{3+})
p38	ß CN f (199-209)	**1151**	576.4 (MH_2^{2+}) + 1151.6 (MH^+)
p39	ß CN f(17-25)2P	**1154**	578.5 (MH_2^{2+}) + 1155.3 (MH^+)
p40	αS1 CN f (169-179)	**1198**	300.2 (MH_4^{4+}) + 400.3 (MH_3^{3+}) + 600.3 (MH_2^{2+})
p41	αS1 CN f (24-34)	**1237**	1237.3 (MH^+) + 619.3 (MH_2^{2+})
p42	αS1 CN f (14-23)	**1246**	1246.3 (MH^+) + 623.8 (MH_2^{2+})
p43	ß CN f(16-25)3P	**1348**	1348.7 (MH^+) + 674.9 (MH_2^{2+})
p44	ß CN f (71-83)	**1403**	702.9 (MH_2^{2+}) + 1404.1 (MH^+)
p45	αS1 CN f (24-36)	**1495**	748.5 (MH_2^{2+}) + 499.3 (MH_3^{3+}) + 430.4 (MH_3^{3+})
p46	ß CN f (82-95)	**1510**	755.9 (MH_2^{2+}) + 504.2 (MH_3^{3+})
p47	αS1 CN f (1-13)	**1535**	768.4 (MH_2^{2+}) + 512.7 (MH_3^{3+})
p48	ß CN f (195-209)	**1589**	795.5 (MH_2^{2+}) + 1589.6 (MH^+)
p49	ß CN f(17-28) 3P	**1590**	795.9 (MH_2^{2+}) + 465.7 (MH_3^{3+})
p50	αS1 CN f (1-14)	**1664**	832.9 (MH_2^{2+}) + 555.8 (MH_3^{3+}) + 431.7 (MH_4^{4+})
p51	ß CN f (16-28)3P	**1703**	852.4 (MH_2^{2+}) + 568.9 (MH_3^{3+}) + 1703.6 (MH^+)
p52	αS1 CN f(24-38)	**1707**	854.4 (MH_2^{2+}) +570.3 (MH_3^{3+})
p53	ß CN f (194-209)	**1717**	1718.5 (MH^+) +859.5 (MH_2^{2+})
p54	ß CN f (15-28)3P	**1790**	895.9 (MH_2^{2+}) + 597.8 (MH_3^{3+})
p55	ß CN f(15-28) 4P	**1870**	936.0 (MH_2^{2+}) + 624.5 (MH_3^{3+})
p56	αS1 CN f (1-16)	**1877**	939.1 (MH_2^{2+}) + 626.6 (MH_3^{3+}) + 549.9 (MH_4^{4+})
p57	ß CN f(193-209)	**1881**	941.2 (MH_2^{2+}) + 628.0 (MH_3^{3+}) + 1881.0 (MH^+)
p58	αS1 CN f (1-17)	**1991**	996.0 (MH_2^{2+}) + 664.6 (MH_3^{3+}) + 582.8 (MH_4^{4+})
p59	ß CN f(14-28) 4P	**1999**	1000.5 (MH_2^{2+}) + 667.5 (MH_3^{3+})
p60	ß CN f (103-119)	**2042**	681.6 (MH_3^{3+}) + 594.2 (MH_4^{4+})
p61	ß CN f(13-28) 4P	**2098**	1050.0 (MH_2^{2+}) + 700.4 (MH_3^{3+})
p62	ß CN f(12-28) 4P	**2212**	1106.6 (MH_2^{2+}) + 738.3 (MH_3^{3+})
p63	ß CN f(11-28) 4P	**2340**	1171.0 (MH_2^{2+}) + 781.0 (MH_3^{3+}) +586.0 (MH_4^{4+})
p64	ß CN f(142-161)	**2340**	781.3 (MH_3^{3+}) + 1171.3 (MH_2^{2+})
p65	ß CN f (159-183)	**2746**	1374.0 (MH_2^{2+}) + 916.6 (MH_3^{3+})
p66	αS1 CN f (1-23)	**2764**	1382.6 (MH_2^{2+}) + 922.1 (MH_3^{3+}) + 691.8 (MH_4^{4+}) + 553.7 (MH_5^{5+})
p67	αS1 CN f (174-199)	**2836**	1418.8 (MH_2^{2+}) + 946.3 (MH_3^{3+}) +710.2 (MH_4^{4+})
p68	ß CN f (98-124)	**3132**	1567.3 (MH_2^{2+}) + 1044.9 (MH_3^{3+}) +784.3 (MH_4^{4+})
p69	αS1 CN f (157-188)	**3452**	1151.2 (MH_3^{3+}) + 863.9 (MH_4^{4+}) +691.3 (MH_5^{5+}) + 576.2 (MH_6^{6+})
p70	αS1 CN f (156-187)	**3477**	1739.6 (MH_2^{2+}) + 1160.0 (MH_3^{3+}) +870.2 (MH_4^{4+})
p71	ß CN f(37-67)	**3580**	512.5 (MH_7^{7+}) +597.6 (MH_6^{6+}) +717.0 (MH_5^{5+}) +896.1 (MH_4^{4+})
p72	αS1 CN f(85-114)	**3601**	901.2 (MH_4^{4+}) +721.2 (MH_5^{5+})
p73	αS1 CN f (155-187)	**3605**	1803.4 (MH_2^{2+}) + 1202.7 (MH_3^{3+}) + 902.3 (MH_4^{4+})
p74	αS1 CN f (83-114)	**3860**	965.6 (MH_4^{4+}) + 773.0 (MH_5^{5+}) + 644.4 (MH_6^{6+})
p75	ß CN f(59-96)	**4024**	1007.1 (MH_4^{4+}) + 1342.6 (MH_3^{3+})

CODE	identification	MW	Characteristic ions
p76	ß CN f(57-93)	**4065**	1355.7 (MH_3^{3+}) + 1016.9 (MH_4^{4+})
p77	αS1 CN f(80-114)	**4238**	1060.2 (MH_4^{4+}) + 848.5 (MH_5^{5+}) +707.4 (MH_6^{6+}) + 606.4 (MH_7^{7+})
p78	ß CN f (55-93)	**4253**	1064.3 (MH_4^{4+}) +1418.6 (MH_5^{5+})
p79	ß CN f (53-93)	**4454**	1114.2 (MH_4^{4+}) +1485.1 (MH_5^{5+})
p80	ß CN f (51-93)	**4696**	1175.0 (MH_4^{4+}) +1566.0 (MH_5^{5+})
p81	ß CN f (99-159)	**7054**	1008.6 (MH_7^{7+}) +882.6 (MH_8^{8+}) +784.8 (MH_9^{9+}) +706.4 (MH_{10}^{10+}) +642.3 (MH_{11}^{11+}) +588.9 (MH_{12}^{12+})
p82	ß CN f (99-160)	**7182**	1026.9 (MH_7^{7+}) + 898.8 (MH_8^{8+}) + 799.0 (MH_9^{9+}) + 719.3 (MH_{10}^{10+}) + 654 (MH_{11}^{11+}) + 599.6 (MH_{12}^{12+})
	Phe-Phe	**312**	313.2 (MH^+) +166.0 (y_1) +120.0 (i_2)

● are Non-Proteolytic peptides, which will be characterized in the following chapter.

Between these peptides, 15 were non proteolytic peptide-like structures (NPPs), which will be discussed in the following chapter, only one has been found to take origin from αS2 casein, and the other 66 identified peptides originate all from αS1 and ß casein. 34 peptides were already identified in other works [2,7,8,9,10,11], 48 have not been previously reported.

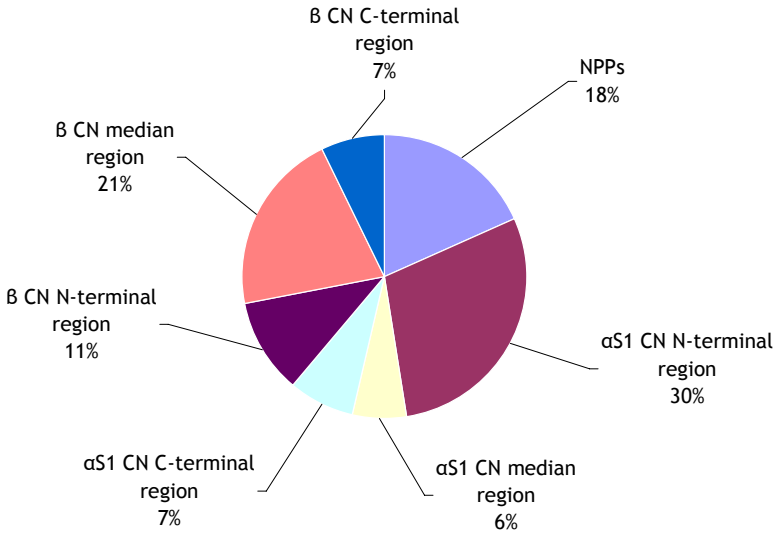


Figure 4.6. Percent distribution of peptides in function of their sequence of origin.

Principal Components Analysis (PCA) was performed in order to have a general overview of the factors affecting the data variability. PCA extracts from a set of *p* variables a reduced set of *m* components (or factors) that account for most of the variance in the *p* variables: *m* can be considered as super-ordinate dimensions.[12]

PCA was performed on data derived from the semiquantification values (corrected for the moisture content) of every peptide in every sample. The score plot including all the analyzed

samples derived from the calculation of the factor score of every sample for each component is reported in Figure 4.7. Samples have been labeled according to their ageing times (in days)

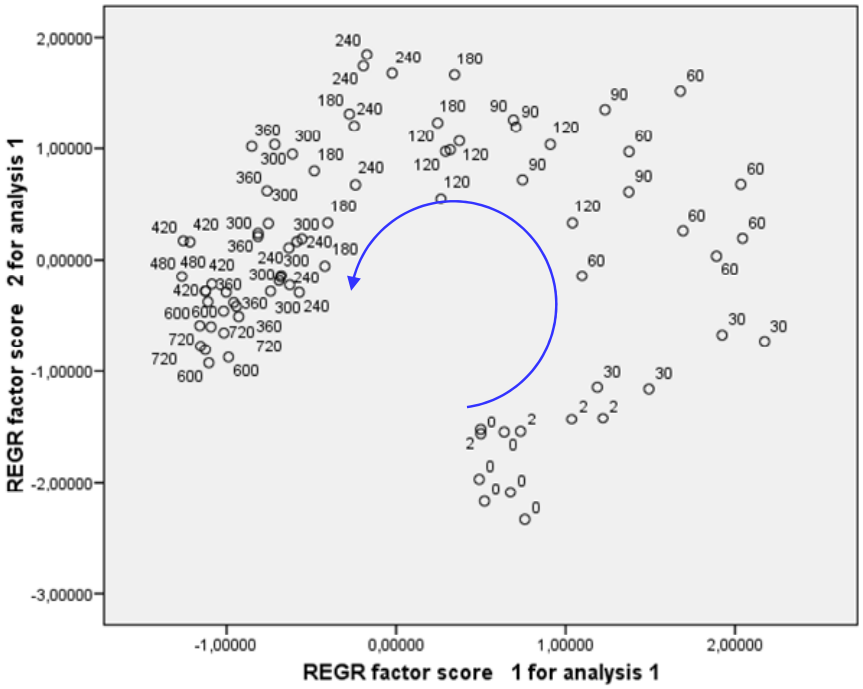


Figure 4.7. Score plot of the collected data: samples are indicated with their ageing time (expressed in days).

It is evident that the distribution of the samples is mainly affected by the ageing time; it could be also noticed that among "younger" samples there was an higher variability than in "older" samples, i.e. great changes occurs in the first months of ripening, and at long ageing time (from 10-12 to 24 months) samples seemed to be progressively more similar.

The trend of the single peptides during the ageing time was studied and peptides were grouped according to the ageing period in which they reached the maximum amount: peptides reaching the maximum in the curd and progressively disappearing (group 1), peptides reaching the maximum during the first three months (Group 2), peptides reaching the maximum around the 4^{th}-6^{th} month (Group 3), peptides having the maximum at the 10^{th}-12^{th} month (Group 4), peptides whose quantity was always increasing during ageing (Group 5).

The Loading plot of the PCA is reported in Figure 4.8: peptides (from P1 to P82 as in Table 4.3) are represented in a 2-dimensional plan generated by the two calculated main factors (component 1 and component 2) which explain, respectively, 40.8 and 18.3% of the variance. Peptides have been grouped according to their trend during the ageing time, as indicated above.

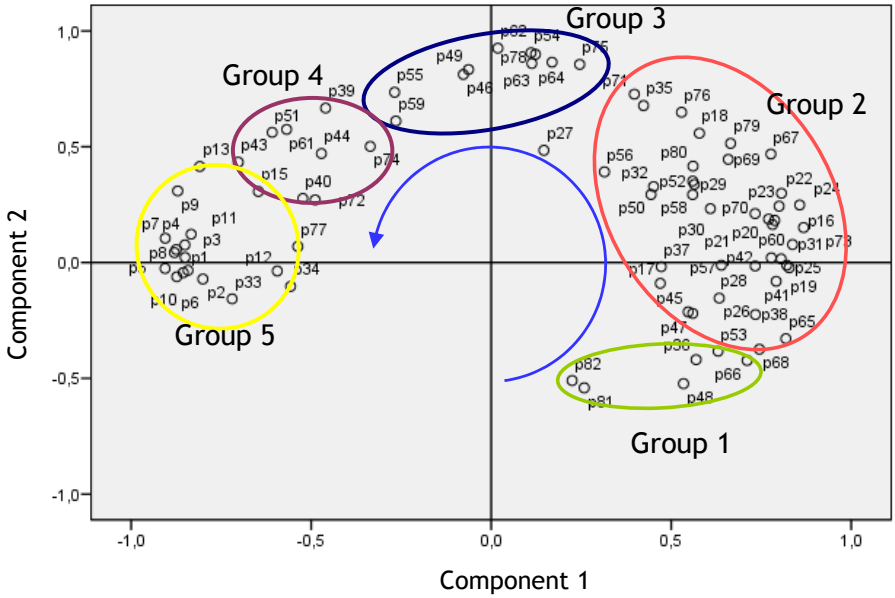


Figure 4.8. Loading plot of Parmigiano-Reggiano peptides.

A comparison of the Score Plot with the Loading Plot confirmed that the major source of the variability was the ageing time and that the oligoepeptide fraction accordingly changed during the ageing: peptides in group 1 were the most abundant in curd samples, whereas group 5 peptides characterized the long ripened cheeses.

The composition of the different peptide groups will be reported in tables and discussed in detail.

4.4.1 Changes of the peptide pattern during the ageing

4.4.1.1 Group 1 peptides (maximum amount in the curd)

The average amounts of these peptides at the different ageing times is summarized in Figure 4.9.

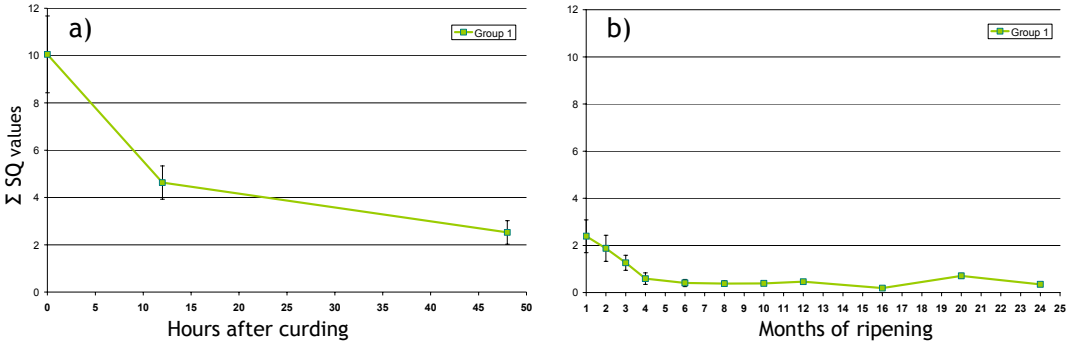


Figure 4.9. Amounts of group 1 peptides in function of the ageing time: a) hours after curding and b) months of ageing. The amount is obtained by the sum of semiquantified values of the peptides of the group[‡].

The peptides belonging to this group are reported in Table 4.4.

Table 4.4. List of group 1 peptides.

Group 1	Group 1 (identity)
P66	αS1 CN f (1-23)
P81	ß CN f (99-159)
P82	ß CN f (99-160)

αS1 CN f (1-23) is a very well known, ubiquitous peptide in dairy products. It is derived from the chymosin action, the main rennet enzyme thus rapidly increasing in the curd at the start of the process. Since this peptide presents many preferential cleavage sites of bacterial proteolytic enzymes, it is rapidly degraded when chymosin activity decreases, although a small amount can be found up to the end of ageing. Peptides ß CN f (99-159) and ß CN f (99-160) were the only peptides found also in the milk used for cheese preparation, so they likely derive from milk endogenous enzymes and they are also rapidly degraded in the first hours of curding.

[‡] SQ value= (Integrated area of the peptide/Integrated area of the internal standard), corrected by moisture content, as described in the experimental section.

4.4.1.2 Group 2 peptides (maximum during the first three months of ageing)

This group is the most rich in peptides. Their relative amount is quite low in the first hours of the cheese-making process, then it reaches a maximum after 1-3 months. Then, they definitively decrease during the subsequent months, whitout disappearing (Figure 4.10).

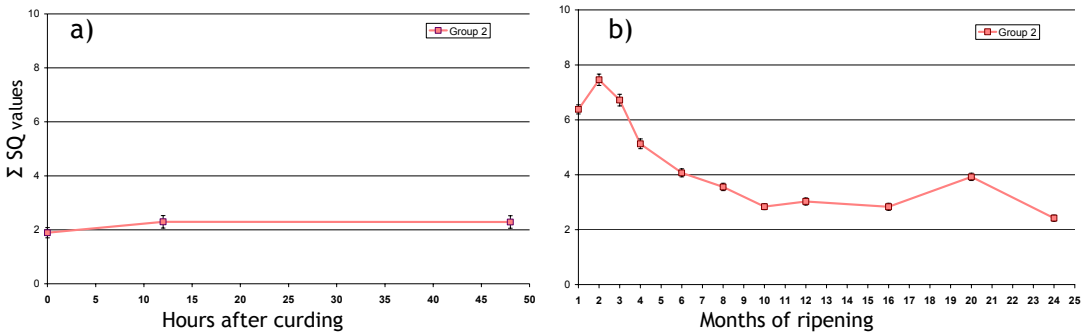


Figure 4.10. Amounts of group 2 peptides in function of the ageing time: a) hours after curding and b) months of ageing. The amount is obtained by the sum of semiquantified values of the peptides of the group.

The identification of these peptides (reported in Table 4.5) allowed to shed some light on the enzymatic activities occurring during the first three months.

Table 4.5. List of group 2 peptides . In BOLD letters, peptides whose origin by a particular enzymatic activity can be inferred (by the N-terminal or C-terminal (or both) end).

max at 1 months	max at 2 months		max at 3 months
ß CN f (1-3)	**ß CN f (1-6)**	**αS1 CN f (18-23)**	**αS1 CN f (1-13)**
ß CN f (1-4)	**αS1 CN f (1-4)**	**αS1 CN f (1-9)**	**αS1 CN f (1-6)**
αS1 CN f (186-192)	**αS1 CN f (10-14)**	**αS1 CN f (174-199)**	**αS1 CN f (24-30)**
ß CN f (47-52)	**αS1 CN f (16-20)**	**αS1 CN f (156-187)**	**αS1 CN f (30-**35)
ß CN f (199-209)	**αS1 CN f (10-16)**	**αS1 CN f (155-187)**	α S2 CN f (154-162)
ß CN f (195-209)	**αS1 CN f (10-13)**	**ß CN f(57-93)**	ß CN f(37-67)
ß CN f (194-209)	**αS1 CN f (1-7)**	**ß CN f (53-93)**	ß CN f (60-**68)**
ß CN f(193-209)	**αS1 CN f (17-23)**	**αS1 CN f (24-34)**	**αS1 CN f (157-188)**
ß CN f (**103**-119)	**αS1 CN f (14-23)**	**αS1 CN f (24-36)**	**ß CN f(59-96)**
ß CN f (159-**183)**	**αS1 CN f (1-14)**	**αS1 CN f(24-38)**	**ß CN f (55-93)**
ß CN f (51-93)	**αS1 CN f (1-16)**		
ß CN f (98-124)	**αS1 CN f (1-17)**		

After one month of ripening, all the peptides deriving from the C-terminal part of ß casein, starting from **ß CN f(193-209)**, which is known to be a chymosin-[13] but also a cell envelope proteinase-derived peptide[2], reach their maximum. In particular, in Parmigiano-Reggiano cheese, the cleavage between the positions 192-193 might be attributed mainly to LAB

proteinases: in fact its trend showed a delayed increase in comparison with αS1 CN f (1-23), it was not consumed as αS1 CN f (1-23) and showed global similarity with other group 2 peptides. The related peptides (Figure 4.11), are likely to be produced by an aminopeptidasic degradation.

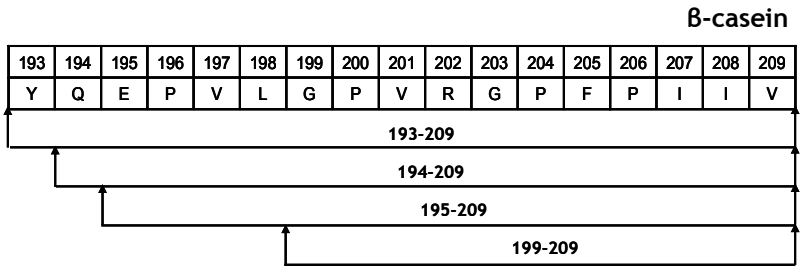


Figure 4.11. Observed peptides derived from the C-terminal end of ß casein.

According to their trend during ageing (reported in details in Figure 4.12) it seems that, although ß CN f(193-209) is present up to the end of the ageing period, the aminopeptidasic activity which give rise to the smaller peptides decreases, since these peptides disappear after 4-6 months of ageing.

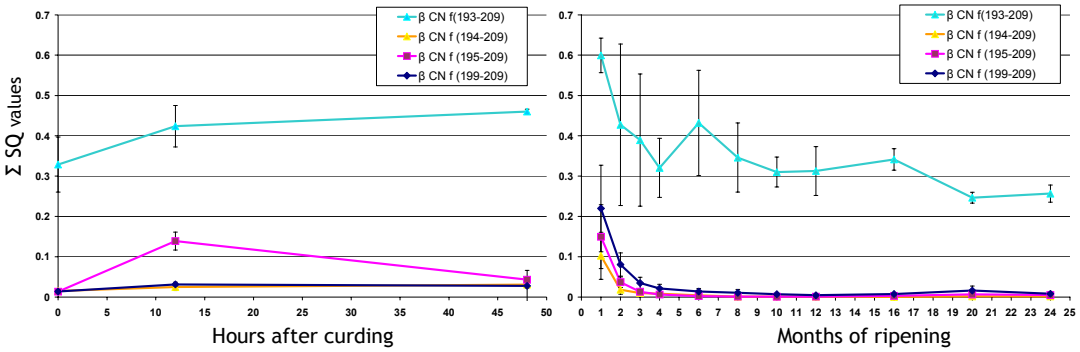


Figure 4.12. Amounts of peptides derived from the C-terminal end of ß casein in the course of the time.

Peptides **ß CN f (1-3)** and **ß CN f (1-4)** (maximum at 1 month) and peptide **ß CN f (1-6)** (maximum at 2 months) derive from LAB cell wall associated proteinases.

Peptides deriving from the **αS1 N-terminal region** reach a maximum at 2 months of ageing, with the simultaneous decrease of the peptide αS1 CN f (1-23). Their N- and C-terminal ends actually correspond to well-known cleavage sites of LAB cell-wall associated proteinases (in red in Figure 4.13) and LAB intracellular endopeptidases (in blue, in Figure 4.13).[14]

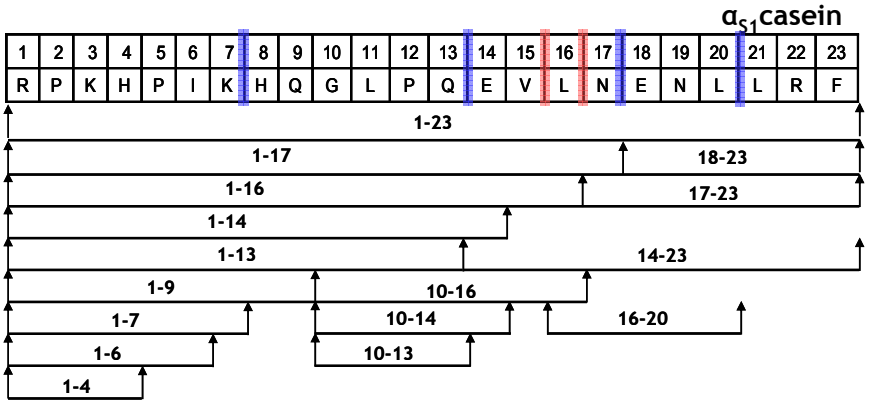


Figure 4.13.Observed peptides derived from the N-terminal part of αS1 casein.

In general, peptides released from casein by proteinases are subsequently hydrolyzed by peptidases inside the cells or by peptidases released in the matrix after cell lysis (Figure 4.1 and ref. 1). In particular, LAB cell-wall associated proteinases are known to cleave the bonds Val_{15}-Lys_{16} Lys_{16}-Asn_{17}, thus forming the observed peptides αS1 CN f(1-16) and f(17-23) starting from the αS1 CN f(1-23) peptide.

The observed trends of other peptides belonging to this group confirmed the action of the LAB enzymes (both proteinases and peptidases) directly on the free substrate αS1 CN f (1-23), indicating that these peptides likely derive from the internal cleavages of the peptide originated by chymosin. The decrease of these peptides after the first months is due to their degradation to amino acids by the action of exopeptidases, i.e. aminopeptidases.

