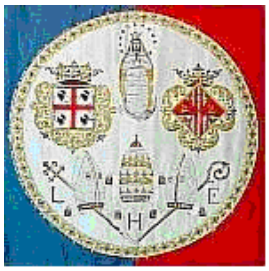

Università degli Studi di Cagliari

DOTTORATO DI RICERCA

Scienze e Tecnologie Chimiche

Ciclo XXIV

METABOLOMIC INVESTIGATION OF FOOD MATRICES BY ^{1}H NMR SPECTROSCOPY

Settore scientifico disciplinare di afferenza
CHIM/02 – Chimica Fisica

Presentata da:	Dott.ssa Cristina Piras
Coordinatore Dottorato	Prof. Mariano Casu
Tutor/Relatore	Dott.ssa Flaminia Cesare Marincola

Esame finale anno accademico 2010 – 2011

Table of Contents

Preface

The term "*metabolomics*" is used to define a scientific study that seeks an analytical description of complex biological samples, by identifying and/or quantifying hundreds or even thousands of distinct chemical identities (low molecular weight metabolites). Initially applied mainly in fields such as medicine, plant sciences, and toxicology, metabolomics has recently emerged as an important tool also in food science, in particular for quality, processing and safety of raw materials and final products. With the advent of "metabolomics" in food science, the analysis of food is now performed in considerably more chemical details in order to understand the molecular details of what gives certain foods their unique flavor, texture, aroma, and color.

High-resolution Nuclear Magnetic Resonance (NMR) spectroscopy has an exceptional place in the chemical analyses of food, offering different advantages with respect to traditional analytical techniques, including simplicity of sample preparation and information on a wide range of compounds present in the food matrix by a single experiment. Among the NMR applications in food science, the ^{1}H NMR-based metabolomic approach has been receiving an increasing interest, since experiments are usually rapid and reproducible and can potentially provide large data sets that turn out to be suitable for statistical interpretation. If coupled with multivariate statistical methods (MVA), with the purpose of unravelling information hidden in complex systems, NMR represents a potent new tool for assessing metabolic function and for highlighting the variations of metabolite concentrations linked to typical conditions. This approach, in particular, opens the possibility of using NMR spectral data for the classification of samples without the use of chemical information, allowing an unbiased chemically comprehensive comparison to be made among different sample.

The work presented in this Ph.D. project shows examples of the successful MVA analysis of liquid-state ^{1}H NMR spectra of foods. This work has resulted in one published paper and two manuscripts in preparation.

The main body of the thesis is divided into the following parts:

Chapter 1 introduces the basic principle of NMR spectroscopy and its applications in food science.

Chapter 2 illustrates the metabolomics field and its terminologies, followed by an introduction of the potentials of using NMR in this research area. Then, some multivariate statistical analysis methods in the data analysis of complex NMR spectra are described together with the importance of the pre-processing of NMR data.

In **Chapters 3** the potential of using NMR spectroscopy for exploring the metabolic profile of bottarga (i.e. salted and dried mullet roes) is investigated. In particular, the usefulness NMR-based metabolomic approach to distinguish bottarga samples according both to the fish geographical origin and the manufacturing processing is illustrated.

In **Chapter 4** a ^{1}H NMR-based metabolomic investigation of Fiore Sardo, a raw ewe's milk cheese produced in the Mediterranean island of Sardinia, is described. In particular, the fermentative performance of autochthonous lactic acid bacteria, used as starter and adjunct cultures in cheese making, was evaluated by a complementary analysis of the metabolic profile of cheese and its microbiological characteristics during ripening. The influence of wild strains on cheese metabolome was compared to that of commercial starters. The results show the potentiality of NMR-based metabolomics for the study of fermentation processes in dairy food matrices.

Chapter 5 illustrates the possibility of using ^{1}H NMR spectroscopy in combination with multivariate statistics in fish research. To this aim, the aqueous and lipid extracts from *Argentina sphyraena* were preliminarily investigated to achieve distinction of fish specimens according to seasonal variations.

Finally, the conclusions of this PhD thesis and the perspective of future work are collected in **Chapter 6.**

1. NMR in food science

1.1 Basic principles of NMR spectroscopy

Some atomic nuclei behave like microscopic bar magnets. The general condition for a nucleus to behave in this way is to have an odd number of protons and/or neutrons (i.e. to have the spin quantum number, I, different from zero). Since the most common nuclei in NMR (^{1}H, ^{13}C, ^{31}P, ^{19}F) have the spin number I equals ½, the following discussion will be particularized for the spin 1/2 nuclei, but the general rules and conclusions are valid for all cases (Abragham, 1961).

The nucleus of an atom, being electrically charged and spinning, generates a magnetic field. The *magnetic moment* (μ) associated with the nucleus is proportional to the *angular momentum (P)*: $\mu=\gamma P$. The proportionality factor γ is named *gyromagnetic ratio* and is a constant for each nucleus. In the absence of other magnetic fields, the magnetic moments μ for a collection of many atoms are randomly oriented and in continuous reorientation due to thermal motions (Figure 1.1A). The result is that the macroscopic sample has no magnetization, because of the averaging to zero of the elementary magnetic moments.