Peptides αS1 CN f (1-4), αS1 CN f (1-6), αS1 CN f (1-14) and αS1 CN f (10-14) do not correspond to known cleavage sites of endoproteases or endopeptidases: they could be due to a typical enzymatic activity characteristic of the microflora of Parmigiano-Reggiano or possibly to carboxypeptidase activities.

After three months of ripening, peptides apparently derived from the action of *Lactococcus* cell wall proteinases (CEP) acting on ß casein (as desumed by their N-terminal or C-terminal ends) also reached a maximum, : ß CN f (53-93), ß CN f (55-93), ß CN f(57-93) and ß CN f(59-96) (N-terminal ends correspond to cleavage sites of cell proteinases); ß CN f (60-68) (both ends); ß CN f (1-6) (C-terminal end). [14]

Also several peptides derived from the **C-terminal region of αS1 casein** can be attributed to the activity of *Lactococcus* cell wall proteinases or endopeptidases: αS1 CN f (157-188), αS1 CN f (156-187), αS1 CN f (155-187), αS1 CN f (174-199).

Cell envelop proteinases of Lactococcus spp[1] also extensively degrade the **αS1 casein region included between Phe_{24} and Asn_{38}**, giving rise to peptides αS1 CN f (24-34), αS1 CN f (24-36), αS1 CN f(24-38), αS1 CN f (24-30) and probably also to αS1 CN f (30-35).

Finally, ß CN f (103-119), ß CN f (159-183) and ß CN f (60-68) showed a C-terminal ending compatible with cathepsin cleavages[15,16]. Cathepsins are proteinases associated with somatic cells in milk, with high activities although present in low quantities.

In summary, the second group of peptides is mainly produced by peptidase and proteinase activities of LAB, generating a large number of peptides during the first months of cheese ripening, which will be subsequently largely degraded during ageing, very likely by exoproteasic activity. In any case, a low but significant amount of these peptides can also be found in the ripened product (Figure 4.10). It is quite evident that LAB enzymes strongly influence the proteolytic modifications in cheese. As it has been said before, LAB in Parmigiano-Reggiano mainly derive from the natural whey starter; the dominant microflora of the natural whey starter are the lactic acid bacteria selected by the process of curd heating and, therefore, these LAB are thermophilic. Among them, *L. helveticus* is the dominant one[17], followed by *S. termophilus*. Since LAB can be found as different strains, the use of a natural whey starter implies a strong link between the product and its original geographic area of production. Thus, some of the unusual cleavage sites here observed could be typical of the strains present in the starter.

4.4.1.3 Group 3 peptides (maximum at 4-6 months of ageing).

Group 3 include peptides with a maximum at 4-6 months of ageing, which strongly decrease after the 10th -12th month (Figure 4.14).

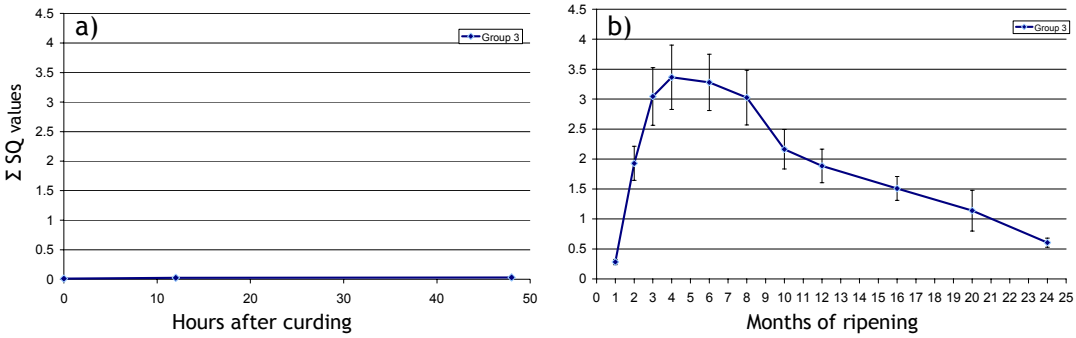


Figure 4.14. Amounts of group 3 peptides in function of the ageing time: a) hours after curding and b) months of ageing. The amount is obtained by the sum of semiquantified values of the peptides of the group.

Peptides given in Table 4.6 belong to this 3rd group.

Table 4.6. List of group 3 peptides.

Group 3	Group 3 (identity)
P49	ß CN f(17-28) 3P
P54	ß CN f (15-28)3P
P55	ß CN f(15-28) 4P
P59	ß CN f(14-28) 4P
P62	ß CN f(12-28) 4P
P63	ß CN f(11-28) 4P
P40	αS1 CN f (169-179)
P46	ß CN f (82-95)
P64	ß CN f(142-161)

These peptides are unrelated to group 1 and group 2 peptides: this suggests that the enzymes involved in the formation of this group act, starting after the first month of ageing, on the intact proteins. Another characteristic feature is that this group mainly includes phosphopeptides of ß-casein ending at position 28: the cleavage between the positions 28-29 is known to derive from the action of plasmin[10]. All these phosphopeptides differ at the N-terminal end and therefore seem to derive from the action of aminopeptidases. Since these peptides originate from plasmin, an endogenous milk enzyme easily denatured by heat treatments, they can be considered as good markers of the technology (raw milk is used for Parmigiano-Reggiano cheese production).
Several hypothesis can be made also on the origin of the other peptides belonging to this group. αS1 CN f (169-179) had only one recognizable cleavage site (Tyr_{168}-Pro_{169}) already ascribed to *Lactococcus* endopeptidasic activity, which can be assumed to act on a longer precursor, such as αS1 CN f (157-188) peptide observed in group 2. The ß CN f (82-95) peptide was consistent with the elastase activity[18], which cleaves Val_{82}-Val_{83} and Val_{95}-Ser_{96} bonds, whereas the ß CN f(142-161) peptide derives from a Cathepsin D activity[19]. Elastase and cathepsins are milk enzymes which derives from somatic cells, very specific towards caseins.

4.4.1.4 Group 4 peptides, with a maximum at 10-12 months of ageing

This group includes peptides which reach a maximum between the 8th and the 12th month, and then slowly decrease and are still present in good amount in the 24-months old cheese (Figure 4.15).

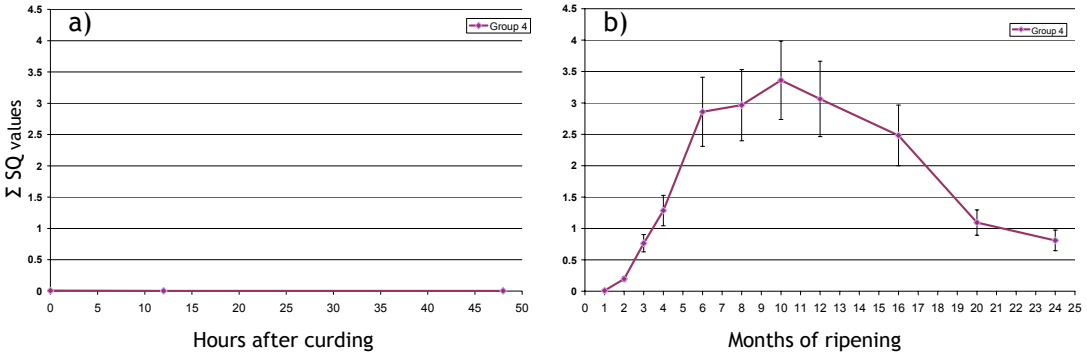


Figure 4.15. Amounts of group 4 peptides in function of the ageing time: a) hours after curding and b) months of ageing. The amount is obtained by the sum of semiquantified values of the peptides of the group.

Group 4 peptides are reported in Table 4.7.

Table 4.7. List of group 4 peptides.

Group 4	Group 4 (identity)
P39	ß CN f(17-25)2P
P43	ß CN f(16-25)3P
P51	ß CN f (16-28)3P
P61	ß CN f(13-28) 4P
P44	ß CN f (71-83)

This group mainly includes other phospho-peptides of ß casein, probably derived from the combined action of plasmin and aminopeptidases.

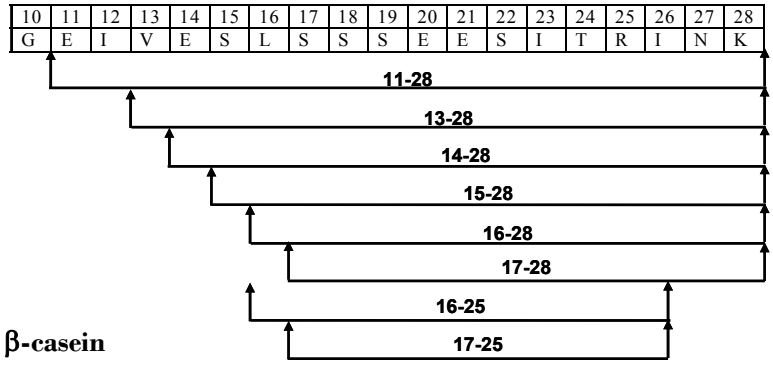


Figure 4.16. Observed phospho peptides from ß-casein in Parmigiano-Reggiano cheese.

4.4.1.5 Group 5 peptides (increasing during the ageing time).

The amount of these peptides incresed during time (Figure 4.17).

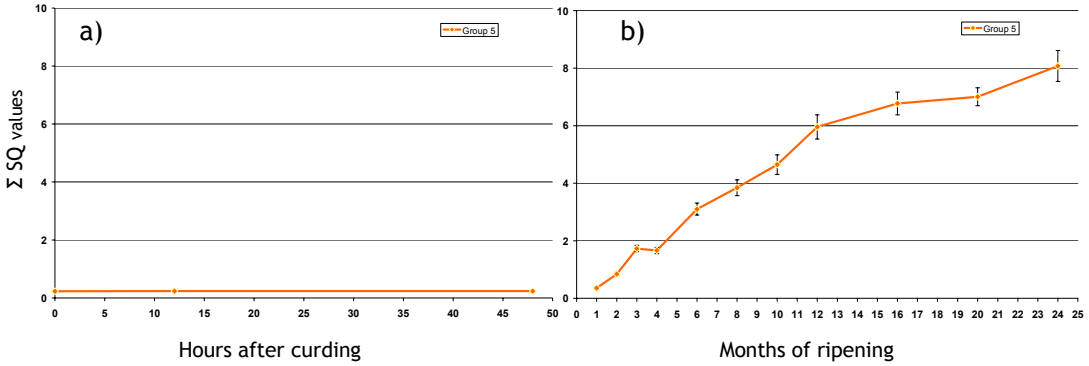


Figure 4.17. Amounts of group 5 peptides in function of the ageing time: a) hours after curding and b) months of ageing. The amount is obtained by the sum of semiquantified values of the peptides of the group.

Peptides of this group are reported in Table 4.8: this group includes two "families": all non-proteolytic peptide-like molecules (NPP, discussed in the following chapter), and other peptides arising from the direct proteolysis of caseins.

Table 4.8. List of group 5 peptides.

Group 5	Group 5 (identity)
P1	Lac-Val
P2	Lac-Leu
P3	Pyr-Ile
P4	γ Glu-Val
P5	γ-Glu-Ile
P6	γ-Glu-Leu
P7	γ-Glu-Phe
P8	γ-Glu-Tyr
P9	Lac-Ile
P10	Lac-Met
P11	Lac-Phe
P12	Pyr-Leu
P13	Lac-Tyr
P14	Pyr-Phe
P15	γ-Glu-Met
P33	αS1 CN f(75-82)1P
P34	ß CN f (111-119)
P72	αS1 CN f(85-114)
P74	αS1 CN f (83-114)
P77	αS1 CN f(80-114)

Being their origin very different, their trend was analyzed separately (Figure 4.18).

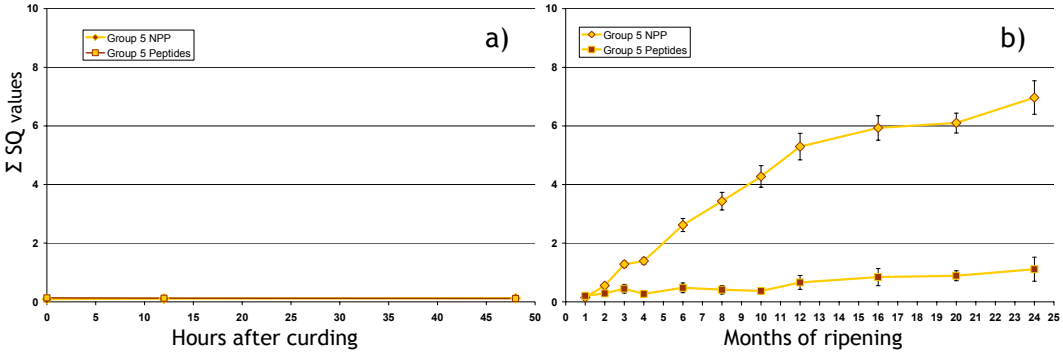


Figure 4.18. Amounts of proteolytic peptides and non-proteolytic peptide-like molecules in function of the ageing time.

Peptides arising from casein proteolysis increase very slowly, whereas the amount of NPPs increase at a much faster rate. This is consistent with the non standard structures of NPPs, which are not recognized and are not cleaved by proteinases. For this reason, they are ideal markers of the ageing of the cheese. The structural characterization of NPPs and the discussion about their formation is reported in the chapter 5.

Peptides αS1 CN f(85-114), αS1 CN f (83-114) and αS1 CN f(80-114) are closely related; they probably derive from the known cleavage sites for plasmin[20] followed by aminopeptidasic, or dipeptidil-peptidasic[4], degradation.

As far as the other peptides are concerned, peptide αS1 CN f(75-82)1P has already been reported[11] and discussed as consistent with lactococcal cell envelope proteinases: this indicate that these activities, as already evidenced by group 2 peptides, are partially present also at long ageing times. Peptide β CN f (111-119) was formed by the cleavage between Phe_{119}-Thr_{120}, known to be typical of elastase[18].

4.4.2 Kinetics of proteolysis in the external and internal parts of the wheels

Data of the trends during ageing in the internal and in the external part of the wheel are given in the figures below, always considering the grouping above presented.

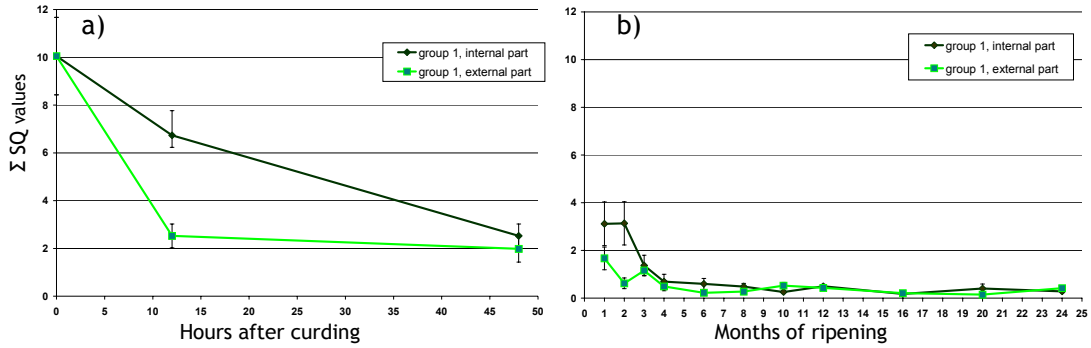


Figure 4.19. Amounts of group 1 peptides, in the external and internal parts of the wheel, during ageing: a) hours after curding and b) months of ageing.

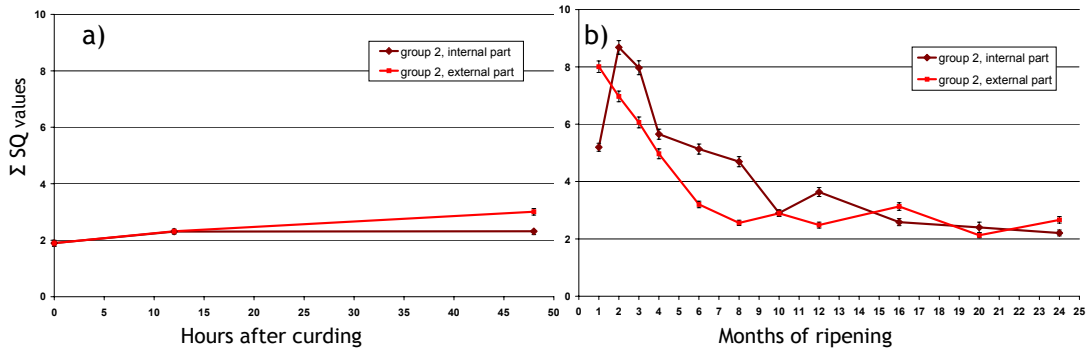


Figure 4.20. Amounts of group 2 peptides, in the external and internal parts of the wheel, during ageing: a) hours after curding and b) months of ageing.

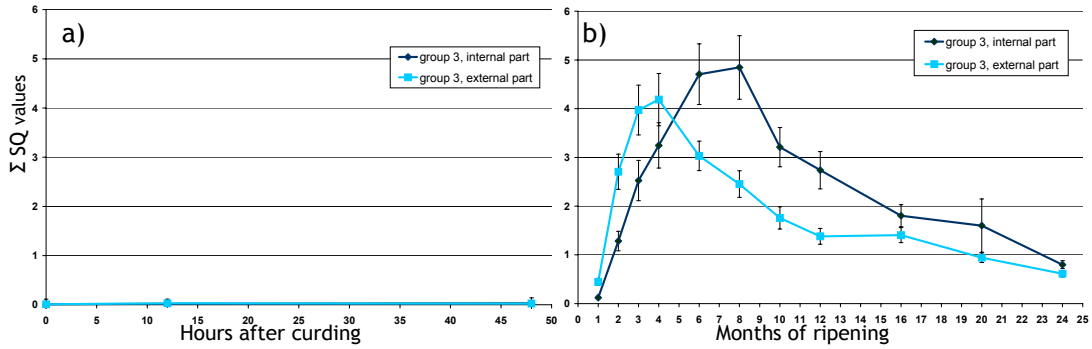


Figure 4.21. Amounts of group 3 peptides, in the external and internal parts of the wheel, during ageing: a) hours after curding and b) months of ageing.

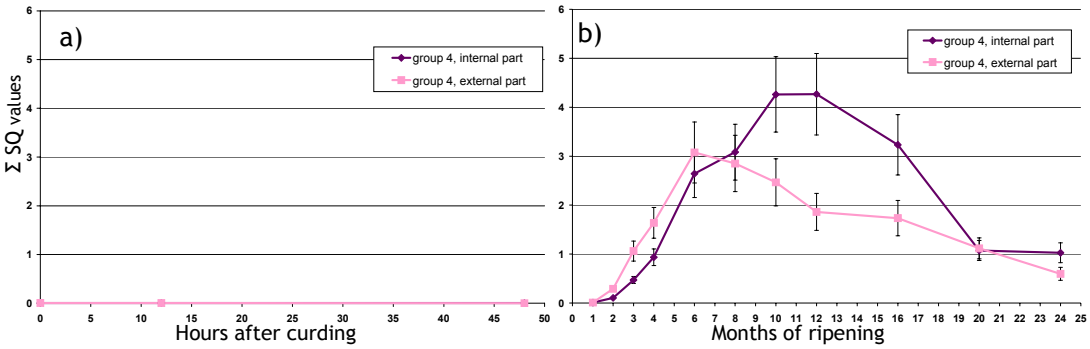


Figure 4.22. Amounts of group 4 peptides, in the external and internal parts of the wheel, during ageing: a) hours after curding and b) months of ageing.

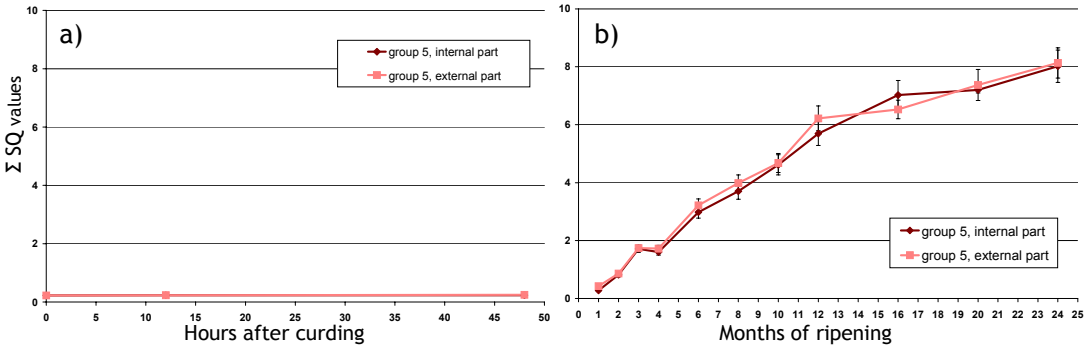


Figure 4.23. Amounts of group 5 peptides, in the external and internal parts of the wheel, during ageing: a) hours after curding and b) months of ageing.

4.4.2.1 Group 1 peptides

From the graphs reported in Figure 4.19 it is evident that group 1 peptides started at the same concentrations and decreased more rapidly in the external part of the wheels, thus indicating that they were hydrolyzed faster in the external part.

4.4.2.2 Group 2 peptides

The graphs reported in Figure 4.20 show that, during the first hours of curding, group 2 peptides increased more rapidly in the external part of the wheels than in the internal part. This confirmed once again that group 2 peptides derived mainly from group 1.

Moreover, in the first months, group 2 peptides of the external part were also degraded more rapidly.

During the production of Parmigiano-Reggiano, the external part is more exposed to temperature decrease after extraction of the cooked curd from vats, and is more rapidly affected by salt during brining. Thus, the internal part reaches the conditions of the external

part lately. The data here presented are consistent with the fact that the internal part of the cheese reaches the conditions of the external part more slowly. Thus, the peptide content is the expression of the enzymatic activity modulated by the medium conditions.

Moreover, these data seem to indicate that LAB enzymes were more active in the external part, probably indicating that LAB typical of Parmigiano-Reggiano natural whey starter grew more rapidly in the external part; these data would be confirmed by Prof. Neviani Microbiological unit working at the same project.

4.4.2.3 Group 3 peptides

Also for group 3 peptides (in Figure 4.21), the delay of proteolysis btween the external and the internal part of the wheel is evident. The amount of peptides increased and were hydrolyzed more rapidly in the external part; in correspondence with the 4th month in the external part and with the 8th month in the internal part. The salt concentration at 8-10 months of ageing in the internal part become as it was at 4 months in the external part (data collected by "Consorzio del Parmigiano-Reggiano" and not shown). The enzymes involved seem to be directly modulated by salt and by the dimension of the wheels, i.e. the technology of production.

4.4.2.4 Group 4 peptides

For the 4th group of peptides, the same conclusion as for group 3 peptides can be drawn, except for the fact that in this case, the amounts of peptides reach the maximum, both in the external and in the internal parts of the cheese, later on (6 months and 10-12 months respectively).

4.4.2.5 Group 5 peptides

Peptides belonging to group 5 are once again very particular: they are fairly homogeneous between the external and the internal part of the wheel, meaning that activities which generated them are less salt-dependant then all the other activities involved in cheese ageing.

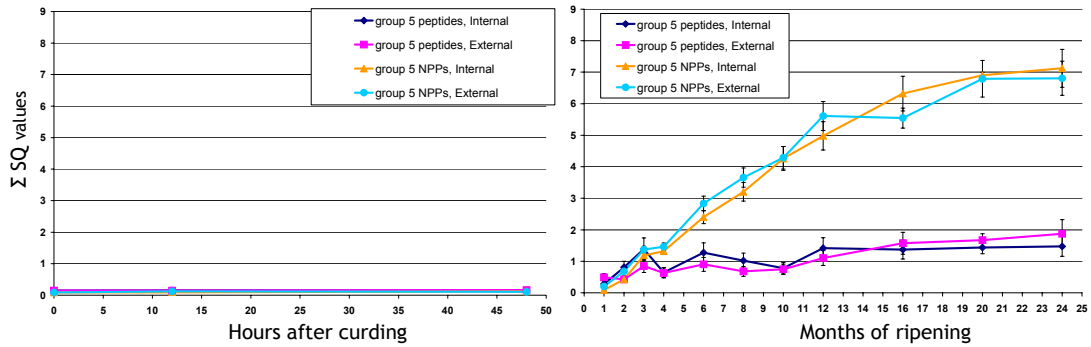


Figure 4.24. Amounts of group 5proteolytic peptides and non-proteolytic peptide-like molecules in function of the ageing time and of the part of the wheel.

4.5 Conclusions and perspectives

In this work an extensive and detailed study of the oligopeptide fraction of Parmigiano-Reggiano cheese, mapped for the first time completely from the caseification process to 24 months of ageing, has been performed by original LC/MS methodologies.

This fraction has been shown to strongly evolve during ageing, leading to dynamic changes of the peptide pattern during ageing. Proteolysis occurs with a different kinetics in the internal and external part of the wheels, and only at long ageing times the differences disappear.

The peptidic fraction studied can be considered a mirror of the enzymatic activities which take place in cheese. Typical cleavage sites of enzymes coming from lactic acid bacteria, rennet and milk have been found in many identified peptides, allowing to assess the most characterizing enzymatic activities which formed the observed peptides at the different ageing times (Figure 4.25).

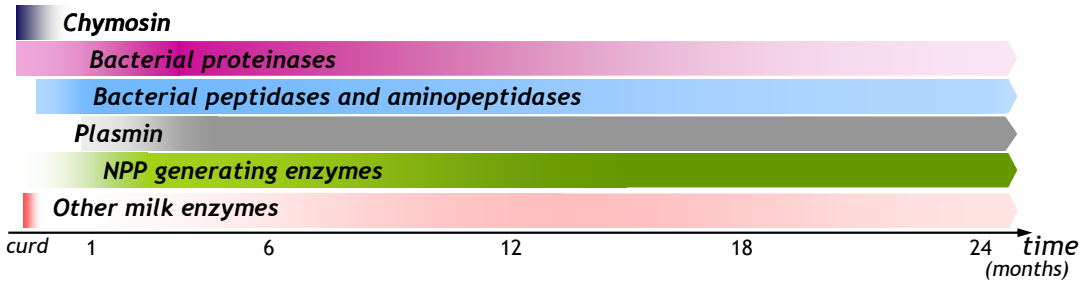


Figure 4.25. Enzymes activities during Parmigiano-Reggiano ageing mapped by their specific peptides released in cheese. A more intense color corresponds to a more intense activity (presumed information based on our data).

The oligopeptide fraction dynamically evolves during the ageing time according to an evolving enzymatic activity determined by the combination of typical starter bacteria and typical technology. These activities are strictly related to the technology (temperature of cooking, amount of salt, wheel dimensions, etc.) and to the area of production (since the starter bacteria directly come from the previous day whey culture), thus the evolving proteolytic pattern constitutes a dynamic fingerprint of the Parmigiano-Reggiano typicality. A parallel study on the microorganism development and on the enzymatic activities has been performed by the Unit of Food Microbiology headed by Prof. E. Neviani on the same samples.

4.6 Acknowledgements

This work was performed with a grant by the Emilia-Romagna Region, in collaboration with the Consorzio del Parmigiano Reggiano.

4.7 References

[1] P.L H McSweeney Biochemistry of cheese ripening *International Journal of Dairy Technology* **2004** 57 (2-3), 127–144.

[2] M.J. Sousa, Y. Ardo, P.L.H. McSweeney, Advances in the study of proteolysis during cheese ripening *International Dairy Journal* **2001** 11, 327–345

[3] http://merops.sanger.ac.uk/index.htm

[4] Kirsi Savijoki . Hanne Ingmer . Pekka Varmanen Proteolytic systems of lactic acid bacteria, *Appl Microbiol Biotechnol* **2006** 71: 394–406

[5] Deutsch S M, Molle D, Gagnaire V, Piot M, Atlan D and Lortal S Hydrolysis of sequenced β-casein peptides provides new insight into peptidase activity from thermophilic lactic acid bacteria and highlights intrinsic resistance of phosphapeptides. *Applied and Environmental Microbiology* **2000**66 5360–5367.

[6] Stefano Sforza, Linda Ferroni, Gianni Galaverna, Arnaldo Dossena, And Rosangela Marchelli, Extraction, Semi-Quantification, and Fast On-line Identification of Oligopeptides in Grana Padano Cheese by HPLC-MS *J. Agric. Food Chem*. **2003**, 51, 2130-2135,

[7] Sforza, S.; Galaverna, G.; Neviani, E.; Pinelli, C.; Dossena, A.; Marchelli, R. Study of the oligopeptide fraction in Grana Padano and Parmigiano-Reggiano cheeses by liquid chromatography-electrospray ionization mass spectrometry. *Eur. J. Mass Spec.*, **2004**, 10, 421-427.

[8] Paolo Piraino, Vivek K. Upadhyay, Rocco Rossano, Paolo Riccio, Eugenio Parente, Alan L. Kelly, Paul L.H. McSweeney, Use of mass spectrometry to characterize proteolysis in cheese, *Food Chemistry* **2007** 101; 964–972

[9] Addeo F, Chianese L, Sacchi R, Musso SS, Ferranti P, Malorni A., Characterization of the oligopeptides of Parmigiano-Reggiano cheese soluble in 120 g trichloroacetic acid/l. *J Dairy Res*. **1994** 61, 3:365-74.

[10] Ferranti P, Barone F, Chianese L, Addeo F, Scaloni A, Pellegrino L, Resmini P. Phosphopeptides from Grana Padano cheese: nature, origin and changes during ripening. *J Dairy Res.* **1997** 64,4:601-15.

[11] Martin Lund And Ylva Ardö Purification and Identification of Water-Soluble Phosphopeptides from Cheese Using Fe(III) Affinity Chromatography and Mass Spectrometry, *J. Agric. Food Chem*. **2004,** 52, 6616-6622,

[12] C.J. Coker, R.A. Crawford, K.A. Johnston, H. Singh, L.K. Creamer, Towards the classification of cheese variety and maturity on the basis of statistical analysis of proteolysis data—a review, *International Dairy Journal* **2005,** 15 631–643

[13] Visser S, slangen K J, On the specificity of chymosin (rennin) in its action on bovine β casein, *Meth. Milk Dairy* J , **1977,** 31, 1162-1166

[14] P.F.Fox, P.L.H. McSweeney, Proteolysis in cheese during ripening, *Food Rev. Int.,* **1996**12(4), 457-509 ,

[15] T. Considine, S. Geary, A.L. Kelly, P.L.H. McSweeney Proteolytic specificity of cathepsin G on bovine as1- and b-caseins, *Food Chemistry* **2002,** 76, 1, 59-67,

[16] T. Considine, A . Healy,A.L. Kelly,P.L.H. McSweeney, Hydrolysis of bovine caseins by cathepsin B,a cysteine proteinase indigenous to milk, *International Dairy Journal* **2004** 14: 117–124,

[17] M. Gatti , C. Lazzi , L. Rossetti , G. Mucchetti And E. Neviani. Biodiversity in Lactobacillus helveticus strains present in natural whey starter used for Parmigiano Reggiano cheese, *Journal of Applied Microbiology* **2003**, 95, 463–470,

[18] T. Considine, A. Healy, A.L. Kelly, P.L.H. McSweeney, Proteolytic specificity of elastase on bovine b-casein, *Food Chemistry* **1999** 66; 463-470

[19] M.J. Hurley, L.B. Larsen, A.L. Kelly, P.L.H. McSweeney, The milk acid proteinase cathepsin D: a review, *International Dairy Journal* **2000**10: 673-681

[20] Addeo, F., Chianese, L., Salzano, A., Sacchi, F, Cappuccio, U., Ferranti, P. and Malorni, A. "Characterization of the 12 % trichloroacetic acid-insoluble oligopeptides of Parmigiano-Reggiano cheese," *J. Dairy Res*. **1992.**59: 401-411.

5 OCCURRENCE OF "NON PROTEOLYTIC PEPTIDE-LIKE MOLECULES" IN CHEESES

5.1 Introduction

Cheese is a biochemically dynamic product which undergoes significant changes during its ageing period. Indeed, cheese ageing is a rather complex process involving lipolysis, glycolysis and proteolysis, and various chemical and biochemical conversions of milk components down to low molecular weight compounds[1].

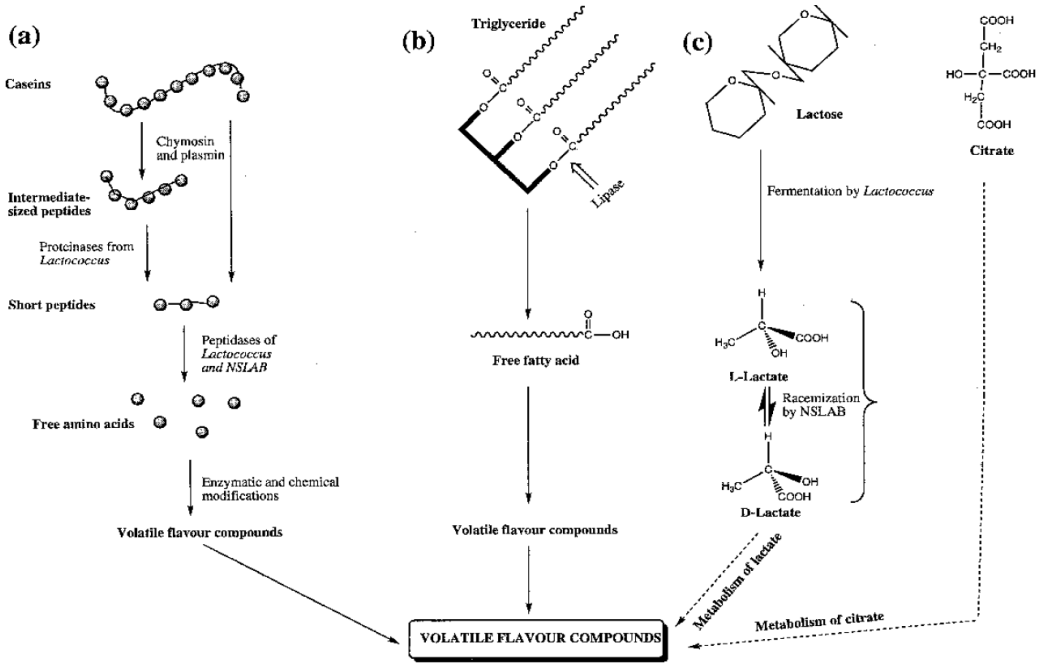


Figure 5.1.General pathways for the biochemistry of cheese ageing: (a) proteolysis, (b) lipolysis, and (c) metabolism of lactose, lactate and citrate.

Glycolysis of Lactose, which leads to lactic acid, is the primary fermentation by the starter culture, allowing Lactic acid bacteria (LAB) to survive in the acidified matrix. This is mainly a pre-aging process, even if starter cultures continue to affect acidity levels for several weeks.

Lipolysis breaks down tryglicerides to free fatty-acids (FFAs), which directly contribute to flavor and texture. FFAs can also indirectly contribute to cheese characteristics, by reacting with alcohols coming from the primary fermentation to yield esters (flavor compounds). In some cases, the FFAs are oxidized when the cheese is perforated with blue molds, aerobically producing specific flavors (as in Gorgozola cheese).

Proteolysis is widely believed to be the most important reactions occurring in cheese.[2] Endogenous enzymes from milk (plasmin), from rennet (chymosin) and enzymes coming from all the bacterial population (starters, nonstarter lactic acid bacteria (NSLAB), secondary or adjunct starters and adventitious bacteria) contribute to proteolysis.[3,4]

Proteolysis transforms part of the curd milk proteins (caseins) into smaller peptides and free amino acids (FAAs), leading to flavor and texture development.[1] FAAs can be further degraded to other acids, alcohols, ammonia and sulfur compounds, and eventually transformed into esters.

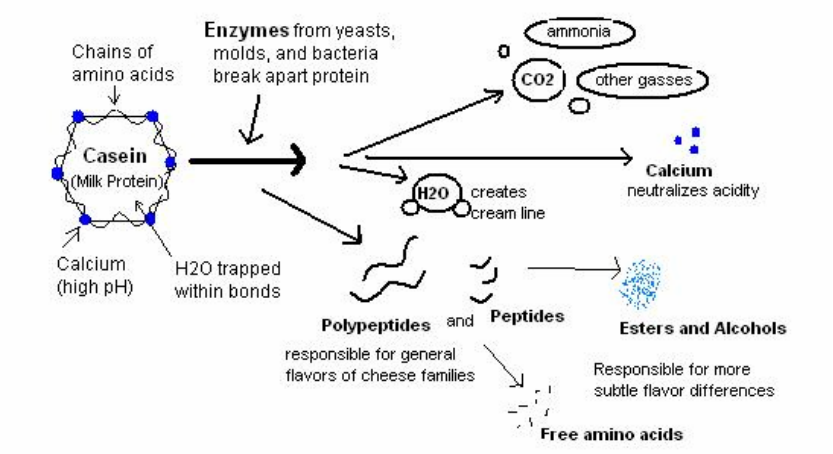


Figure 5.2. Main events and products of proteolysis.

The initial stages of proteolysis have been extensively studied; the later stages, where amino acids are catabolized, are less well understood, although important from a flavour perspective.[1] At this stage, rearrangements between molecules are scarcely known. In the last few years, γ-glutamyl amino acids were reported in literature as present in foods. γ-L-glutamyl-L-leucine, γ-L-glutamyl-L-valine and γ-L-glutamyl-L-cysteinyl-β-alanine were isolated in edible beans[5] and have been characterized as providing the "Kokumi" taste (term used in the flavour industry to describe characteristics such as continuity, mouthfulness, richness and thickness in contrast to monosodium glutamate, which is usually characterized as having the “Umami” taste)). γ-Glu-Phe, γ-Glu-Tyr, and γ-Glu-Leu have been isolated and characterized in Compté cheese[6]; moreover, synthetic compounds corresponding to the identified dipeptides were tasted: γ-Glu-Tyr was sour and salty; γ-Glu-Phe had a more intense and complex taste, which was brothy and slightly sour, salty, and metallic. γ-Glu-Phe, γ-Glu-Ile and γ-Glu-Leu were also reported by our group to be present in Parma ham[7], where they were demonstrated to be related to the ageing time and possibly linked to the pleasant aged flavour. γ-Glutamyl compounds have been reported (most abundant of all, γ-Glutamyl trans-prop(I)enyl cysteine sulphoxide and γ-glutamyl phenylalanine) in the *Allium* species, where they store compounds of nitrogen and sulfur.[8,9] γ-glutamyl peptide compounds, and the enzymatic methods to prepare them, have been the

object of a recent patent (which comprehend also ß-asparagyl peptides) as compounds which provide a kokumi flavour to consumables and flavour compositions.[10]

Pyroglutamic acid and the enzymatic reaction which catalyze internal cyclization of glutamic acid were also found in cheeses.[11] Pyroglutamyl peptides have been reported in wheat[12] and mushrooms[13] but never in cheese, although the presence of pyroglutamic acid in cheese has been reported. Derivatives of pyroglutamylpeptides have been patented by Nestlé S.A. as taste enhancers which impart umami taste to food-stuffs.[14]

To the best of our knowledge, lactoyl-amino acids are here reported for the first time to be present in foods.

5.2 Aim of the work

In this study, we intended to identify and characteriza unusual peptide-like molecules in ripened cheeses, which derive from amino acids but that are not produced by proteolysis but rather are the result of a peptide neo-synthesis occurring during the ageing period, likely catalyzed by an enzymatic activity. These unusual aminoacidic derivatives (γ-glutamyl-dipeptides, pyroglutamyl-dipeptides and lactoyl-amino acids) will be called "non-proteolytic peptide-like molecules" (NPPs), which do not derive from proteolysis. These compounds will be identified on the basis on spectral data and by comparison with synthetic authentic standards. Their occurrence in cheeses will be investigated and their origin will be discussed.

5.3 Experimental part

5.3.1 Solvents and reagents

For cheese sample preparation, see section 4.3.2 of the previous chapter.

For synthesis of authentic specimens:

- (R)-(+)-2-(Benzyloxy)propionic acid (FLUKA)
- L-Phenylalanine benzyl ester hydrochloride (FLUKA)
- n-Butanol
- Acetic acid CH_3COOH (CARLO ERBA)
- (S) - (-) - 2-Acetoxypropionic acid (ALDRICH)
- FMOC-Phe-WANG resin, 100-200mesh, 0.65 mol/g (NOVABIOCHEM)
- Hexafluorophosphate Benzotriazolyl Tetramethyl Uronium (HBTU) (NOVABIOCHEM)
- BOC-D-Glu-OBzl resin (NOVABIOCHEM)
- Thioanisol (ALDRICH)
- m-cresol (FLUKA)
- NaCl (CARLO ERBA)
- $NaHCO_3$ (CARLO ERBA)

- $KHSO_4$ (RIEDEL DE HAËN)
- Benzotriazole (HOBt) (NOVABIOCHEM)
- N^{α}-Boc-L-Glutamic acid α-^{t}butyl ester (NOVABIOCHEM)
- Trifluoromethanesulphonic acid (TFMSA) (FLUKA)
- Trifluoroacetic acid (TFA) (ACROS)
- methylene chloride (CH_2Cl_2) (J.T.BAKER)
- Diisopropylethylamine (DIPEA) (FLUKA)
- THF dry and not stabilized by quinones
- Piperidine (BIOSOLVE LTD)
- DMF (CARLO ERBA)

5.3.2 Instrumentation

For cheese sample preparation, see section 4.3.2 of the previous chapter. The column used was a Gemini C18 column (5µm, 110 Å, 250x4.6mm) (PHENOMENEX) and the mass spectrometer was a SQ (WATERS).

In the synthesis of authentic specimens:

- Bench scales (ORMA)
- Rotavapor Water bath 461, (BÜCHI)
- pH meter, 691 (METRÖHM)
- Centrifuge HETTICH UNIVERSAL 320 R

5.3.3 Samples

Samples of Parmigiano-Reggiano, Grana Padano and Asiago cheeses were obtained from the market.

5.3.4 Extraction and concentration of the oligopeptide fraction

See section 4.3.4 of the previous chapter.

5.3.5 LC/MS analysis of the oligopeptide fraction

HPLC-MS analysis conditions: eluent A: H_2O (0.2% CH_3CN and 0.1% HCOOH); eluent B: CH_3CN (0.2% H_2O and 0.1% HCOOH); elution: 0-15 min isocratic 100% A, 15-60 min linear gradient from 100% A to 67% A, 60-69 min isocratic 67% A, 69-70 min from 67% to 60% A, 70-80 min from 60%A to 0%A, 80-85 min isocratic 0% A, and reconditioning. Flow rate: 1 ml/min. MS conditions: ESI interface with 80% splitting of the column flow, positive ions, single quadrupole analyzer. Capillary voltage 3.2 kV, cone voltage 30 V, source temperature 100 °C, desolvation temperature 150 °C, cone gas (N_2) 100 L/h, desolvation gas (N_2) 450 L/h. Acquisition in total ion mode (TIC, 100-1900 m/z), scan time 2 s. Data were collected and analyzed with Mass Lynx 4.0 software.

5.3.6 Synthesis of the authentic specimens of NPPs

5.3.6.1 γ – D – Glutamyl – L – Phenylalanine

γ-D-Glu-L-Phe was synthesized by solid phase methodologies. 0.065 mmol of Fmoc-Phenylalanine immobilized on a WANG resin (FMOC-Phe-WANG) were placed on a reaction vessel for solid phase peptide synthesis, mixed for 30 min in CH_2Cl_2 and washed twice with DMF. The Fmoc protective group was removed with a solution of 20 % piperidine in DMF (2 x 5 min). The resin was washed twice with DMF.

A solution containing 0.325 mmol of Boc-D-Glu-OBzl (5 equivalents), 0.310 mmol of HBTU, 0.650 mmol DIPEA in 1.625 ml of DMF was activated for two min and then immediately added to the resin. The reaction vessel was stirred for 30 min and washed twice with DMF. The reaction was monitored by standard Kaiser test . The resin was washed again twice with CH_2Cl_2 and dried under vacuum.

The peptide was cleaved from the resin, with simultaneous removal of the Boc and benzyl protecting groups, by adding 5 ml of a 6:2:1:1 solution of TFA:TFMSA:m-cresol:thioanisol to the reaction vessel and slightly mixing for 60 min. The reaction mixture was washed several times with TFA and the combined eluates were dried under N_2 flux. Diethyl ether was added and the product was extracted with H_2O. The water solution was extracted twice with ethyl ether, then dried under vacuum, redissolved in H_2O:CH_3CN:HCOOH = 50:50:0.2 and analyzed by HPLC/ESI-MS (conditions as in paragraph 5.3.5 of this chapter). Yield = 60 %. Characterization: MH^+ (ESI-MS): expected 295.1 m/z, found 295.1 m/z.

5.3.6.2 γ – L – Glutamyl – L - Phenylalanine

γ-L-Glu-L-Phe was synthesized by solution methodologies. N^{α}-Boc-L-Glutamic acid α-^{t}butyl ester (1mmol) was dissolved in DMF dry together with HBTU and HOBt (0.95 mmol) and the mixture was stirred at room temperature for 10 min, in order to activate the carboxylic function. Phenylalanine ^{t}butyl ester (0.95 mmol) was added together with DIPEA (2 mmol) and the reaction was left under magnetic stirring for 3 h at room temperature. DMF was then evaporated under reduced pressure and the residue was dissolved in methylene chloride. The organic solution was washed with a saturated solutions of $KHSO_4$ (three times), $NaHCO_3$ (three times) and NaCl (three times), dried with Na_2SO_4 and filtered. The solvent was evaporated under reduced pressure and the residue was dissolved in 10 ml of a trifluoroacetic acid / methylene chloride 1:1 solution, containing also 2 % (v/v) of m-cresol and thioanisole as scavengers. The mixture was stirred at room temperature for 1 h and then the solvent was evaporated under reduced pressure. Ethyl ether was added to the residue in order to precipitate the free dipeptide. The precipitate was washed several times with ethyl ether and isolated by centrifugation. It was not possible to calculate the yield as the product dried at rotavapor was obtained as a gel. The product was diluted in H_2O:CH_3CN:HCOOH = 50:50:0.2 and analyzed by means of HPLC-MS. Characterization: MH^+ (ESI-MS): expected 295.1 m/z, found 295.1 m/z.

5.3.6.3 D - Lactoyl – L –Phenylalanine (D-Lac-L-Phe)

D-Lac-L-Phe was synthesized by solution methodologies. (R)-(+)-2-(Benzyloxy)propionic acid (3mmol) was dissolved in 6 ml of CH_2Cl_2, together with 2.85 mmol of HBTU; the mixture was stirred at room temperature for 30 min, in order to activate the carboxylic function. L-Phenylalanine benzyl ester hydrochloride (3 mmol), dissolved in 4 ml of CH_2CL_2, together with DIPEA (7.5 mmol) were added and the reaction was left under magnetic stirring for 4 h at room temperature. The reaction was monitored by TLC (ethyl acetate as eluent, UV absorbance detector, Rf = 0.8). The organic solution was washed with saturated solutions of $KHSO_4$ (three times) and $NaHCO_3$ (three times), dried with $MgSO_4$ and filtered. The product was dried under vacuum (Yield 73%).

The benzyl protective group was removed by a 6-hours hydrogenation catalyzed by Pd/activated carbon powder of the product dissolved in methanol. The yield of hydrogenation was 81%. The dried product was redissolved in H_2O:CH_3CN:HCOOH = 50:50:0.2 and analyzed by means of HPLC-MS. Characterization MH^+ (ESI-MS): expected 238.1 m/z, found 238.1 m/z.

5.3.6.4 L - Lactoyl – L –Phenylalanine (L-Lac-L-Phe)

L-Lac-L-Phe was synthesized by solution methodologies. It was also necessary to synthesize the reagent L-Phenylalanine methyl ester hydrochloride.

Synthesis of L-Phenylalanine methyl ester hydrochloride

2 g of L-Phenylalanine were dissolved in 100 ml of methanol and kept under continuous stirring in an ice bath; $SOCl_2$ was added to a final concentration of 1 M (corresponding to 24 ml of $SOCl_2$). The reaction was monitored by TLC (Partridge (n-butanol:acetic acid:H_2O= 4:1:1) as eluent, UV and ninidrine detection; Rf = 0.6) and took place overnight. The reaction mixture was dried under vacuum and washed with methanol and dried under vacuum 4 times to eliminate completely HCl. Yield = 98%.

(S)-(-)-2-Acetoxypropionic acid (3.12 mmol) was diluted in 4 ml CH_2Cl_2 together with 2.96 mmol of dried HBTU and the mixture was kept under continuous stirring at room temperature for 30 min, in order to activate the carboxylic function. L-Phenylalanine-methyl ester (3.12 mmol) was dissolved in 4 ml of CH_2CL_2, together with DIPEA (9.36 mmol), up to basic pH and then added to the activated acetoxypropionic acid. The reaction was left under magnetic stirring for 4 h at room temperature. The reaction was monitored by TLC (ethyl acetate as eluent, UV absorbance detector, Rf = 0.8). The organic solution was washed with saturated solutions of $KHSO_4$ (three times) and $NaHCO_3$ (three times), dried with $MgSO_4$ and filtered. The product was dried under vacuum, and 0.648 mmol were recovered. (Approximate yield 46 %).

The methyl and the acetyl protecting groups were removed by reaction in 20 ml of dry THF + 1.296 mmol of $Ba(OH)_2$ (which corresponded to 2 equivalents of the protected molecules) for 20 min, on an ice bath. THF was then eliminated under vacuum, and the aqueous solution was acidified up to pH 3 with HCl. The resulting solution (containing Barium salts) was analyzed by means of HPLC-MS. Characterization MH^+ (ESI-MS): expected 238.1 m/z, found 238.1 m/z.

5.3.6.5 D – Pyroglutamyl – L – Phenylalanine

D-Pyr-L-Phe was synthesized by solution methodologies. 3 mmol of D-pyroglutamic acid were diluted in 10 ml CH_2Cl_2 together with 2.85 mmol of dried HBTU and the mixture was kept in continuous stirring at room temperature for 30 min, in order to activate the carboxylic function. L-Phenylalanine-methyl ester (3 mmol) was dissolved in 4 ml of CH_2CL_2, together with DIPEA (9 mmol, up to pH=9) and then added to the activated pyroglutamic acid. The reaction was left under magnetic stirring for 5 h at room temperature. and monitored by TLC (ethyl acetate as eluent, UV absorbance detector, rf = 0.65). The organic solution was washed with saturated solutions of $KHSO_4$ (three times) and $NaHCO_3$ (three times), dried with $MgSO_4$ and filtered. The product was dried under vacuum. (Approximate yield 60 %)

The methyl group was removed by dissolving the ester in 50 ml of dry THF and adding 2 equivalents of $Ba(OH)_2$, 20 min on an ice bath. THF was eliminated under vacuum, and the aqueous solution was acidified up to pH 3 with HCl. The product was precipitated by centrifugation and dried under vacuum. Yield: 54 %. The product was redissolved 100 µg/ml in H_2O:CH_3CN:HCOOH = 50:50:0.2 and analyzed by means of HPLC-MS. Characterization MH^+ (ESI-MS): expected 277.1 m/z, found 277.1 m/z.

5.3.6.6 L – Pyroglutamyl – L – Phenylalanine

L-Pyr-L-Phe was synthesized by solution methodologies. This synthesis was carried exactly as described in paragraph 5.3.6.5, but the product did not precipitated so the entire reaction mixture was analyzed by means of HPLC-MS; the yield was not calculated. Characterization MH^+ (ESI-MS): expected 277.1 m/z, found 277.1 m/z.

5.3.7 Sample spiking with the synthesized standards

50 µl of cheese samples prepared as described in paragraph 5.3.4 were spiked with 2.5, 5 or 10 µl of the synthesized standards at an approximate concentration of 100 µg/ml. A H_2O:CH_3CN:HCOOH solution (40:60:0.1) was added up to a final volume of 100 µl. 20 µl of spiked samples were injected and analyzed by HPLC-MS (conditions as in paragraph 5.3.5 of this chapter).

5.4 Results and discussion

The LC-MS analysis of the oligopeptide fraction of the cheese samples, performed according to the method already reported in the literature,[15] outlined the presence of interesting compounds with low molecular weight, mainly present in long-ripened samples. A chromatogram corresponding to a HPLC-MS analysis of a Parmigiano-Reggiano oligopeptide fraction is reported in Figure 5.3, as an example. The arrows indicate peaks of MW ranging between 200 and 300 Da. It is worth noticing that these compounds accumulated during ageing, as described in the

previous chapter, looking as being unaffected by protease activity. Moreover, their MWs actually did not correspond to those expected for peptides coming from caseins (as checked by a software developed by our research group[16]). Their MWs and characteristic ions found in ESI spectra are reported in Table 5.1. Examples of mass spectra are reported from Figure 5.4 to Figure 5.9.

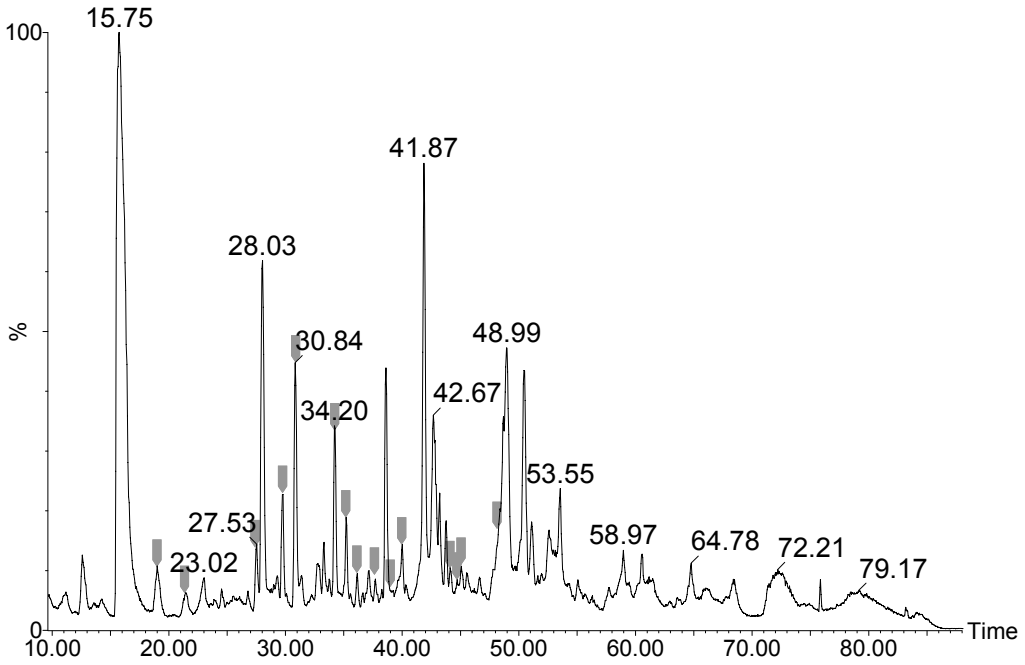


Figure 5.3. HPLC-MS of a Parmigiano-Reggiano cheese ripened 12 months. Arrows indicate peaks of interests.

Table 5.1. List of peaks of interest and relative characteristic ions.

Retention time, Rt (min)	Characteristic ions (m/z)	MW (Da)
19.5	279.3, 150.1	278
21.7	247.2 , 118.1	246
28.4	311.0, 182.0	310
30.7	261.2, 132.1	260
31.7	261.2, 132.1	260
34.7	295.2, 166.0, 120.0	294
35.8	222.1, 176.1	221
36.6	144.1, 190.2	189
37.3	254.2, 208.2, 136.1	253
38.7	243.2, 197.1, 132	242
40.5	243.2, 197.1, 132	242
44.2	158.2, 204.2	203
44.5	158.2, 204.2	203
44.8	277.2, 231.1, 166.0, 120.0	276
48.6	238.1, 192.1, 120	237

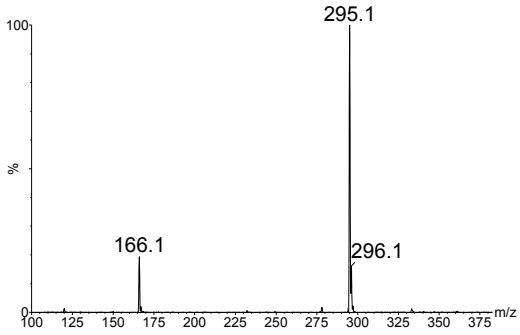


Figure 5.4. Spectrum of the peak at 34.7 min.

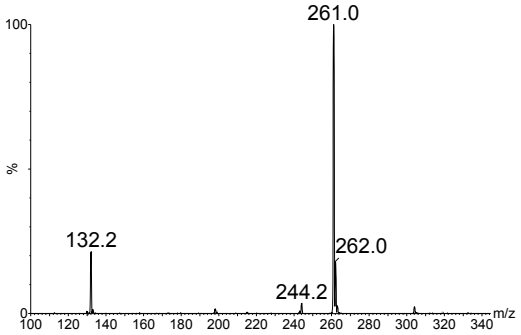


Figure 5.5. Spectrum of the peak at 31.7 min.

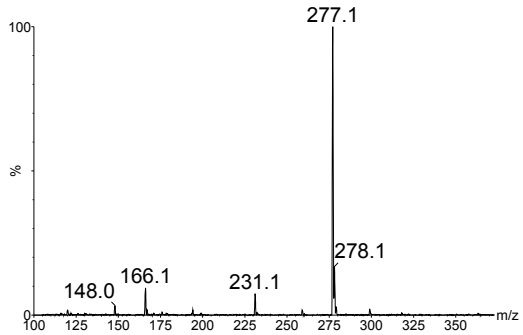


Figure 5.6. Spectrum of the peak at 44.8 min.

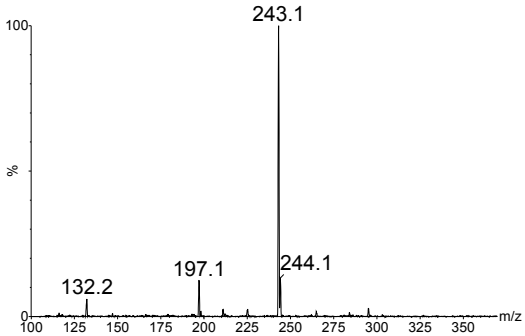


Figure 5.7. Spectrum of the peak at 40.5 min.

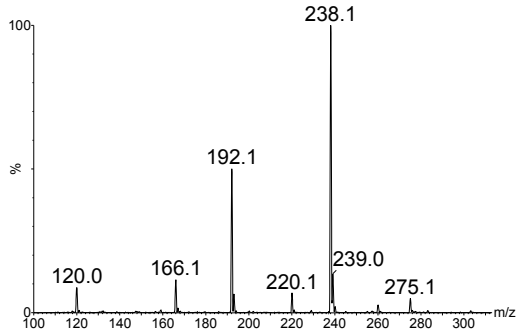


Figure 5.8. Spectrum of the peak at 48.6 min.

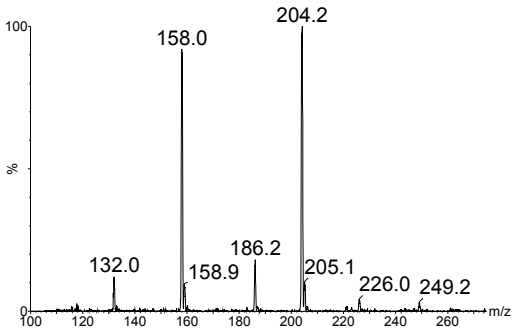


Figure 5.9. Spectrum of the peak at 44.5 min

Fragment ions corresponding to free apolar amino acids were identified in several spectra. For example, ions at m/z 132 (Leu or Ile) were identified at Rt 30.7 and 31.7 min (Mw = 260 Da), Rt 38.7 and 40.5 min (MW = 242 Da), Rt 44.2 and 44.5 min (MW = 203 Da). Analogously, ions at m/z 166 together with the corresponding immonium ion at m/z 120 (diagnostic for Phe) were identified at Rt 34.7 min (MW = 294 Da), Rt 44.8 min (MW = 276 Da), Rt 48.6 min (MW = 237 Da). The difference between the mass of the full molecules and the free amino acids identified corresponded respectively to residues of glutamic acid, pyroglutamic acid and lactic acid.

Since glutamyl peptides seemed to accumulate during ageing, indicating a resistance to proteolysis, we hypothesized that an amidic bond with the γ-carboxylic side chain of glutamyl residue had been formed.

Thus, we hypothesized these compounds being γ-glutamyl-amino acids, pyroglutamyl amino acids and lactoyl amino acids respectively. The putative attributions are summarized in Table 5.2:

Table 5.2. List of presumed identifications.

Retention time, Rt (min)	MW (Da)	Identification
19.5	278	γ-Glu-Met
21.7	246	γ-Glu-Val
28.4	310	γ-Glu-Tyr
30.7	260	γ-Glu-Ile
31.7	260	γ-Glu-Leu
34.7	294	γ-Glu-Phe
35.8	221	Lactoyl-Met
36.6	189	Lactoyl-Val
37.3	253	Lactoyl-Tyr
38.7	242	Pyr-Ile
40.5	242	Pyr-Leu
44.2	203	Lactoyl-Ile
44.5	203	Lactoyl-Leu
45.0	276	Pyr-Phe
48.6	237	Lactoyl-Phe

All these compounds were named “non proteolytic peptide-like molecules” (NPPs). Several structural attributions according to mass spectral data are here reported.

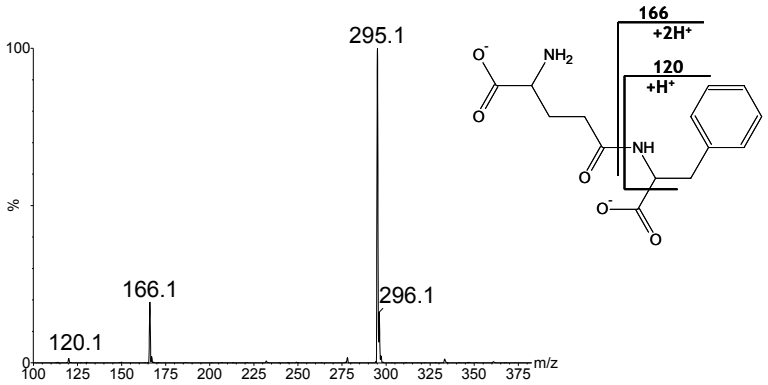


Figure 5.10. Presumed formula of peak at Rt 34.7 min and $[MH]^+$=295.1 m/z based on mass spectral data: γ-glutamyl-phenylalanine.

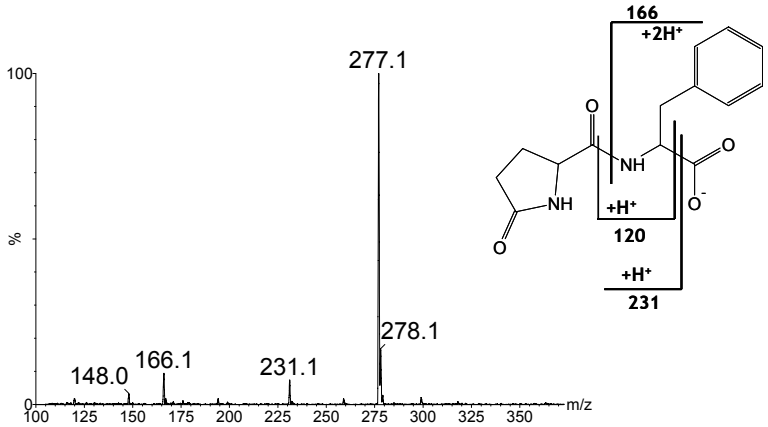


Figure 5.11. Presumed formula of peak at Rt 48.6 min and $[MH]^+$= 277.1 m/z based on mass spectral data: pyroglutamyl-phenylalanine.

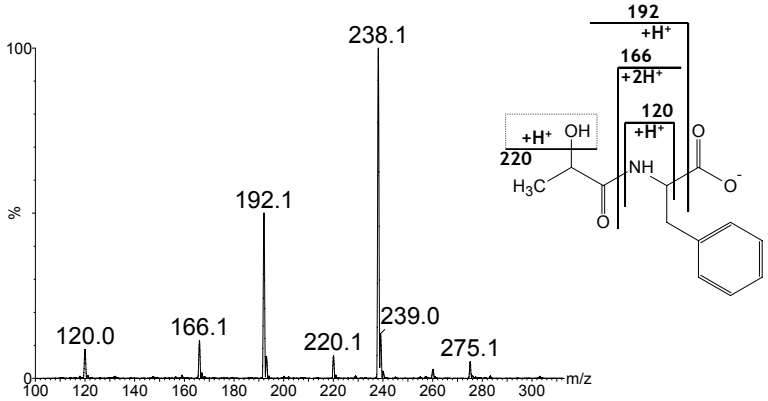


Figure 5.12. Presumed formula of peak at Rt 48.6 min and $[MH]^+$= 238.1 m/z based on mass spectral data: lactoyl-phenylalanine.

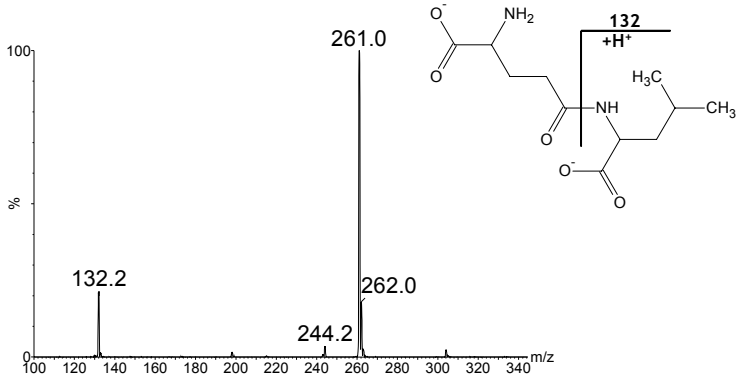


Figure 5.13.Presumed formula of peak at Rt 40.5 min and $[MH]^+$= 261.0 m/z based on mass spectral data: γ-glutamyl-leucine.

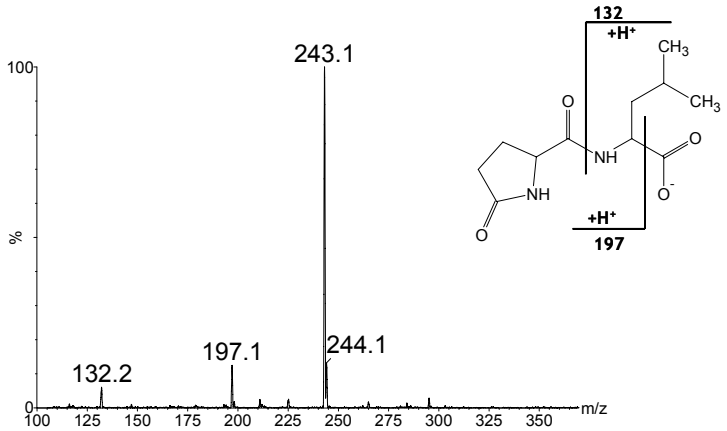


Figure 5.14.Presumed formula of peak at Rt 40.5 min and $[MH]^+$= 243.1 m/z based on mass spectral data: pyroglutamyl-leucine.

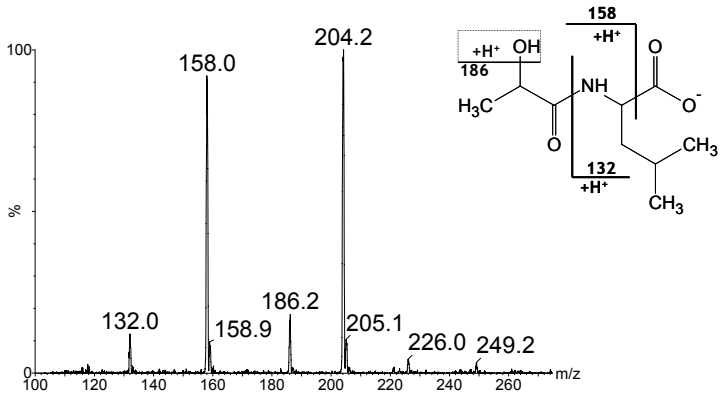


Figure 5.15.Presumed formula of peak at Rt 44.5 min and $[MH]^+$= 204.2 m/z based on mass spectral data: Lactoyl-leucine.

In order to confirm the hypothetical structures obtained by MS, it was necessary to make a comparison with the authentic specimens, also to verify the stereochemistry. In fact, if apolar amino acids present in hard cheese are known to be of proteolytic origin and therefore to have the L configuration, glutamic acid and lactic acid derive also from lactic bacteria, and they can be either D- or L-. pyroglutamic acid derives from D- and L-glutamic acid by cyclization. Thus, in order to overcome these uncertainties, some of the hypothesized molecules were synthesized. In particular, all the possible derivatives containing L-phenylalanine were obtained: γ-L-Glu-L-Phe, γ-D-Glu-L-Phe, L-Pyr-L-Phe, D-Pyr-L-Phe, L-Lac-L-Phe, D-Lac-L-Phe. The syntheses were performed by following standard protocols for peptide synthesis either in homogeneous phase or in solid phase (details are reported in the experimental section). The standards were used without any further purification and being a diastereomeric compounds in all cases were eluted at different retention times, except in the case of γ-glutamyl-phenylalanine where the difference was quite small (although still significant, Table 5.3).

Table 5.3. Retention times of synthesized diastereoisomers.

	LL	DL
	Retention time (min)	Retention time (min)
γ-Glu-L-Phe (MW 294)	34.7	34.4
Pyr-L-Phe (MW 237)	48.5	46.0
Lactoyl-L-Phe (MW 276)	45.0	44.7

The retention times of the compounds identified in the hard cheeses corresponded in all cases to the LL diastereoisomers, thus confirming the hypothesized structure and defining the stereochemistry.

MS spectra of the LL standards perfectly matched the MS spectra of the unknown compounds as reported in Figure 5.16, Figure 5.17 and Figure 5.18.

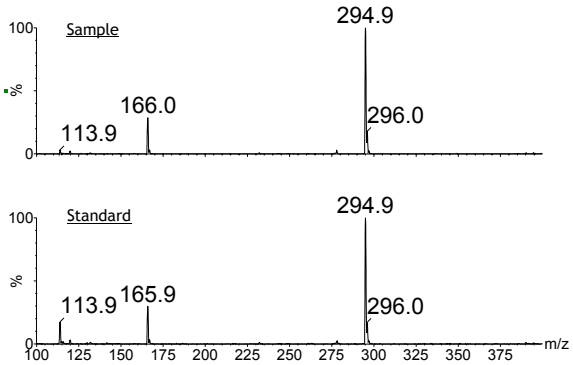


Figure 5.16. MS spectra of the peaks eluting at Rt 34.7 min in the sample and of the γ-L-Glu-L-Phe synthesized standard.

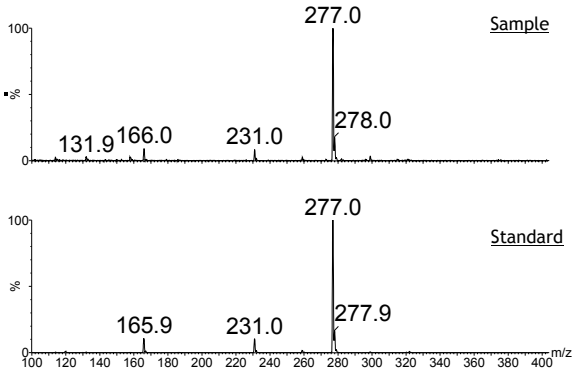


Figure 5.17. MS spectra of the peaks eluting at 45.0 min in the sample and of the L-Pyr-L-Phe synthesized standard.

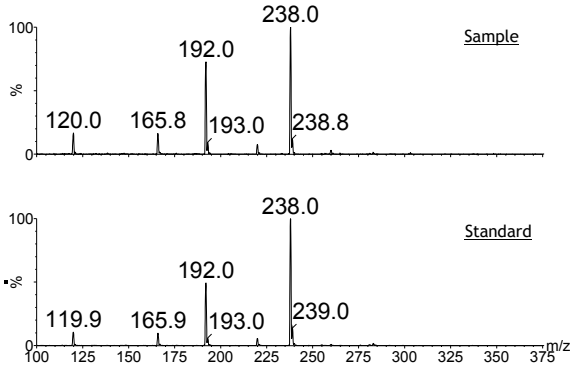


Figure 5.18. MS spectra of the peaks eluting at 48.6 min in the sample and of the L-Lac-L-Phe synthesized standard.

In order to get a definitive confirmation, samples of Parmigiano-Reggiano cheese were spiked with the standards. Spiked samples are reported in the following figures.

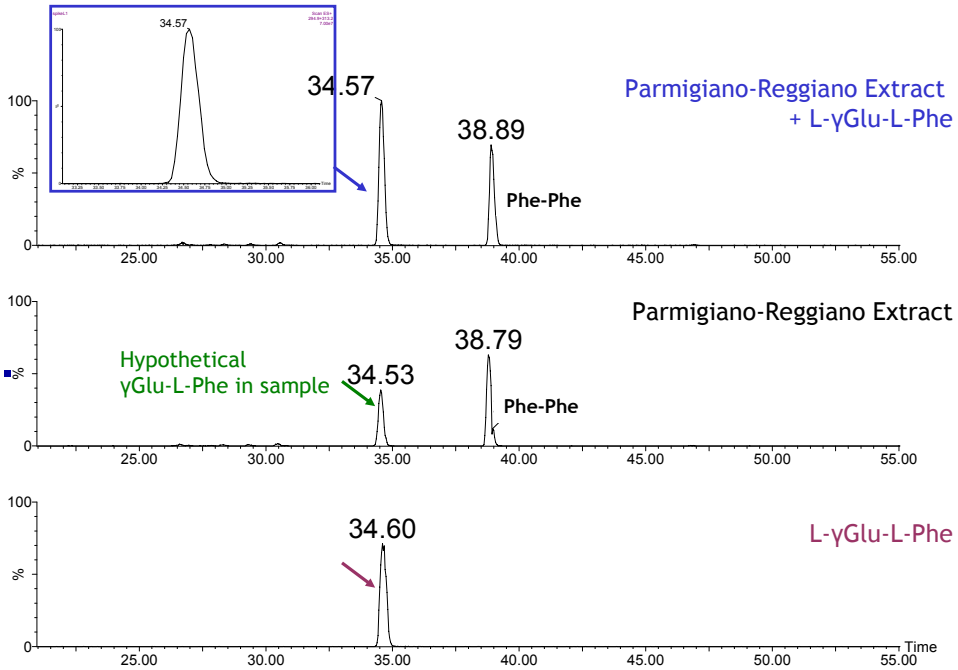


Figure 5.19. XICs chromatograms of MS ions 313.2 (Phe-Phe, internal standard) and 294.9 (γ-Glu-Phe) in the chromatograms of γ-L-Glu-L-Phe synthesized standard (bottom), sample (middle) and spiked sample (top). The unknown compound exactly corresponds to the standard.

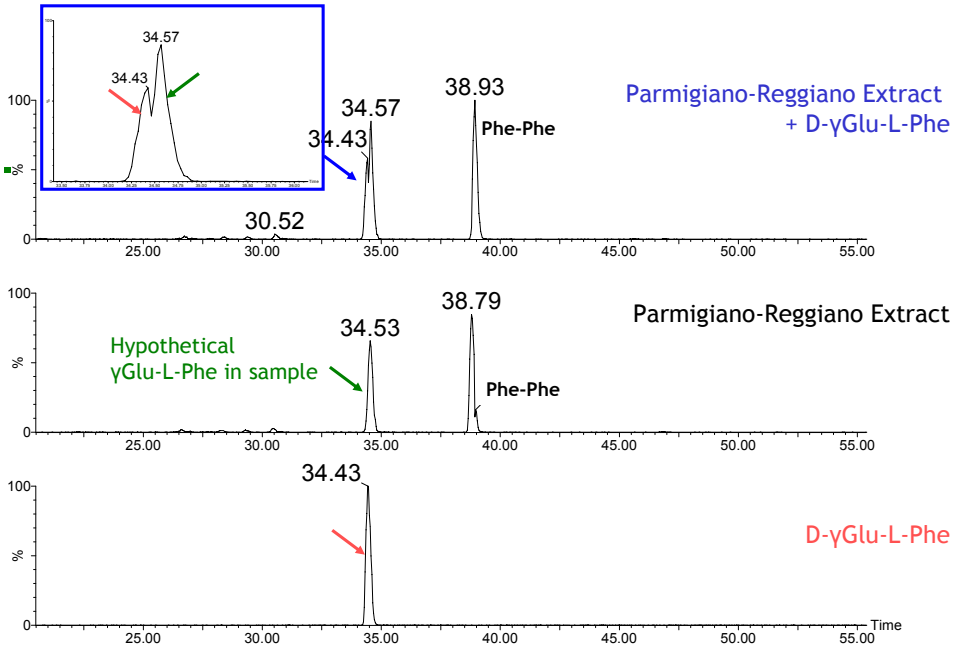


Figure 5.20. XICs chromatograms of MS ions 313.2 (Phe-Phe, internal standard) and 294.9 (γ-Glu-Phe) in the chromatograms of γ-D-Glu-L-Phe synthesized standard (bottom), sample (middle) and spiked sample (top). The unknown compound does not correspond to the standard.

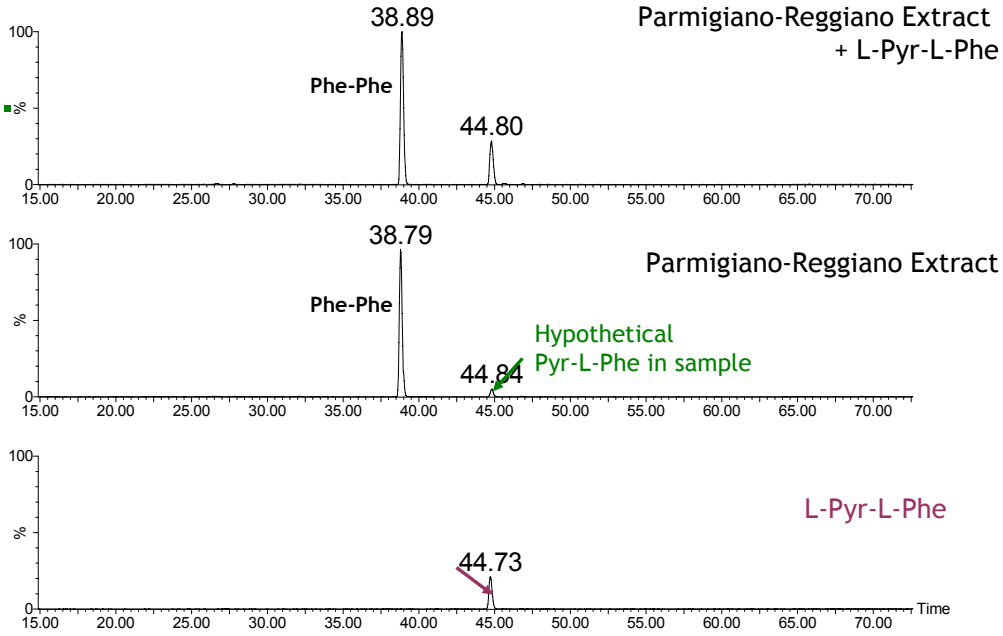


Figure 5.21. XICs chromatograms of MS ions 313.2 (Phe-Phe, internal standard) and 277.2 (Pyr-Phe) in the chromatograms of L-Pyr-L-Phe synthesized standard (bottom), sample (middle) and spiked sample (top). The unknown compound exactly corresponds to the standard.

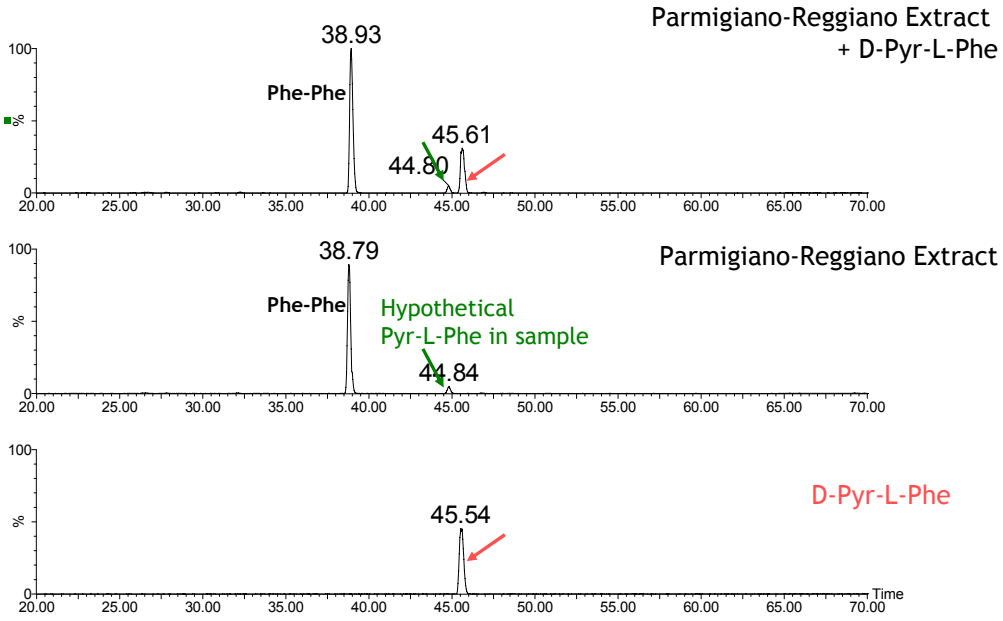


Figure 5.22. XICs chromatograms of MS ions 313.2 (Phe-Phe, internal standard) and 277.2 (Pyr-Phe) in the chromatograms of D-Pyr-L-Phe synthesized standard (bottom), sample (middle) and spiked sample (top). The unknown compound does not correspond to the standard.

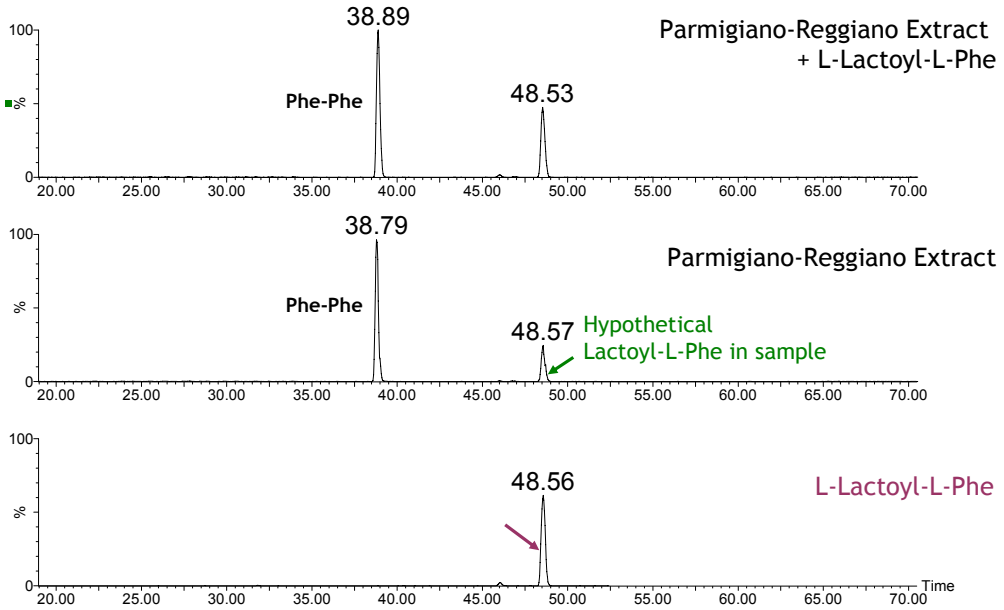


Figure 5.23. XICs chromatograms of MS ions 313.2 (Phe-Phe, internal standard) and 238.1 (Lac-Phe) in the chromatograms of L-lactoyl-L-Phe synthesized standard (bottom), sample (middle) and spiked sample (top). The unknown compound exactly corresponds to the standard.

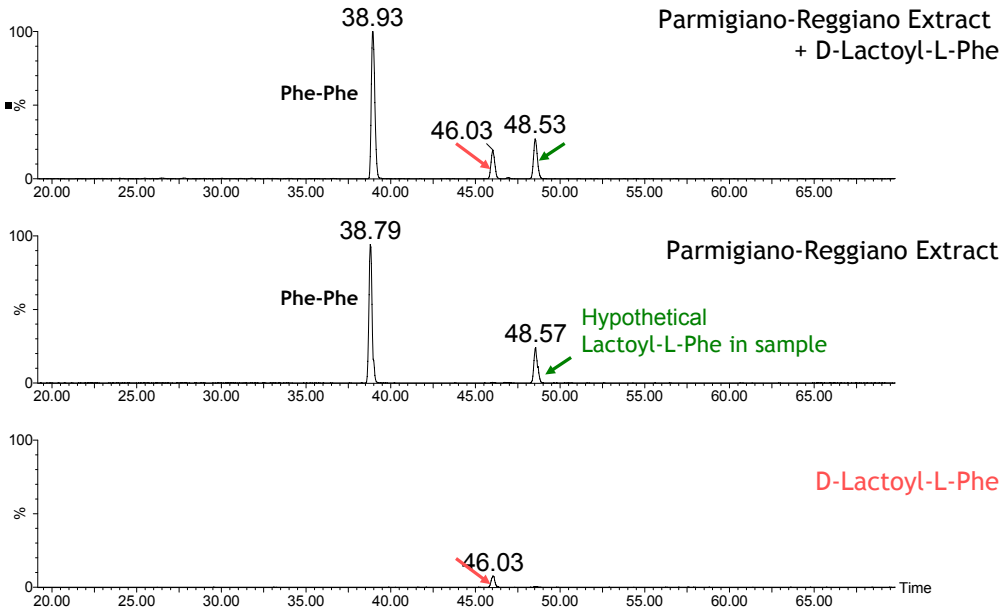


Figure 5.24. XICs chromatograms of MS ions 313.2 (Phe-Phe, internal standard) and 238.1 (Lac-Phe) in the chromatograms of D-lactoyl-L-Phe synthesized standard (bottom), sample (middle) and spiked sample (top). The unknown compound does not correspond to the standard.

Thus the identity of the unknown compounds were confirmed by spiking as L-lactoyl-L-phenylalanine, L-pyroglytamyl-L-phenylalanine and γ-L-glutamyl-L-phenylalanine. Similarly, also the other NPP attributions likely with the same stereochemistry are likely to be correct.

5.5 Conclusions and perspectives

In this work we have demonstrated the occurrence of quite unique peptide-like molecules in cheeses. The term "Non-Proteolytic Peptide-like molecules" (NPPs) was chosen in order to underline the non proteolytic origin of these compounds, since they appear to be originated from the coupling of some apolar free amino acids with pyroglutamic acid, lactic acid or with the γ-carboxylic group of glutamic acid.

It is quite tempting to speculate on a common origin, very likely enzymatic, of the NPPs identified. Several data are actually consistent with this hypothesis. First, they all derive from the same reaction, the coupling of a carboxylic acid to a nitrogen compound to form a new amide bond. Moreover, the enzymatic origin is indicated by the common nature of the nitrogen donor compounds (all apolar amino acids: Phe, Val, Met, Ile, Leu, Tyr) and from the conserved stereochemistry of the carboxylic acids (all L, although D-analogues are known to be present in hard cheeses). These enzymes likely derives from the (intracellular) pool of enzymes of thermophilic LAB.[17]

There are many possible consequences from the production of these molecules. First, the removal of carboxylic moieties from the medium might increase the pH, thus affecting both enzymatic activities and taste; moreover, this reaction sequestrates bitter amino acids from the medium, transforming them in derivatives with a different, or more "umami" taste.

5.6 References

[1] McSweeney P.L.H., Sousa M.J., Biochemical pathways for the production of flavour compounds in cheeses during ripening: A review, *Lait* **2000** 80; 293–324

[2] Fox, P.F. Proteolysis during cheese manufacture and ripening. *Journal of Dairy Science* **1989** 72, 1379-1400.

[3] Leszek Stepaniak, Dairy enzymology *International Journal of Dairy Technology* **2004** 57, 153-171

[4] M.J. Sousa, Y. Ardo, P.L.H. McSweeney, Advances in the study of proteolysis during cheese ripening, *International Dairy Journal* , **2001**, 11: 327–345

[5] Andreas Dunkel, Jessica Köster, Thomas Hofmann, Molecular and Sensory Characterization of gamma-Glutamyl Peptides as Key Contributors to the Kokumi Taste of Edible Beans (*Phaseolus vulgaris L.*), *J Agric Food Chem.* **2007** 55, 16:6712-9.

[6] F. Roudot-Algaron, L. Kerhoas,' D. Le Bars, J. Einhorn.L And J. C. Gripon Isolation of γ-Glutamyl Peptides from Comte Cheese *J Dairy Sci* **1994** 77:1161-1166,

[7] Stefano Sforza, Gianni Galaverna, Cristina Schivazappa, Rosangela Marchelli, Arnaldo Dossena, and Roberta Virgili, Effect of Extended Aging of Parma Dry-Cured Ham on the Content of Oligopeptides and Free Amino Acids, *J. Agric. Food Chem.* **2006**, 54, 25, 9422 -9429,.

[8] Mutsch-Eckner M, Meier B, Wright AD, Sticher O. Gamma-glytamyl peptides from allium sativum bulbs. *Phytochemistry.* **1992** 31, 7: 2389-91.

[9] Martin L Shaw, Jane E Lancaster, Geoffrey A Lane Quantitative analysis of the major γ glutamyl peptides in onion bulbs (*Allium cepa*), *Journal of the Science of Food and Agriculture* **2006,** 48, 4 , 459 – 467

[10] Thomas Frank Hofmann & Andreas Dunkel, KOKUMI FLAVOUR COMPOUNDS AND USE *Publication Number: WO/2007/042273 International Application No.: PCT/EP2006/009817*, Publication Date: 19.04.**2007**,

[11] G. Mucchetti, F. Locci, P. Massara, R. Vitale, and E. Neviani, Production of Pyroglutamic Acid by Thermophilic Lactic Acid Bacteria in Hard-cooked Mini-Cheeses *, J. Dairy Sci.* **2002.**85:2489–2496,

[12] Higaki-Sato Noriko; Sato Kenji ; Esumi Yasuaki; Okumura Takeshi; Yoshikawa Hideki; Tanaka-Kuwajima Chie; Kurata Akie; Kotaru Makoto; Kawabata Makoto; Nakamura Yasushi; Ohtsuki Kozo; Isolation and identification of indigestible pyroglutamyl peptides in an enzymatic hydrolysate of wheat gluten prepared on an industrial scale, *Journal of agricultural and food chemistry,* **2003,** 51, 1, 8-13

[13] M. R. Altamura, R. E. Andreotti , L. Longjr, Pyroglutamyl dipeptides in mushroom, *Agaricus Campestris*, *Journal of Food Science* **1970** 35 (2), 134–139.

[14] Schlichtherle-Cerny, Hedwig, Amado, Renato, Applicants: SOCIETE DES PRODUITS NESTLE S.A., Flavouring compositions, *WO/2003/043444, PCT/EP2002/012720,* **2003**

[15] Sforza, S.; Galaverna, G.; Neviani, E.; Pinelli, C.; Dossena, A.; Marchelli, R. Study of the oligopeptide fraction in Grana Padano and Parmigiano-Reggiano cheeses by liquid chromatography-electrospray ionization mass spectrometry. *Eur. J. Mass Spec.,* **2004,** 10, 421-427.

[16] Stefano Sforza, Linda Ferroni, Gianni Galaverna, Arnaldo Dossena, And Rosangela Marchelli, Extraction, Semi-Quantification, and Fast On-line Identification of Oligopeptides in Grana Padano Cheese by HPLC-MS *J. Agric. Food Chem.* **2003**, 51, 2130-2135,

[17] G. Mucchetti, F. Locci, P. Massara, R. Vitale, and E. Neviani. Production of Pyroglutamic Acid by Thermophilic Lactic Acid Bacteria in Hard-cooked Mini-Cheeses, *J. Dairy Sci.* **2002** 85:2489–2496,

6 OLIGOPEPTIDES IN ASIAGO CHEESE: MARKERS OF TECHNOLOGY AND TIPICALITY

6.1 Introduction

Asiago is an Italian, semi-fat and semi-cooked cheese, made only from cow's milk. It bears the Protected Designation of Origin (PDO, EU Reg. 2081/92), which means that the product can be considered as authentic according to the European law only if it is produced in its specific geographic area, and according to a specific protocol, known as the "disciplinare di produzione". Each wheel of the Asiago cheese is guaranteed by the "Consorzio di Tutela", the Regulatory Board for the Supervision of the Production of Asiago Cheese.[1]

There are different types of Asiago cheese, with different textures, ageing times and also different methods of production, according to the PDO disciplinary: Asiago "Fresco" or "Pressato" (sweet, smooth, fresh milk-like flavored, ripened from 20 to 40 days) and aged Asiago "d'Allevo", with a very concentrated, complex variety of flavors ranging from the fruity, to the nutty, sharp, piquant and pungent. The aged cheese is classified, according to the ripening time in "Mezzano" (4 to 6 months), "Vecchio" (more than 10 months) and "Stravecchio" (more than 15 months).

The Asiago cheese has been traditionally produced for hundreds of years in the alpine area of "Altopiano di Asiago", an Alpine plateau characterized by meadows. The milk-collection areas and the specified traditional zones of production, which have been established by the disciplinary, are defined as the entire provinces of Vicenza and Trento, and two areas of the lowland provinces of Treviso and Padova, in the north-eastern region of Italy.

The Asiago is generally produced in industrial dairies, but it is also still produced in "*coops*" (*malghe*), specialized groups of local dairies which provide grass-fed cow's milk[2]. The grazing of dairy cows during the summer takes often place in alpine pastures. Alpine meadows have a larger variety of grass species, medicinal plants and flowers all of which contribute to a better tasting milk with a higher protein content.[3,4] Asiago cheese is produced from two milkings, one of which is partially skimmed overnight. The raw milk is heated to a temperature of about 35°C ± 2°C, and rennet is added. The half-cooking process is continued until the temperature is 46°C ± 2°C, which takes about 10-12 minutes to be reached. After half an hour, the resulting curd is cut with a tool called a "cipollina" until the cheese curd is broken up into granules; then, the curd is extracted and the cheese is portioned. The portions are placed into molds lined with canvas so that the cheese rind acquires the typical textile texture. After a few hour rest (3-5 h) on tables, the cheese pieces are turned up side down. The product is then sent to the pre-salting area, where the removal of whey is completed and the forms are marked with the seals

of the regulatory board. The whole process is completed in three to five days, meanwhile the forms are turned various times. The salting operation can be done either with the surface dry-salting technique or with brine. The final phase is the ageing period, in which the cheese forms are left in storage bays to ripen. At this point, the important factors are temperature (11-14 °C) and the degree of humidity (80-85 %). The cheese is aged for a period of up to two years. During the ageing period the rind is often scraped and oiled.

As reported in the previous chapter, the starter is essential for determining the characteristics of the final product.[5] While Parmigiano-Reggiano cheese must use the natural whey starter of the previous day, Asiago d'Allevo disciplinary of production does not give specific restrictions, and each plant can choose if using a commercial starter, a natural milk-based starter or exploiting natural acidification. Traditional productions avoid the use of the starter. In this latter case, Non-starter microflora plays an important, scarcely controlled, role. Non-starter lactic acid bacteria (NSLAB) are lactic acid bacteria that are not part of the normal starter bacteria added to the milk for cheese making.[6,7,8] NSLAB are (for the most part) *Lactobacilli* , *Enterococci* and (occasionally) *Pediococci* and *Leuconostocs*. In most cheeses (regardless whether they are made from raw or pasteurized milk), the dominant organisms are homofermentative lactobacilli, especially *Lactobacillus casei*, *Lactobacillus plantarum*, and *Lactobacillus curvatus*. Obligate heterofermenters, particularly *Lactobacillus brevis*, are found in lower numbers.[9]

6.2 Aim of the work

In this work, we intended to study the oligopeptide fraction of Asiago cheese during ageing in order to make correlations with the different techniques of production. Samples of Asiago d'Allevo (PDO and non PDO) produced both in dairy or in "coops", have been analyzed after 6, 12 and 18 months of ageing. By means of HPLC-MS techniques, the oligopeptide fraction of the cheese has been characterized.

6.3 Experimental part

6.3.1 Solvents and reagents

See section 4.3.1.

6.3.2 Instrumentation

See section 4.3.2.

The column used was a PHENOMENEX GEMINI C18 column (5 µm, 110 Å, 250 × 4.6 mm).

6.3.3 Samples

Wheels of Asiago cheese were produced in six plants (A, B, C, D, E and F) in three different months, May (1), July (2) and September (3), of the summer 2005. Samples produced in May

derived from milk obtained from cows reared in stables; samples produced in July and September derived from milk obtained from cows reared in alpine pastures and eating meadow grass. Samples analyzed in this work are reported in Table 6.1. Factories have different techniques of production, principally reviewed in Table 6.2.

Table 6.1. Samples used in this work.

Sample	Moisture %	Production time	Months of ageing	Factory
2	32.460	1	6	B
7	31.578	1	6	A
9	31.574	1	6	D
1	31.723	2	6	A
4	34.439	2	6	C
5	32.900	2	6	B
11	33.051	2	6	D
13	34.908	2	6	F
14	34.247	2	6	E
18	30.361	3	6	A
21	33.121	3	6	B
22	29.871	3	6	C
25	28.440	3	6	F
26	33.632	3	6	E
28	27.093	1	12	D
29	27.954	1	12	B
30	28.340	1	12	A
33	28.792	2	12	B
36	30.843	2	12	C
37	30.117	2	12	F
40	28.711	2	12	A
43	28.532	2	12	D
47	32.257	2	12	E
51	29.800	3	12	E
54	28.521	3	12	C
57	29.221	3	12	B
62	27.587	3	12	A
64	29.607	3	12	F
69	26.070	1	18	A
70	25.290	1	18	D
71	26.200	1	18	B
67	26.630	2	18	E
72	27.540	2	18	F
74	27.200	2	18	C
76	25.680	2	18	A
80	26.930	2	18	D
82	26.499	2	18	B
84	26.025	3	18	E
86	26.452	3	18	F
90	27.044	3	18	A
94	26.304	3	18	C

Table 6.2. Technologies of production and caseification parameters in the different dairies.

Factory	STARTER	Cooking process temperature	Curd pH	Salting	Production	Ageing	PDO
A	Yes	45 °C	6.35	Brine	dairy	dairy	
B	Yes	45.5 °C	6.3	Brine	coop	dairy	
C	No	48 °C	6.3	Brine	coop	dairy	*
D	Yes	45 °C	6.35	Brine	dairy	dairy	
E	Yes	46 °C	5.85	Brine	coop	coop	
F	No	47.5 °C	6.5	Salt	coop	dairy	*

* These plants can label their product as PDO.

6.3.4 Extraction and concentration of the oligopeptide fraction

See section 4.3.4.

6.3.5 LC/MS analysis of the oligopeptide fraction

See section 5.3.5.

6.3.6 Data analysis

See section 4.3.6. Moisture data for each sample are reported in Table 6.1.

6.3.7 Statistical analysis

See section 4.3.7.

Discriminant Analysis (DA) was performed with the Wilks lambda method, with a stepwise strategy: at each step, the variable that mostly contribute to the separation of the groups was entered into the discriminant functions. Wilks lambda is used to test the null hypothesis that the populations have identical means on discriminant function.

The criteria for entry and removal was based on probability p of full hypothese (entry p = 0.05, removal p= 0.1) and were set as defaults. Ageing time, month of production and dairy of origin were chosen as discriminant factors.

6.4 Results and discussion

6.4.1 HPLC/MS characterization of Asiago cheese peptides.

The peptide profile was analyzed in the different samples, at the different ageing times. TIC chromatograms were very different, as it is evident in Figure 6.1, and in particular the amounts of peptides appeared to be lower at higher ageing times.

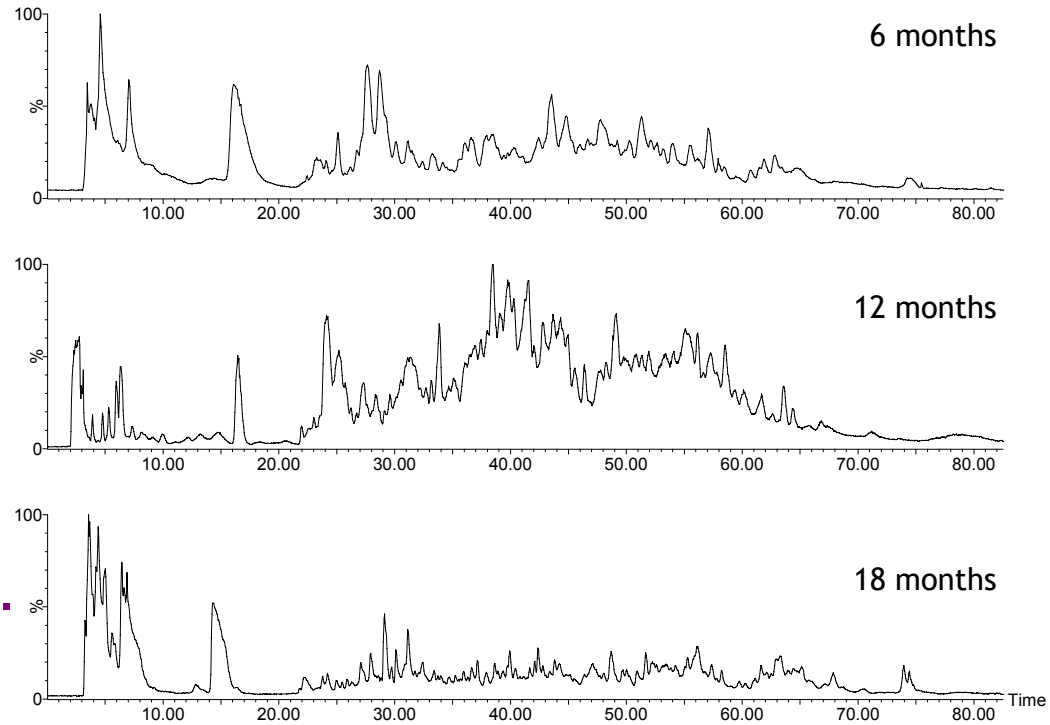


Figure 6.1. HPLC/MS profiles of an Asiago cheese at different ageing times.

Several peptides were traced and identified in the different samples, and they are reported in Table 6.3.

Table 6.3. List of peptides present in Asiago cheeses.

CODE	Identification	MW (Da)	Characteristic ions (m/z)	Rt (min)
p1	Lac-Val	**189**	144.1 (i_1) + 190.2 (MH^+)	36.6
p2	Lac-Ile	**203**	158.2 (i_1) + 204.2 (MH^+)	44.2
p3	Lac-Leu	**203**	158.2 (i_1) + 204.2 (MH^+)	44.8
p4	Lac-Met	**221**	222.1 (MH^+) + 176.1 (i_1)	35.8
p5	Lac-Phe	**237**	238.1 (MH^+) + 192.1 (i_1) + 120 (i_2)	48.6
p6	Pyr-Ile	**243**	243.2 (MH^+) + 197.1 (i_1) + 132 (y_1)	38.7

CODE	Identification	MW (Da)	Characteristic ions (m/z)	Rt (min)
p7	Pyr-Leu	**243**	243.2 (MH^+) + 197.1 (i_1) + 132 (y_1)	40.5
p8	γ Glu-Val	**246**	247.2 (MH^+) + 118.1 (y_1)	18.7
p9	Lac-Tyr	**253**	254.2 (MH^+) + 208.2 (i_1) + 136.1 (i_2)	37.3
p10	γ Glu-Ile	**260**	261.2 (MH^+) + 132.1 (y_1)	30.7
p11	γ-Glu-Leu	**260**	261.2 (MH^+) + 132.1 (y_1)	31.7
p12	Pyr-Phe	**276**	277.2 (MH^+) + 231.1 (i_1) + 166(y_1) + 120 (i_2)	44.8
p13	γ Glu-Met	**278**	279.3 (MH^+) + 150.1 (y_1)	21.0
p14	γ Glu-Phe	**294**	295.2 (MH^+) + 166 (y_1)+ 120 (i_2)	34.9
p15	γ Glu-Tyr	**310**	311.0 (MH^+) + 182.0 (y_1)	28.4
p16	β CN f(4-6)	**389**	390.1 (MH^+)	27.5
p17	α S1 CN f(18-21)	**487**	488.4 (MH^+) + 357.4 (b_3)	38.0
p18	α S1CN f(27-30)	**489**	489.3 (MH^+) + 342.3 (y_3)	29.5
p19	β CN f(3-6)	**502**	503.3 (MH^+) + 372.3 (b_3)	36.0
p20	β CN f(10-14)	**545**	546.4 (MH^+) + 399.3 (b_4)	25.5
p21	α S1CN f(17-21)	**601**	602.4 (MH^+)	49.0
p22	β CN f(2-6)	**631**	632.4 (MH^+) + 501.4 (b_4)	38.9
p23	β CN f(47-52)	**755**	756.4 (MH^+) + 378.9 (MH_2^{2+})	30.9
p24	β CN f(1-6)	**787**	788.4 (MH^+) + 657.3 (b_5)	33.3
p25	α S1 CN f (24-30)	**805**	806.4 (MH^+) + 403.7 (MH_2^{2+})	50.8
p26	NI*	**850**	425.8 (MH_2^{2+}) + 851.0 (MH^+)	26.6
p27	β CN f(7-14)	**855**	856.5 (MH^+) + 709.5 (b_7) + 643.4 (y_6)	38.1
p28	αS1 CN f (17-23)	**905**	905.6 (MH^+) + 453.5 (MH_2^{2+})	51.7
p29	αS1 CN f(80-87)	**964**	965.4 (MH^+) + 751.1 (b_6)	43.0
p30	NI	**1025**	513.5 (MH_2^{2+}) + 1025.6 (MH^+)	47.5
p31	αS1 CN (f24-32)	**1052**	1052.6 (MH^+) + 527.5 (MH_2^{2+})	56.6
p32	NI	**1088**	363.6 (MH_3^{3+}) + 545.1 (MH_2^{2+})	22.8
p33	αS1 f(111-119)1P	**1094**	548 (MH_2^{2+})+ 365.5 (MH_3^{3+})	33.3
p34	β CN f(84-93)	**1156**	579.0 (MH_2^{2+})+ 1156.7 (MH^+)	51.8
p35	α S1 CN f (24-34)	**1237**	1237.7 (MH^+) + 619.6 (MH_2^{2+})+ 413.0 (MH_3^{3+})	48.9
p36	β CN f(83-93)	**1255**	628.5 (MH_2^{2+})+ 1255.8 (MH^+)	53.5
p37	αS1 CN f(3-13)	**1283**	642.5 (MH_2^{2+})	26.3
p38	αS1 CN f(2-13)	**1380**	460.9 (MH_3^{3+})+ 690.9 (MH_2^{2+})	24.3
p39	α S1 CN f(2-14)	**1509**	503.9 (MH_3^{3+})+ 755.0 (MH_2^{2+})	25.6
p40	β CN f (82-95)	**1510**	756.1 (MH_2^{2+})+ 504.3 (MH_3^{3+})	35.2
p41	αS1 CN f(1-13)	**1536**	385.0 (MH_4^{4+}) + 513.0 (MH_3^{3+})+ 768.7 (MH_2^{2+})	23.5
p42	β CN f (193-206)	**1555**	1556.8 (MH^+) + 778.7 (MH_2^{2+})+ 519.6 (MH_3^{3+})	46.0
p43	β CN f(17-28) 3P	**1590**	795.8 (MH_2^{2+})+ 530.9 (MH_3^{3+})	39.9
p44	NI	**1601**	801.4 (MH_2^{2+})	29.6
p45	αS1 CN f(1-14)	**1665**	555.9 (MH_3^{3+}) + 833.1 (MH_2^{2+})	24.8
p46	αS1 CN f(24-38)	**1708**	570.3 (MH_3^{3+}) + 854.5 (MH_2^{2+})	44.9

CODE	Identification	MW (Da)	Characteristic ions (m/z)	Rt (min)
p47	α S1CN f(2-16)	**1721**	574.6 (MH_3^{3+}) + 861.2 (MH_2^{2+})	34.7
p48	ß CN f(193-208)	**1781**	891.7 (MH_2^{2+}) + 1781.9(MH^+)	52.1
p49	ß CN f (15-28)3P	**1790**	896.0 (MH_2^{2+})+ 597.7 (MH_3^{3+})	43.2
p50	α S1 CN f(1-16)	**1877**	626.7 (MH_3^{3+})+ 939.6 (MH_2^{2+})+ 470.0 (MH_4^{4+})	31.9
p51	ß CN f(193-209)	**1881**	941.2 (MH_2^{2+}) + 627.7(MH_3^{3+})	53.6
p52	α S1 CN f(1-17)	**1991**	664.7 (MH_3^{3+}) + 996.2 (MH_2^{2+})	31.6
p53	ß CN f(12-28) 4P	**2212**	1106.7 (MH_2^{2+})+ 738.1 (MH_3^{3+})+ 553.8 (MH_4^{4+})	49.7
p54	αS2 CN f(189-207)	**2332**	467.4 (MH_5^{5+}) + 583.9 (MH_4^{4+})+ 778.3 (MH_3^{3+})	42.9
p55	ß CN f(11-28) 4P	**2340**	1170.9 (MH_2^{2+})+ 780.8 (MH_3^{3+})+ 585.8 (MH_4^{4+})	43.3
p56	αS1CN f(1-20)	**2347**	783.5 (MH_3^{3+})+ 587.8 (MH_4^{4+})+ 1174.8 (MH_2^{2+})	36.0
p57	α S1 CN f(1-21)	**2461**	616.4 (MH_4^{4+}) +821.2 (MH_3^{3+}) + 1231.8 (MH_2^{2+})	39.7
p58	ß CN f(8-28)4P	**2594**	1298.0 (MH_2^{2+}) + 865.5 (MH_3^{3+}) + 649.7 (MH_4^{4+})	54.9
p59	αS1 CN f(1-22)	**2617**	655.3 (MH_4^{4+}) + 873.4 (MH_3^{3+})+ 1309.3 (MH_2^{2+})	38.3
p60	NI	**2627**	1314.7 (MH_2^{2+}) + 876.7 (MH_3^{3+})	51.1
p61	αS1 CN f(6-28)	**2709**	903.6 (MH_3^{3+}) + 1354.9 (MH_2^{2+})	55.0
p62	αS1 CN f(1-23)	**2764**	550.0 (MH_5^{5+}) + 692.0 (MH_4^{4+})+ 922.1 (MH_3^{3+})	43.5
p63	αS1 CN f(89-110)	**2780**	695.8 (MH_4^{4+}) + 927.5 (MH_3^{3+})	40.3
p64	αS1CN f(1-24)	**2908**	728.0 (MH_4^{4+}) + 970.0 (MH_3^{3+}) + 1454.6 (MH_2^{2+}) + 582.8 (MH_5^{5+})	38.4
p65	αS2 CN f(183-207)	**3116**	520.3 (MH_6^{6+}) + 624.3 (MH_5^{5+}) + 780.4 (MH_4^{4+})	41.7
p66	ß CN f(98-124)	**3133**	1567.3 (MH_2^{2+}) + 1044.9 (MH_3^{3+}) + 784.3 (MH_4^{4+})	41.0
p67	αS2 CN f(182-207)	**3217**	537.2 (MH_6^{6+})+ 644.4 (MH_5^{5+})+ 805.2 (MH_4^{4+})+ 1073.1 (MH_3^{3+})	40.8
p68	NI	**3594**	1199.2 (MH_3^{3+}) + 899.6 (MH_4^{4+})+ 719.8 (MH_5^{5+})	63.6
p69	NI	**3707**	1236.5 (MH_3^{3+}) + 927.8 (MH_4^{4+}) + 742.5 (MH_5^{5+})	64.9
p70	NI	**3824**	1275.9 (MH_3^{3+}) + 957.1 (MH_4^{4+}) + 765.8 (MH_5^{5+})	66.1
p71	NI	**3864**	1289.1 (MH_3^{3+}) + 967.0 (MH_4^{4+}) + 774 (MH_5^{5+})	59.8
p72	ß CN f(59-96)	**4024**	1342.7 (MH_3^{3+}) + 1007.1 (MH_4^{4+}) + 805.8 (MH_5^{5+})	67.9
p73	NI	**4035**	1346.1 (MH_3^{3+}) + 1009.5 (MH_4^{4+}) + 808.0 (MH_5^{5+})	60.3
p74	ß CN f(57-93)	**4065**	1355.9 (MH_3^{3+}) + 1017.0 (MH_4^{4+})	60.8
p75	αS1CN f(1-36)	**4235**	706.6 (MH_6^{6+})+ 848.0 (MH_5^{5+})+ 1059.5 (MH_4^{4+})+ 1412.6 (MH_3^{3+})	56.4
p76	NI	**4736**	1185.2 (MH_4^{4+}) + 948.3 (MH_5^{5+})+ 790.6 (MH_6^{6+})	58.0
	Phe-Phe	**312**	313.2 (MH^+) +166.0 (y_1) +120.0 (i_2)	38.5

* "NI" means "Non Identified" peptides

Among these peptides, 11 peptides were not identified, but their signals were so strong that they were considered important for sample profiles and descriptions. 15 specimens were identified as non proteolytic peptide-like compounds (NPPs) discussed in the previous two chapters, although their signals were often weak. Three peptides originate from αS2 casein, and

all the other identified peptides come from αS1 or ß casein. The identification of the peptides has been made possible by comparison with literature data on the peptide fraction of other cheeses[10,11,12,13,14,15] and the in-source fragmentation, according to the method previously presented[16]. It is the first time that the oligopeptide fraction of Asiago cheese has been characterized.

The casein portions originating the observed peptides are depicted in the figures here below. 45 of the 47 peptides identified derive from the casein regions reported below.

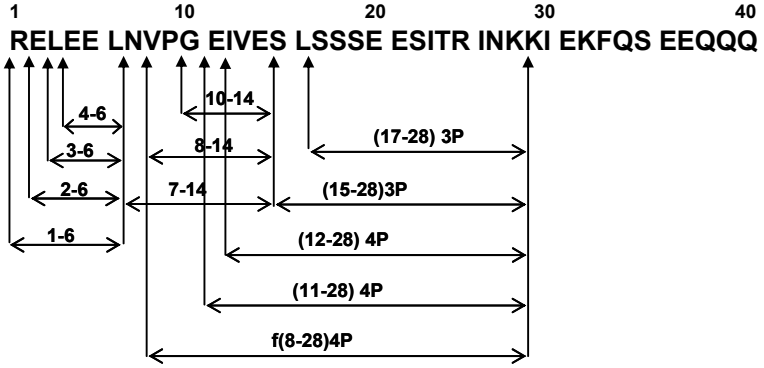


Figure 6.2. Peptides found in Asiago cheeses coming from the N-terminal part of ß casein.

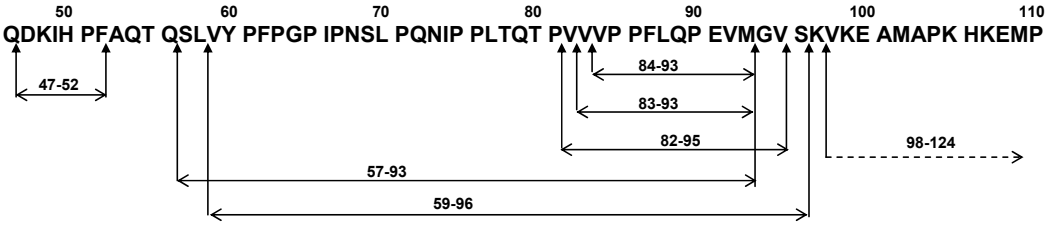


Figure 6.3. Peptides found in Asiago cheeses coming from the median part of ß casein.

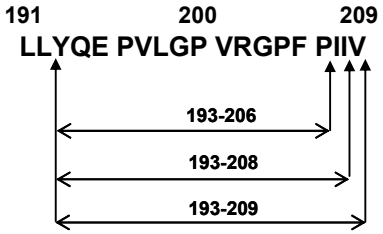


Figure 6.4. Peptides found in Asiago cheeses coming from the C-terminal part of ß casein.

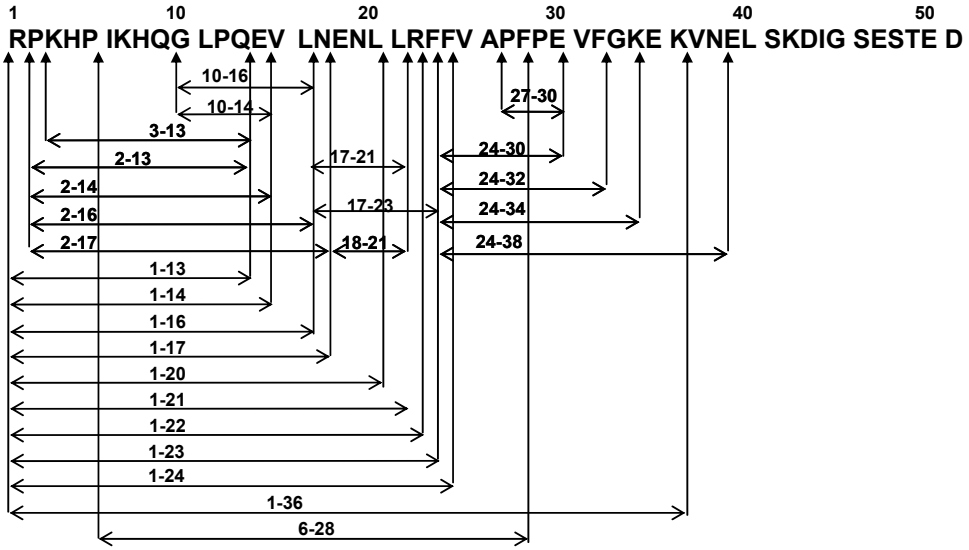


Figure 6.5. Peptides found in Asiago cheeses coming from the N-terminal part of α S1 casein.

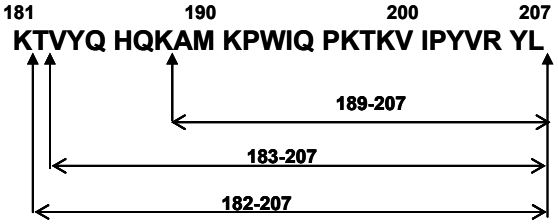


Figure 6.6.Peptides found in Asiago cheeses coming from the C-terminal part of α S2 casein.

Although many C-terminal ends and N-terminal ends corresponded to cleavage sites of enzymes well known to be active in cheeses, several cleavage sites of unknown origin were present, such as Arg_1-Pro_2, Pro_2-Lys_3 in α S1 casein, Arg_1-Glu_2, Glu_2-Leu_3, Leu_3-Glu_4 and Pro_{81}-Val_{82}, Val_{82}-Val_{83}, Val_{83}-Val_{84} in ß casein. This is likely to be related to the complexity and the variability of the microflora characterizing these cheeses.

6.4.2 Principal Component Analysis

All the semiquantitative data concerning the 76 peptides in the 41 samples were analyzed by Principal Component Analysis, in order to have a general overview of the factors affecting the data variability.[17]

The Loading plot is reported in Figure 6.7: peptides are represented in a 2-dimensional plan generated by the two calculated main factors (component 1 and component 2) which explain, respectively, 28.9 and 19.3% of the variance.

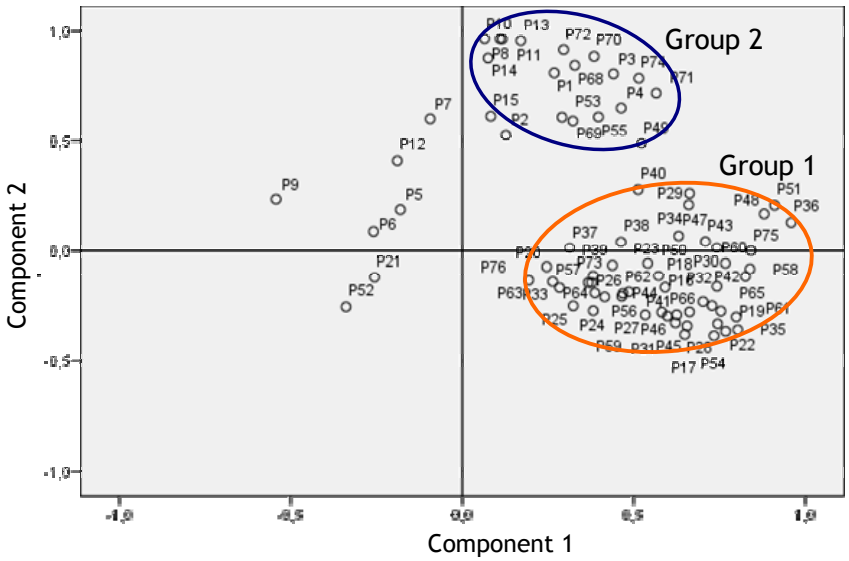


Figure 6.7. Loading plot of peptides (named from P1 to P76) in Asiago cheeses.

From the loading plot, peptides appeared to be clustered into two main groups. Group 1 includes most peptides, whereas group 2 includes some NPPs and several high molecular weight peptides.

Each factor of the production technology, i.e. time, plant, etc., can influence the peptidic profile. In fact, each dairy/coop has its own breed fed, its own kind of pasture/fodder, and its technique of caseification. As an example, dairies A and D collect the milk from their cows in pastures and bring it to the Asiago social cheese plant. Moreover, each period of production is correlated to seasonal variations of temperature and humidity, which affect pasture herbs composition and temperature of the rooms of caseification and ageing (if the conditions are not controlled); the period of production also influences the different energetic request of cows, which is higher if they graze. Within these months, the cows' feed is very different too: in May, coops keep their cows in stables and consequently the feed is based on hay or silage; on the contrary, in July and September, the feed comes from the paddock which is harvested on the same day or eaten on the premises by the cows.

Each ripening period also bears multiple effects too: actually, some plants age the wheels in controlled conditions, coops B, C and F age the wheels in their own rooms till November, then move the wheels to the Asiago social cheese factory. One coop (E) keeps the wheels in a ageing room till the sale.

The score plot including all the analyzed samples and derived from the calculation of the factor score of every sample for each component is reported in figures here below. Samples have been labeled according to their ageing times (in months) in Figure 6.8, to their dairy of origin in Figure 6.9 and to the period of production in Figure 6.10.

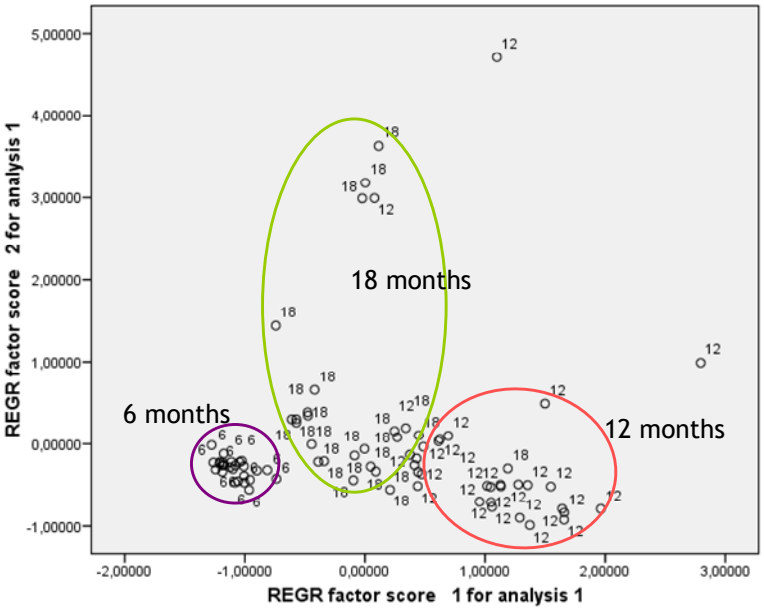


Figure 6.8. Score plot of the Asiago cheese analyzed samples, labeled and regrouped (where possible) by the ageing period (expressed in months).

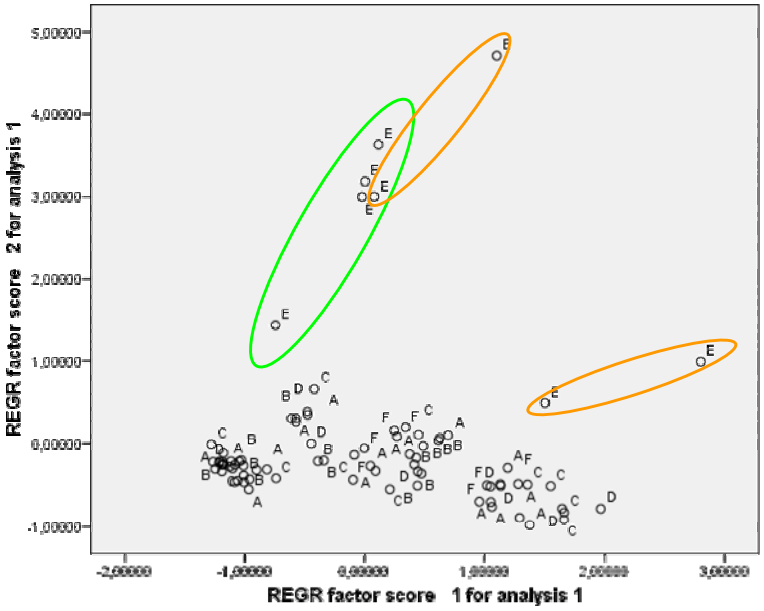


Figure 6.9. Score plot of the Asiago cheese analyzed samples, labeled by the dairy or coop of production (see Table 6.1) and re-grouped by the ageing period (as in Figure 6.8).

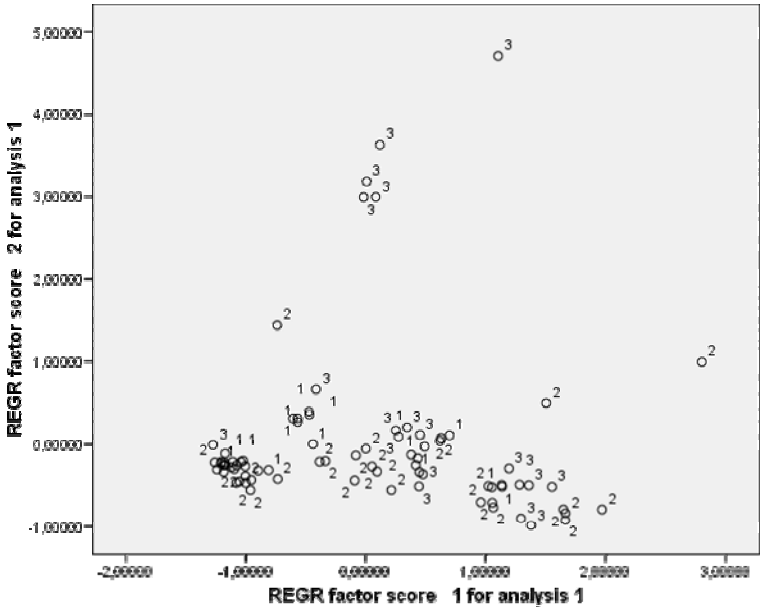


Figure 6.10. Score plot of the Asiago cheese analyzed samples, labeled by the month of production (1=May, 2=July, 3=September) and re-grouped by the ageing period (as in Figure 6.8).

The score plot in Figure 6.8 indicates that one of the most important factors affecting sample variability is the ageing time.

In any case, it is quite clear that, although most samples tend to cluster according to their ageing time, few samples having high positive score along PC2 appear to be very different from all the others, indicating that the differentiating factor is not the ageing period.

By considering the same score plot with the samples labeled according to the different factories of production (Figure 6.9), it is very clear that all the anomalous samples belong to factory E. The main difference between coop E and the other plants is the ageing conditions: in coop E ageing takes place in its own ageing rooms, in natural conditions, while in the other production sites it takes place in rooms at controlled moisture and temperature. It is very likely that the natural ambient microflora affects the proteolysis in factory E, thus modifying the peptide pattern. The month in which the cheeses are produced does not appear to strongly influence the variability of the samples (Figure 6.10). Analogously, in these conditions, samples produced by addition of starter bacteria appeared to be quite similar to samples made without the starter, or the variables which mostly affect the two main components are different.

6.4.3 Discriminant analysis

Discriminant Analysis (DA) was also performed in order to classify the samples according to the different production technologies and also to verify the peptides more influenced by the different ways of production. This analysis is used to classify or assign objects (e.g. individual

cheeses) to a small number of groups (i.e. ageing time or caseification month) that are predicted by a number of variables that should be uncorrelated, in order to achieve classification.[17] The attempted classifications were based on ageing time (Figure 6.11), production month (Figure 6.12) and dairy of origin (Figure 6.13), and in all cases, a cross validation was performed.

A stepwise DA was performed in order to select the most discriminating peptides in each case. Classification based on ageing time involved mainly the following peptides (variables), which thus are the most different at the different ageing times: ß CN f(83-93), ß CN f(98-124), ß CN f(10-14), ß CN f(47-52), γ Glu-Tyr, αS2 CN f(183-207), Lac-Ile and Lac-Tyr. The cross validated classification allowed to identify the correct ageing period in the 100 % of the cases, confirming that the peptide fractions are highly influenced by the ageing time (and thus can be easily differentiated).

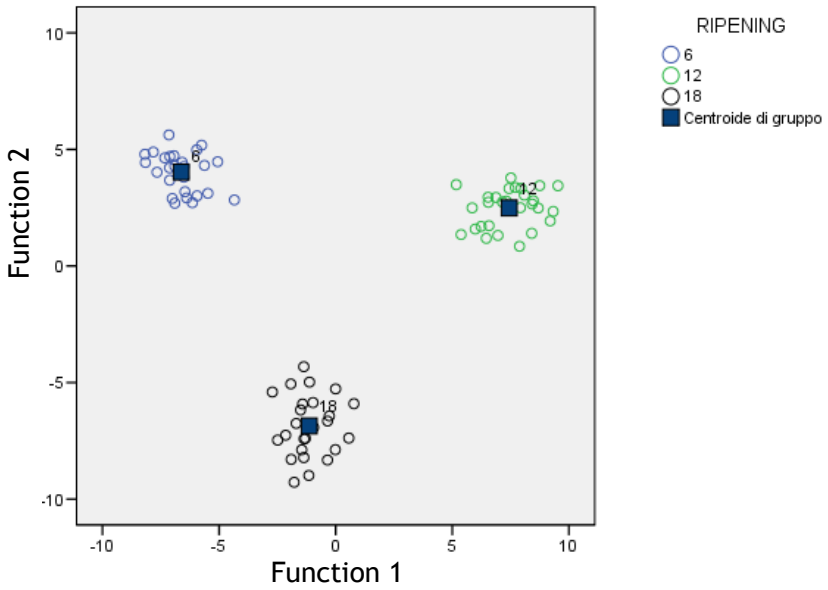


Figure 6.11. DA of the Asiago cheese samples, based on ageing time.

Classification based on the production month (thus, feeding of cows, temperature, etc) involved mainly the following peptides: α S1 CN f(2-14), ß CN f(193-209), αS2 CN f(189-207), α S1 CN f(1-21), α S1 CN f(1-16), unidentified p30 (MW 1025), ß CN f(84-93), ß CN f (193-206), Pyr-Phe, α S1 CN f (24-30), αS2 CN f(182-207) and αS1 CN f (17-23).

The cross-validated classification allowed the correct classification in the 93.8 % of the cases: in particular, group 2 (July) was always perfectly classified, 5.6 % of samples of group 1 (May) were incorrectly classified as belonging to group 2, and 14.3 % of samples of group 3 (September) were incorrectly classified as belonging to group 2. A group 1 sample was never classified as

belonging to group 3, and *vice versa*. Not surprisingly, this indicates that the production of May and September are the most different, and July is in between.

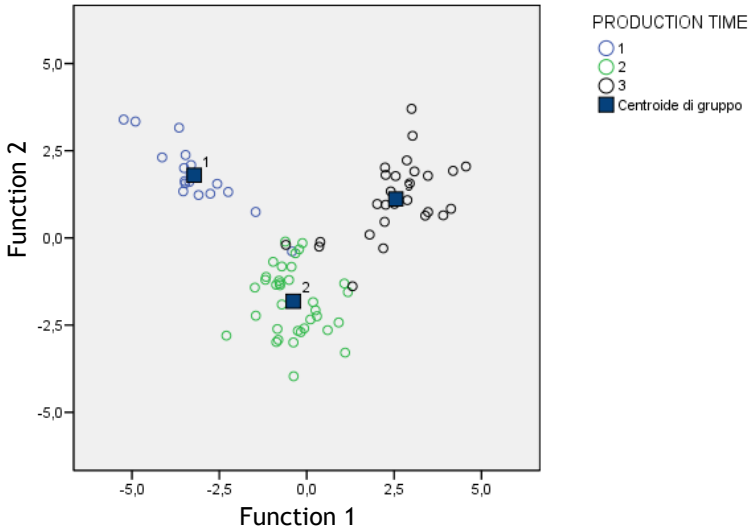


Figure 6.12. DA of the Asiago cheese samples, based on production time.

Classification based on dairies involved a lower number of peptides: unidentified p68 (MW 3594), αS1 CN f(89-110), unidentified p44 (MW 1601) and αS1 CN f(1-23).

Cross validated classification matched the 70 % of the cases, and the specific percentages of correct and wrong attributions are reported in Table 6.4.

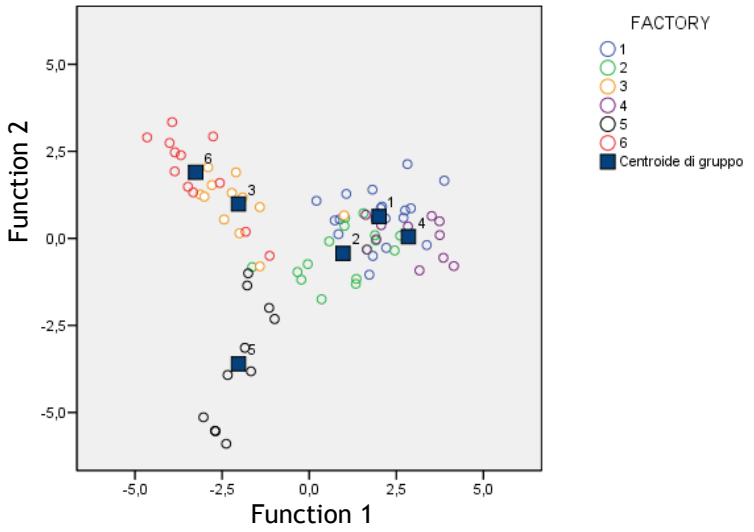


Figure 6.13. DA of the Asiago cheese samples, based on dairy of origin (1=A, 2=B, 3=C, 4=D, 5=E, 6=F).

Table 6.4. Results of the cross validated classification of samples, based on plant of origin.

			Hypothesized group						total
		FACTORY	1	2	3	4	5	6	
Cross validation	Count	1	11	4	0	3	0	0	18
		2	3	10	1	1	1	0	16
		3	1	0	10	0	0	1	12
		4	6	0	0	6	0	0	12
		5	0	1	0	0	10	0	11
		6	0	1	1	0	0	10	12
	%	1	61.1	22.2	.0	16.7	.0	.0	100.0
		2	18.8	62.5	6.3	6.3	6.3	.0	100.0
		3	8.3	.0	83.3	.0	.0	8.3	100.0
		4	50.0	.0	.0	50.0	.0	.0	100.0
		5	.0	9.1	.0	.0	90.9	.0	100.0
		6	.0	8.3	8.3	.0	.0	83.3	100.0

Samples coming from factories A and D are confirmed to be very similar, which is consistent with the analogous technologies of these dairies, as reported in Table 6.2. Again, Coop E turned out to be the most different, with the highest percentage of samples identified.

6.4.4 Non proteolytic peptide-like molecules in Asiago cheese

NPP abundance was distributed within samples as in Figure 6.14. Their amounts in the different samples are reported in Figure 6.15.

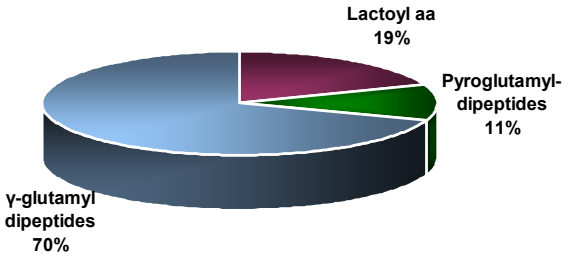


Figure 6.14. Percent distribution of NPP within all Asiago cheese samples.

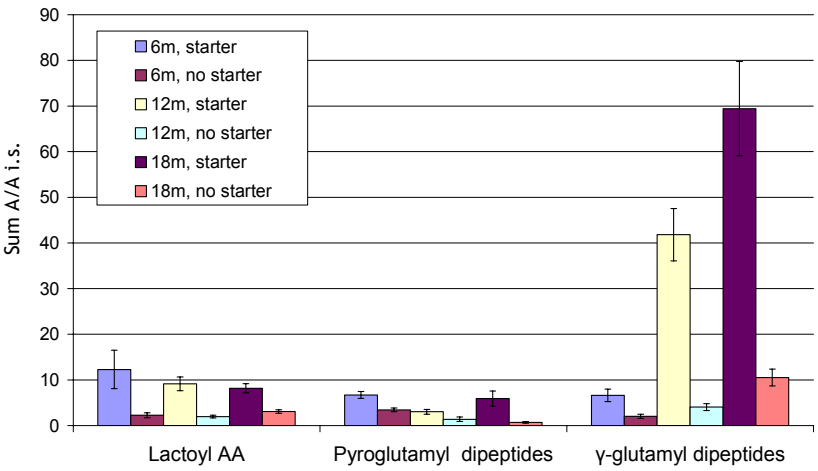


Figure 6.15. Total content of NPPs (differenced by class of compounds on the X axe) in Asiago cheeses, according to time and the use of starter.

Unlike what observed in the samples of Parmigiano-Reggiano cheese, the most differentiating factor appeared to be the use of the starter, rather than the ageing time. From these data it might be hypothesized that NPP formation is likely related to LAB, since, when the starter is used, their quantity is sensibly higher than in cheeses made without the starter.

6.5 Acknowledgements

Prof. Enrico Novelli (University of Padova) is gratefully acknowledged for having provided the samples.

6.6 References

[1] www.asiagocheese.it

[2] IL SISTEMA DELLE MALGHE ALPINE Aspetti agro-zootecnici, paesaggistici e turistici Quaderni Sozooalp N. 1 – **2004** http://www.sozooalp.it/docs/Quad_1/Quaderno%20SZA%20n.1%20Testo.pdf

[3] L'ALIMENTAZIONE DELLA VACCA DA LATTE AL PASCOLO Riflessi zootecnici, agro-ambientali e sulla tipicità delle produzioni, QUADERNI SOZOOALP N. 2 – **2005**
http://www.sozooalp.it/docs/Quad_2/Quaderno%20SZA%20n.2%20Testo.pdf

[4] Quaderno SOZOOALP n° 3 – **2006**, QUALE ZOOTECNIA DA LATTE PER LA MONTAGNA ALPINA? (http://www.sozooalp.it/docs/Quad_3/Quaderno%20SZA%20n.3%20Copertina.pdf)

[5] Paul L H McSweeney, Biochemistry of cheese ripening International *Journal of Dairy Technology* **2004** 57 (2-3), 127–144.

[6] Beresford T.P., Non-starter lactic acid bacteria (NSLAB) and cheese quality, *Dairy processing – Improving quality,* Smit G. (Ed.), Woodhead Publishing Limited, Cambridge, England, **2003**, pp. 448–469.

[7] S. D. PETERSON and R. T. MARSHALL, Nonstarter Lactobacilli in Cheddar Cheese: A Review *J Dairy Sci* **1990** 73:1395-1410,

[8] Michael G. Casey, Jean Pierre Häni, Josef Gruskovnjak, Walter Schaeren, Daniel Wechsler, Characterisation of the non-starter lactic acid bacteria (NSLAB) of Gruyère PDO cheese *Lait* **2006** 86: 407–414, ,

[9] Leszek Stepaniak, Dairy enzymology, *International Journal of Dairy Technology* **2004,** 57, 2/3 153-171

[10] Paolo Piraino, Vivek K. Upadhyay, Rocco Rossano, Paolo Riccio, Eugenio Parente, Alan L. Kelly, Paul L.H. McSweeney, Use of mass spectrometry to characterize proteolysis in cheese, *Food Chemistry* **2007** 101 964–972

[11] Addeo F, Chianese L, Sacchi R, Musso SS, Ferranti P, Malorni A., Characterization of the oligopeptides of Parmigiano-Reggiano cheese soluble in 120 g trichloroacetic acid/l. *J Dairy Res*. **1994** 61,3:365-74.

[12] Ferranti P, Barone F, Chianese L, Addeo F, Scaloni A, Pellegrino L, Resmini P. Phosphopeptides from Grana Padano cheese: nature, origin and changes during ripening. *J Dairy Res.* **1997** 64, 4:601-15.

[13] Stefano Sforza, Linda Ferroni, Gianni Galaverna, Arnaldo Dossena, And Rosangela Marchelli, Extraction, Semi-Quantification, and Fast On-line Identification of Oligopeptides in Grana Padano Cheese by HPLC-MS *J. Agric. Food Chem.* **2003**, 51, 2130-2135,

[14] Martin Lund, Ylva Ardö. Purification and Identification of Water-Soluble Phosphopeptides from Cheese Using Fe(III) Affinity Chromatography and Mass Spectrometry, *J. Agric. Food Chem.* **2004**, 52, 6616-6622

[15] M.J. Sousa, Y. Ard.o, P.L.H. McSweeney Advances in the study of proteolysis duringcheese ripening, *International Dairy Journal,* **2001,** 11: 327–345

[16] Sforza, S.; Galaverna, G.; Neviani, E.; Pinelli, C.; Dossena, A.; Marchelli, R. Study of the oligopeptide fraction in Grana Padano and Parmigiano-Reggiano cheeses by liquid chromatography-electrospray ionization mass spectrometry. *Eur. J. Mass Spec.,* **2004,** 10, 421-427.

[17] C.J. Coker, R.A. Crawford, K.A. Johnston, H. Singh, L.K. Creamer, Towards the classification of cheese variety and maturity on the basis of statistical analysis of proteolysis data—a review, *International Dairy Journal* **2005** 15631–643

7 PROTEOLYTIC OLIGOPEPTIDES AS MOLECULAR MARKERS FOR THE PRESENCE OF COWS' MILK IN FRESH CHEESES DERIVED FROM SHEEP MILK

7.1 Introduction

The production of cheeses derived from sheep milk is very common in European and Asian countries: Greece, France, Spain and Italy are the world leaders. Most of them bear a Protected Designation of Origin (PDO), as Feta (Greece), Roquefort (France), Idiazabal (Spain) and Pecorino Romano (Italy). In many varieties the use of different milks, together with sheep milk, is allowed by national legislations (for example, cows' milk is also used for Caciotta di Urbino or Robiola di Roccaverano PDO cheeses), but food manufacturers, at least in EU, are required to state the species of origin of the milk used for cheese production or other dairy products.[1] The undeclared addition of cow milk to sheep milk is a common fraud[2], which can also have other kinds of implication. Some consumers are allergic or intolerant to cows' milk[3], whereas others prefer to avoid it for religious, ethical or cultural reasons[4]. Therefore, for legal, medical and ethical reasons, cheeses should be authentic and correctly labeled. The undeclared use of cows' milk for the production of cheeses labeled as made from "pure" sheep milk is a relatively common fraud, due to seasonal fluctuations of the availability of sheep milk and its higher price compared to cows' milk.[2]

Although several methods based on protein determination are reported in the literature for detecting sheep milk adulteration with cows' milk,[5,6,7] the application in cheeses is often hampered by the protein breakdown exerted by proteolytic enzymes.[8] In addition to the quite cumbersome European official method, based on different mobility of γ casein obtained after plasmin hydrolysis[9], capillary electrophoresis of whey proteins[2,10], liquid chromatography of denatured caseins[11], immunological methods[12,13,14] were proposed for identifying cows' milk in sheep cheeses. More recently, a PCR-based method aimed at identifying cows' DNA have also been applied[15]. Although immunological, chromatographic and electrophoretic methods are usually robust and reliable, the interference due to the proteolytic processes occurring in cheese led to the conclusion that, in order to have a sound evidence of milk adulteration in cheeses, these methods should always be combined with PCR-based methods[16].

7.2 Aim of the work

In this work we explored the possibility to use short peptides, spontaneously formed during the proteolytic process of the milk caseins, as molecular markers for assessing the milk origin of samples of a typical Italian cheese made from sheep milk ("Vastedda del Belice").

7.3 Experimental part

7.3.1 Solvents and reagents

See section 4.3.1.

7.3.2 Instrumentation

See section 4.3.2.

The MS spectrometer used was a ZMD single quadrupole (MICROMASS).

7.3.3 Samples

Four model "Pecorino" cheeses, aged 13 and 21 days, were produced according to a literature procedure[17] in the Department of Food Sciences of the University of Bologna. Five "Vastedda del Belice", pasta filata cheeses aged from 30 to 60 days made from raw sheep milk without starter addition, were purposely made in local factories with the addition of known percentages of cows' milk[18]. Other eleven "Vastedda" cheeses (no cows' milk declared on the label) were bought on the Italian market. Four commercial non-PDO "Pecorino" cheeses made from mixtures of cows' and sheep milk, as declared on the label, were bought on the Italian market. Cows' casein content of the eleven commercial "Vastedda del Belice" cheeses was determined by IEF electrophoresis according to the European official method (EC Reg. 213/2001).

7.3.4 Extraction and concentration of the oligopeptide fraction

See section 4.3.4.

7.3.5 LC/MS analysis of the oligopeptide fraction

See section 5.3.5.

MS conditions: ESI interface with 80% splitting of the column flow, positive ions, single quadrupole analyzer. Capillary voltage 3kV, cone voltage 30 V, source temperature 100 °C, desolvation temperature 150 °C, cone gas (N_2) 90 L h-1, desolvation gas (N_2) 450 L h-1. Acquisition of total ion chromatograms (TIC, 100-1900 m/z), scan time 4 s, inter-scan delay 0.1 s.

7.3.6 Data analysis

Extract ion chromatograms (XIC) were obtained from the TICs by extracting the following ions:

- sheep α S1 CN f(1-14): m/z 812.8, 542.2, 460.8;
- cows α S1 CN f(1-14): m/z 833.6, 555.9, 431.9;
- sheep α S1 CN f(1-23): m/z 1362.1, 908.4, 681.6, 569.2, 536.2;
- cows α S1 CN f(1-23): m/z 1382.7, 922.1, 691.8, 582.9, 550.0.

Chromatographic peaks of cows' milk peptides were integrated and the areas were normalized to the areas of the homologous sheep milk peptides. Every analysis was performed in duplicate.

7.4 Results and discussion

The LC/ESI-MS method was applied in this study to the analysis of the oligopeptide fractions of different fresh cheeses produced with sheep milk, either pure or mixed with various percentages of cows' milk.

In Figure 7.1 we report the LC/ESI-MS Total Ion Chromatogram (TIC) obtained for the peptide fraction of a "Vastedda" cheese bought on the market, by applying the extraction procedure and the analytical method reported in the experimental section.

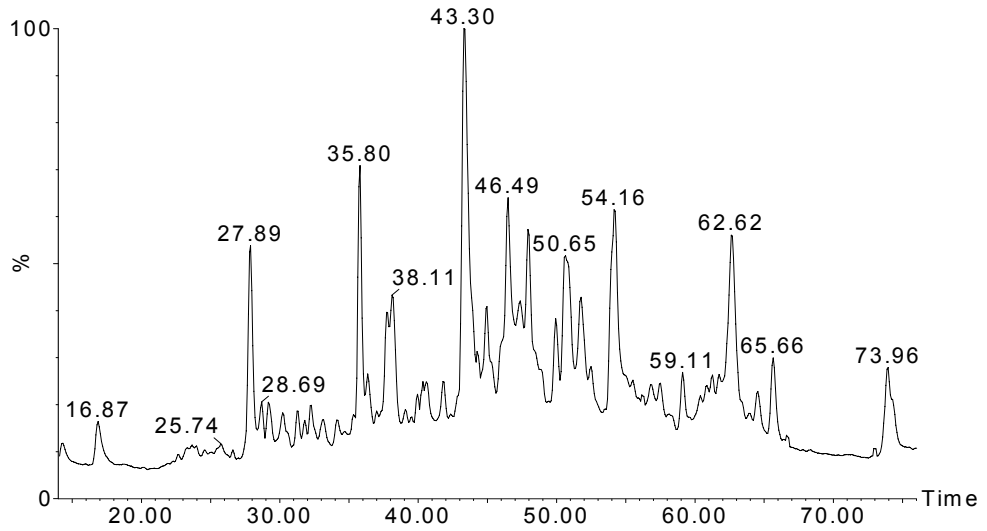


Figure 7.1. Example of the Total Ion Current chromatogram of a commercial Vastedda del Belice cheese.

All cheeses were first screened in order to identify peptides with the characteristics to be considered as good molecular markers for the presence of cows' milk in sheep cheeses: a) simultaneously present homologous peptides derived from the same sequences of vaccine and ovine caseins due to the same proteolytic cleavage; b) one or more different amino acids present between the homologous sequences in order to distinguish them by chromatography and mass spectrometry; c) present in all samples; d) among the most abundant peptides in the chromatograms.

Two peptides were chosen as the most promising considering all the characteristics required: the fragments 1-23 and 1-14 of the α S1 casein. The origin of the first one in cheeses is well known: αS1 CN f(1-23) derives from the cleavage between the positions 23-24 of α S1 casein exerted by chymosin, the main rennet enzyme[19]. It is to be underlined that in some cheeses this peptide has been reported to be undetectable after long ageing times[20], thus its determination could turn out to be unsuitable for long aged cheeses. α S1 CN f(1-14) derives from fragmentation of the α S1 CN f(1-23) between positions 14 and 15. Its formation has been ascribed to the action of lactic acid bacteria (lactococcal or thermophilic lactobacilli) neutral

oligoendopeptidase (PepO), whose activity in cheese is a consequence of the starter lysis[21]. Although the latter peptide can be observed in many different cows' cheese types, usually other peptides derived from α S1 CN f(1-23) are more common, such as α S1 CN f(1-9), α S1 CN f(1-13), α S1 CN f(1-16)[22]. Actually, these peptides were also observed in our samples, but, in all cheeses here considered, α S1 CN f(1-14) was always more abundant than the previous ones. The sequences of the marker peptides, both from ovine and vaccine α S1 caseins, are reported in Table 7.1, together with their molecular weights.

Table 7.1. Cow and sheep peptides used for the detection of cows' milk in sheep cheese. Amino acid substitutions between cows' and sheep caseins are highlighted.

Origin	Peptide	Sequence	Average Mw (Da)	Retention time (min)
Cow	α S1 CN f(1-14)	RPKHPIKHQGLPQE	1665	28.5
Sheep	α S1 CN f(1-14)	RPKHPIKHQGLSPE	1624	27.2
Cow	α S1 CN f(1-23)	RPKHPIKHQGLPQEVLNENLLRF	2764	50.8
Sheep	α S1 CN f(1-23)	RPKHPIKHQGLSPEVLNENLLRF	2723	50.2

Annotated mass spectra of the above reported peptides, extracted from the TIC chromatogram in Figure 7.1, are reported in Figure 7.2.

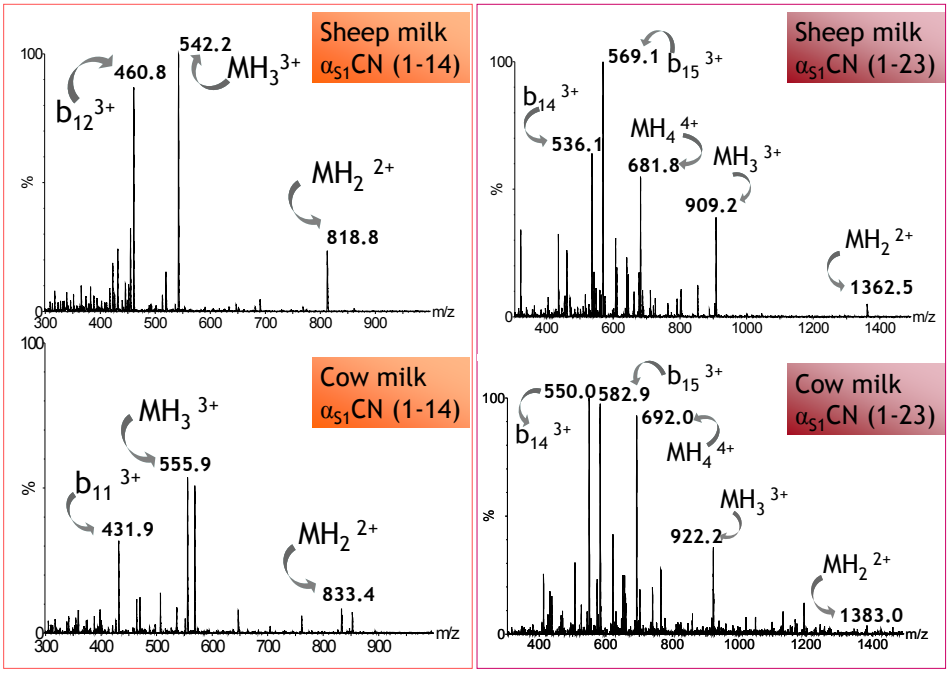


Figure 7.2. Annotated Electrospray-Mass Spectrometry (ESI-MS) spectra of sheep and cows' αS1 CN f(1-23) and αS1 CN f(1-14) peptides.

The α S1 CN f(1-23) peptides (from sheep and from cow) can be identified by the multicharged molecular ions and by two characteristic ions corresponding to b_{15} and b_{14} fragments. The α S1 CN f(1-14) peptides (from sheep and from cow) can be identified by the multicharged molecular ions and by the diagnostic b ions originating by the fragmentation between the proline and the preceding amino acid residue (Ser in the case of sheep casein and Leu in the case of cows' casein).

EXtract Ion Chromatograms (XICs) were obtained by extracting from TICs the diagnostic ions reported in Figure 7.2. As example, XICs obtained by the analysis of the peptide fraction of the commercial "Vastedda del Belice" cheese are given in Figure 7.3. In this sample, no cows' milk was declared, but these data evidenced a small but significant presence of peptides derived from cows' α S1 casein. Since ovine and vaccine peptides have different chromatographic retention times and characteristic mass spectra they could be univocally identified, when present.

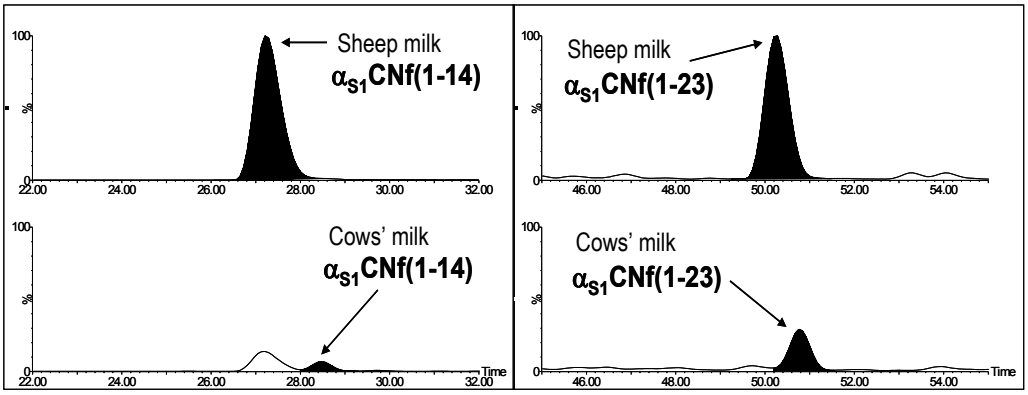


Figure 7.3. Extract Ion Chromatograms (XICs) obtained from the peptide fraction of a commercial "Vastedda del Belice" cheese obtained by extracting from the Total Ion Chromatograms (TICs) the diagnostic ions of peptides derived from cows' and sheep α S1 casein.

For each sample, the areas of the chromatographic peaks corresponding to α S1 CN f(1-14) and to α S1 CN f(1-23) coming from cow casein were integrated and divided for the corresponding areas of the homologous peptides from sheep casein, taken as references. This approach was performed on all samples and all results are reported in Table 7.2.

Table 7.2. Ratio between cow- and sheep-derived peptides in "Vastedda del Belice" samples

Cheese	Bought (market)	Cows' milk in the sheep milk used for caseification	α_{S1}CN f(1-14) cow/sheep ratio	α_{S1}CN f(1-23) cow/sheep ratio
Mixed "Pecorino" [a]	No	5% (added)	0.03	0.08
Mixed "Pecorino" [a]	No	5% (added)	0.03	0.29
Mixed "Pecorino" [b]	No	5% (added)	0.02	- [c]
Mixed "Pecorino" [b]	No	5% (added)	0.02	- [c]
"Vastedda del Belice"	No	1% (added)	<0.01	0.22
"Vastedda del Belice"	No	3% (added)	<0.01	0.29
"Vastedda del Belice"	No	5% (added)	<0.01	0.28
"Vastedda del Belice"	No	10% (added)	0.02	0.33
"Vastedda del Belice"	no	20% (added)	0.04	0.49
"Vastedda del Belice"	Yes	27% (det. by electrophoresis)	0.08	1.15
"Vastedda del Belice"	Yes	n.d. (det. by electrophoresis)	<0.01	<0.01
"Vastedda del Belice"	Yes	11% (det. by electrophoresis)	0.02	0.46
"Vastedda del Belice"	Yes	n.d. (det. by electrophoresis)	<0.01	<0.01
"Vastedda del Belice"	Yes	n.d. (det. by electrophoresis)	<0.01	0.04
"Vastedda del Belice"	Yes	n.d. (det. by electrophoresis)	<0.01	<0.01
"Vastedda del Belice"	Yes	25% (det. by electrophoresis)	0.04	0.23
"Vastedda del Belice"	Yes	n.d. (det. by electrophoresis)	<0.01	<0.01
"Vastedda del Belice"	Yes	31% (det. by electrophoresis)	0.04	1.25
"Vastedda del Belice"	Yes	57% (det. by electrophoresis)	0.19	2.07
"Vastedda del Belice"	Yes	n.d. (det. by electrophoresis)	<0.01	<0.01
Mixed "Pecorino"	Yes	Unknown (presence declared)	2.87	142.2
Mixed "Pecorino"	Yes	Unknown (presence declared)	- [d]	49.4
Mixed "Pecorino"	Yes	Unknown (presence declared)	0.08	0.54
Mixed "Pecorino"	Yes	Unknown (presence declared)	<0.01	0.07

a) 13 days of ageing, **b**) 21 days of ageing
c) Sheep α S1CNf(1-23) not quantitatively measurable (S/N<10)
d) Sheep α S1CNf(1-14) not quantitatively measurable (S/N<10)

In model "Pecorino" cheeses 5% of cows' milk was added before cheese making and also five "Vastedda del Belice" cheeses were made from sheep milk containing known percentages of cows' milk. In commercial "Vastedda del Belice" cheeses cows' milk was not supposed to be present: the eventual presence of cows' caseins was first independently verified by IEF electrophoretic analysis according to the European official method (EC Reg. 213/2001). In commercial mixed "Pecorino" cheeses the presence of cows' milk was declared on the label, although the exact percentage was not known.

The data on model "Pecorino" and "Vastedda del Belice" cheeses, made with sheep milk added of cows' milk, demonstrated the feasibility of the applied method for detecting also very small amounts of cows' milk. In model "Pecorino" samples with 5% of cows' milk, cows' α S1 CN f(1-23) was easily detected after 13 days of ageing, whereas after 21 days of ageing neither sheep or cows' peptide were detectable, probably due to the proteolytic action which lowered their amounts below the detection limit. In any case, the α S1 CN f(1-14) signals were intense in all samples and cows' α S1 CN f(1-14) was always detectable. In "Vastedda del Belice" cheeses made from milk containing various percentages of cows' milk, cows' α S1 CN f(1-23) was clearly detectable down to 1% of cows' milk. Sheep α S1 CN f(1-14) was detectable in all samples, but cows' α S1 CN f(1-14) was below the detection limit for contaminations lower than 10%.

In commercial "Vastedda del Belice" fresh cheeses, previously tested by IEF electrophoresis in order to detect and quantify cows' caseins (cows' milk should not be present in this cheese when sold on the market), both cows' α S1 CN f(1-14) and α S1 CN f(1-23) were consistently detected only in samples in which cows' caseins had previously been found. In all samples devoid of cows' caseins neither cows' peptide was detected, with the exception of one sample which revealed a small, but significant amount of cows' milk α S1 CN f(1-23). In this case it is very likely that a very small amount of cows' milk, undetectable for the electrophoretic method and possibly due to an adventitious contamination, was present in the original sheep milk used for cheese making.

In the commercial "Pecorino" cheeses made from mixed cows' and sheep milk (as declared on the label), cows' α S1 CN f(1-23) was detected in all samples in various amounts. Quite interestingly, cows' α S1 CN f(1-23) peptide was present in two samples in much higher amount than the homologous sheep peptide, suggesting that the nominal "mixed" milk was actually constituted mainly from cows' milk with only small percentages of sheep milk. In these two samples, cows' α S1 CN f(1-14) was also easily detected. As far as the other two samples, in one of them the sheep-derived α S1 CN f(1-14) was not detectable, nor any other peptide derived from α S1 CN f(1-23), suggesting a very low degree of proteolysis due to lactic acid bacteria. In the last one, presumably with less cows' milk, sheep's α S1 CN f(1-14) was detected, but not the analogue from cow.

According to the above results, the simultaneous use of two marker peptides allowed to assess the presence of cows' milk either from both peptides or from only one peptide, thus making the

method more flexible and robust, allowing its application also in cases in which either peptide is below the limit of detection or missing due to lack of endopeptidasic activity (α S1 CN f(1-14)) or enzymatic degradation (α S1 CN f(1-23))[23]. All the data here reported indicate that the cows' α S1 CN f(1-23) is a good marker which allows determining the presence of cow's milk with very good sensitivity, down to very small percentages. The use of cows' α S1 CN f(1-14) as molecular marker is somewhat less sensitive, but still useful as a confirmation or in the cases in which the first one is absent. It is to be underlined that, should in some samples α S1 CN f(1-13) or α S1 CN f(1-16) give a more intense signal than α S1 CN f(1-14), the approach here proposed could be applied to these peptides in the same way.

7.5 Conclusions and perspectives

This is the first time, to the best of our knowledge, that the oligopeptide fraction which spontaneously arises as the result of the proteolytic process occurring in cheese is used in order to evaluate cheese authenticity. The α S1 CN f(1-14) and α S1 CN f(1-23) peptides are good molecular markers for rapidly assessing the presence of cows' milk in fresh cheeses made from sheep milk by LC/MS.

Here, one or the other peptide were present in all the analyzed cheeses, thus allowing detecting the presence of cows' milk, even down to 1%.

There is a need to test more samples in order to fully assess the method applicability, but it is to be underlined that the multiple detection points for cows' derived peptides (two peptides with characteristic retention times and diagnostic mass ions) seems to make negligible the possibility of false positives or false negatives in fresh cheeses. The reported method has all the characteristics to represent a nearly ideal confirmatory test for fresh sheep cheeses, alternative to the official electrophoretic method, eliminating the uncertainty and variability of plasmin activity towards the casein extracted from cheese.

Moreover, this work represents also an approach for the detection of undeclared cow milk added to sheep milk. The considered peptides originate from casein by two different and common proteolytic activity (peptidases of starter lactic acid bacteria and chymosin), present in different products. Notwithstanding, in other cheeses, at different ageing times, other peptides can, in principle, be used for the same purpose. A preliminary study of the peptide composition would underline which peptides can be chosen.

7.6 Acknowledgements

Prof. Maria Elisabetta Guerzoni is gratefully acknowledged for having provided the samples of mixed "Pecorino" made with known percentages of cows' milk.

7.7 References

[1] European Commission. EC 213/2001 Methods for the analysis and quality evaluation of milk and milk products. *Official Journal of the European Communities*, (**2001**). 44, L37/31–L37/99

[2] Herrero-Martinez, J. M., Simo-Alfonso, E. F., Ramis-Ramos, G., Gelfi, C., and Righetti, P. G. Determination of cow's milk in nonbovine and mixed cheeses by capillary electrophoresis of whey proteins in acidic isoelectric buffers. *J. Chrom. A,* (**2000**). 878:261–271

[3] Monaci, L., Tregoat, V., van Hengel, A.J., and Anklam, E. Milk allergens, their characteristics and their detection in food: a review. *Eur. Food Res. Tech.*, (**2006**). 223:149-179.

[4] Shatenstein, B., and Ghadirian, P. Influences on diet, health behaviours and their outcome in select ethnocultural and religious groups. *Nutrition*, (**1998**). 14:223-230

[5] Molina, E., Martin-Alvarez, P.J., and Ramos, M. Analysis of cows', ewes' and goats' milk mixtures by capillary electrophoresis: quantification by multivariate regression analysis. *Int. Dairy J.* . (**1999**). 9:99-105

[6] Veloso, A.C.A., Teixeira, N., and Ferreira, I.M.P.L.V.O. Separation and quantification of the major casein fractions by reverse-phase high-performance liquid chromatography and urea–polyacrylamide gel electrophoresis - Detection of milk adulterations. *J. Chromatography A,* (**2002**). 967:209–218

[7] Hurley, I.P., Coleman, R.C., Ireland, H.E., and Williams, J.H.H. Measurement of Bovine IgG by Indirect Competitive ELISA as a Means of Detecting Milk Adulteration. *J. Dairy Sci.,* (**2004**). 87:543–549

[8] Amigo, L., Ramos, M., Martin Alvarez, P.J., and Barbosa, M. Effect of technological parameters on electrophoretic detection of cows milk in ewes milk cheeses. *J. Dairy Sci.,* (**1991**). 74:1482-1490

[9] Addeo F., Moio L., Chianese L., Stingo C., Resmini P., Berner I., Krause I., Di Luccia A., and Bocca A. Use of plasmin to increase the sensitivity of the detection of bovine milk in ovine and/or caprine cheese by gel isoelectric focusing of γ2-caseins. *Milchwissenschaft,* (**1990**). 45:708-711

[10] Cartoni, G., Coccioli, F., Jasionowska, R., and Masci, M. Determination of cows' milk in goats' milk and cheese by capillary electrophoresis of the whey protein fractions. *J. Chrom. A,* (**1999**). 846:135–141

[11] Bramanti, E., Sortino, C., Onor, M., Beni, F., and Raspi G. Separation and determination of denatured αS1-, αS2-, γ- and β-caseins by hydrophobic interaction chromatography in cows', ewes' and goats' milk, milk mixtures and cheeses. *J. Chrom. A,* (**2003**). 994:59–74

[12] Colak, H., Aydin, A., Nazli, B., and Ergun, O. Detection of presence of cow's milk in sheep's cheeses by immunochromatography. *Food Control,* (**2006**). 17:905–908

[13] Anguita, G., Martin, R., Garcia, T., Morales, P., Haza, A.I., Gonzalez, I., Sanz, B., and Hernandez, P.E. Immunostick ELISA for detection of cow's milk in ewe's milk and cheese using a monoclonal antibody against bovine beta-casein. *J. Food Prot*., (**1996**). 59:436-437

[14] Hurley, I.P., Coleman, R.C., Ireland, H.E., Williams, J.H.H. Use of sandwich IgG ELISA for the detection and quantification of adulteration of milk and soft cheese. *Int. Dairy J.,* (**2006**). 16:805–812.

[15] Lopez-Calleja, I., Gonzalez, I., Fajardo, V., Martın, I., Hernandez, P.E., Garcıa, T., and Martın, R. Application of polymerase chain reaction to detect adulteration of sheep's milk with goats' milk. *J. Dairy Sci*. (**2005**). 88:3115–3120

[16] Hurley, I.P., Coleman, R.C., Ireland, H.E., and Williams, J.H.H. Application of immunological methods for the detection of species adulteration in dairy products. *Int. J. Food Sci. Tech.*. (**2004**). 39:873–878

[17] Lanciotti, R., Vannini, L., Patrignani, F., Lucci, L., Vallicelli, M., Ndagijimana, M., Guerzoni, M.E. Effect of high pressure homogenisation of milk on cheese yield and microbiology, lipolysis and proteolysis during ripening of Caciotta cheese. *J. Dairy Res.* (**2006**). 73:216-226.

[18] Mucchetti, G., Bonvini, B., Remagni, M.C., Ghiglietti, R., Locci, F., Barzaghi, S., Francolino, S., Perrone, A., Rubiloni, A., Campo, P., Gatti, M., and Carminati, D. Influence of cheese-making technology on composition and microbiological characteristics of Vastedda cheese, *Food Control,* (**2008**). 19:119-125

[19] McSweeney, P.L.H., Olson, N.F., Fox, P.F., Healy, A., and Hojrup, P. Proteolytic specificity of chymosin on bovine alpha-s1-casein. *J. Dairy Res.* (**1993**). 60:401-412.

[20] Fernandez, M., Singh, T.K., Fox, P.F. (**1998**). Isolation and identification of peptides from the diafiltration permeate of the water-soluble fraction of cheddar cheese. J. Agr. Food Chem., 46:4512-4517.

[21] Baankreis, R., van Schalkijk, S., Alting, A.C., and Exterkate, F.A. The occurrence of two intracellular oligoendopeptidases in Lactococcus lactis and their significance for peptide conversion in cheese. *Appl. Microbiol. Biotechnol.* (**1995**).44:386–392

[22] Piraino, P., Upadhyay, V.K., Rossano, R., Riccio, P., Parente, E., Kelly, A.L., and McSweeney, P.L.H. Use of mass spectrometry to characterize proteolysis in cheese. *Food Chem.* (**2007**). 101:964–972

[23] Caira, S., Ferranti, P., Gatti, M., Fornasari, M.E., Barone, F., Lilla, S., Mucchetti, G., Picariello, G., Chianese, L., Neviani, E., and Addeo F.. Synthetic peptides as substrate for assaying the proteolytic activity of Lactobacillus helveticus. *J. Dairy Res.,* **2003** 70:315-325

Summary

The present Ph.D. thesis concerns the development of Mass Spectrometry (MS) methods for studying proteins and peptides in food, as an innovative approach for assessing food safety and food quality. Casein traces contaminating hypoallergenic infant formulas were determined by a proteomic bottom-up approach; structures of the allergens of peach (Pru p 3) and plum (Pru d 3) were elucidated by means of a combination of bottom-up and top-down approach. MALDI Mass Spectrometry Imaging was used to detect the two allergen distribution in fruit sections. Peptides derived from simulated gastro-intestinal digestion of peach LTP have also been determined by MS techniques and their allergenic potential has been evaluated by means of dot-blotting technique.
In the second part of this thesis, LC/MS based methodologies have been exploited in order to study oligopeptides in cheeses. Several peptides have been identified, many of them never previously reported, allowing new insights on the influence of ripening time and technologies of production on the proteolysis occurring in cheese. New peptide-like structures constantly increasing during ageing have been identified for the first time. An LC/MS method has also been developed in order to easily detect the undeclared presence of cows' milk in sheep milk cheeses.

Nel presente lavoro di tesi si sono sviluppati metodi di spettrometria di massa (MS) per lo studio di proteine e peptidi in alimenti. Nel contesto della sicurezza alimentare, si sono affrontati diversi temi. Tecniche proteomiche basate su elettroforesi e MALDI-TOF sono state utilizzate per individuare tracce di caseine in alimenti ipoallergenici per l'infanzia. Mediante ESI-MS sono state determinate le strutture primarie di due allergeni, uno di pesca e l'altro di susina, e la loro localizzazione all'interno della sezione del frutto è stata determinata mediante MALDI IMAGING. Inoltre, mediante HPLC/MS, sono stati identificati i peptidi derivati dalla digestione gastro-intestinale simulata sull'allergene della pesca, anche valutati per il loro potenziale allergenico.
Riguardo alla qualità degli alimenti, lo studio ha riguardato la caratterizzazione di peptidi derivati dalla proteolisi nei formaggi. Mediante HPLC/MS, sono stati identificati numerosi peptidi che hanno permesso di ottenere nuove informazioni sull'effetto della stagionatura e delle tecnologie produttive sulla proteolisi nei formaggi. Sono anche stati identificati per la prima volta nuovi derivati amminoacidici che si formano e si accumulano durante la stagionatura. È infine stato sviluppato un metodo HPLC/MS per rivelare l'adulterazione con latte di vacca nella produzione di formaggio Pecorino.

Publications

Manuscript in preparation

V. Cavatorta, S. Sforza, G. Mastrobuoni, G. Pieraccini, G. Moneti, A. Dossena, E.A. Pastorello, R. Marchelli

UNAMBIGUOUS CHARACTERIZATION AND TISSUE LOCALIZATION OF PRU P 3 PEACH ALLERGEN BY ELECTROSPRAY MASS SPECTROMETRY AND MALDI IMAGING

Manuscript accepted for publication

S. Sforza, G. Aquino, V. Cavatorta, G. Galaverna, G. Mucchetti, A. Dossena, R. Marchelli

PROTEOLYTIC OLIGOPEPTIDES AS MOLECULAR MARKERS FOR THE PRESENCE OF COWS' MILK IN FRESH CHEESES DERIVED FROM SHEEP MILK

International Dairy Journal, **2008**

V. Cavatorta, G. Galaverna, S. Sforza, L. Lignitto, E. Novelli

PEPTIDI PROTEOLITICI COME MARCATORI MOLECOLARI DI TIPICITA' NELL'ASIAGO D'ALLEVO

La produzione e trasformazione del latte nel comprensorio montano vicentino. **2007** Ed. Fondazione Studi Universitari di Vicenza e Fondazione CARIVERONA

Pubblications in peer-reviewed journals

V. Cavatorta, S. Sforza, F. Lambertini, M. Bellini, F. Stefani, G. Galaverna, A. Dossena, R. Marchelli

EVOLUZIONE DELLA FRAZIONE AMMINOACIDICA E PEPTIDICA DEL PARMIGIANO-REGGIANO FINO ALL'OTTAVO MESE DI STAGIONATURA

Scienza e Tecnica Lattiero-Casearia, **2007**, 58, 183-203, ISSN 0390-6361.

Publications in Proceedings

Oral Communications

S. Sforza, V. Cavatorta, A. Dossena, R. Marchelli

OLIGOPEPTIDES IN PARMIGIANO-REGGIANO CHEESE: MOLECULAR MARKERS OF TIPICALITY, TECHNOLOGY, AGEING

"*Food quality, an issue of molecule based science*", Proceedings of EuroFoodChem XIV, **2007**, Societe Francaise de Chimie, Cercle des Sciences Analytiques, Volume 2, 408-411.

G. Mastrobuoni, V. Cavatorta, F. Boscaro, S. Sforza, G. Pieraccini, R. Marchelli, A. Dossena, G. Moneti
CHARACTERIZATION AND LOCALIZATION OF ALLERGENIC PROTEINS IN ROSACEAE FRUITS BY HIGH RESOLUTION MASS SPECTROMETRY AND IMAGING MASS SPECTROMETRY
Massa 2007, Italian Annual Meeting on Mass Spectrometry, Lucca, 2-5 settembre **2007**, pp. 51-52, OR-14.

C. Contado, L. Bregola, S. Sforza, V. Cavatorta, R. Marchelli, F. Dondi
A SEDIMENTATION FFF, STUDY TO MONITOR BIO-INTERACTIONS ON SURFACE-MODIFIED PS MICRO-SPHERES.
FFF 2007, 13th International Symposium on Field-and Flow-based Separation, June 27–30, **2007** Salt Lake City, Utah, USA, L2

Poster Communications

V. Cavatorta, S. Sforza, E.A. Pastorello, R. Marchelli
SIMULATED GASTROINTESTINAL DIGESTION OF THE ALLERGENIC PEACH PROTEIN PRU P 3: MOLECULAR CHARACTERIZATION OF THE PRODUCTS BY LC/ESI-MS
4th MS Pharmaday, Parma, Italy, 30-31 ottobre **2006**, Po-06, p 48-49.

V. Cavatorta, S. Sforza, G. Aquino, E.A. Pastorello, R. Marchelli
CARATTERIZZAZIONE MOLECOLARE DELLA STRUTTURA PRIMARIA DEL MAGGIORE ALLERGENE DELLA PESCA, LA PROTEINA PRU P 3, E DEI SUOI PRODOTTI DI DIGESTIONE GASTORINTESTINALE TRAMITE LC-ESI-MS
Atti del VI Congresso Nazionale di Chimica degli Alimenti, **2007**, Editrice Taro, 818-822, ISBN: 978-88-87359-52-7.

S. Sforza, V. Cavatorta, G. Aquino, G. Mucchetti, A. Dossena, R. Marchelli
OLIGOPEPTIDES AS MOLECULAR MARKERS FOR ASSESSING THE PRESENCE OF COW'S MILK IN CHEESE MADE FROM SHEEP'S MILK
"*Food quality, an issue of molecule based science*", Proceedings of EuroFoodChem XIV, **2007**, Societe Francaise de Chimie, Cercle des Sciences Analytiques, Volume 2, 504-507.
Awarded as "Best Poster Presentation" of the Conference

Acknowledgements

I would like to thank my Supervisor Prof. Rosangela Marchelli, for her scientific and moral support, and Prof. Stefano Sforza, who has always been by my side, inspiring my work and offering continual support.

I would like to thank Prof. Arnaldo Dossena, Prof. Gianni Galaverna, Dr. Tullia Tedeschi, Dr. Chiara Dall'Asta, Dr. Andrea Germini and Dr. Andrea Faccini for the academic, scientific and laboratory support.

I would like to thank all the students who have been working with me during their trainings, especially Elisa Sgarbi and Francesca Lambertini.

I would like to thank also all the rest of the academic staff of the Department of Organic and Industrial Chemistry at the University of Parma and the staff of the Centro Interdipartimentale di Spettrometria di Massa (C.I.S.M.) of the University of Florence.

Curriculum Vitae

Valeria Cavatorta
Born in Parma (Italy), August 6th, 1980
Nationality, Italian
valeria.cavatorta@gmail.com

Studies

September 1994 – July 1999

Scientific High school at the "liceo Scientifico G.Ulivi" in Parma, with final grading of 100/100.

September 1999 – October 2004

5 years degree in Food Science and Technology at the University of Parma, with final grading of 110/110 magna cum laude. The dissertation after the final year of laboratory experience was "Attività immunomodulante ed oppiode di estratti idrosolubili di formaggi" ("Immunomodulating and opioid biological activity of water-soluble extracts from cheeses"). This work was carried out both at the Organic and Industrial Chemistry department laboratories of the Univ. of Parma (supervisor Prof. Rosangela Marchelli), both in the "Laboratoires de Recherche en Génie Industriel Alimentaire" of the Univ. of Lyon (supervisor Prof. Adèle Martial). This work comprehended a preparation of samples and peptides and their evaluation on cell culture of human T lymphocytes. This study was possible during an Erasmus' 6 months experience in Bourg en Bresse, France.

January 2005 – December 2007

PhD in Food Science and Technology, University of Parma, Italy, at the Organic and Industrial Chemistry Department (supervisors, Prof. Marchelli and Prof. Sforza), granted by the Italian Ministry of University and Research.

Experiences abroad

From March to July 2004: "Laboratoires de Recherche en Génie Industriel Alimentaire", University of Lyon (France), Erasmus laboratory experience during my degree.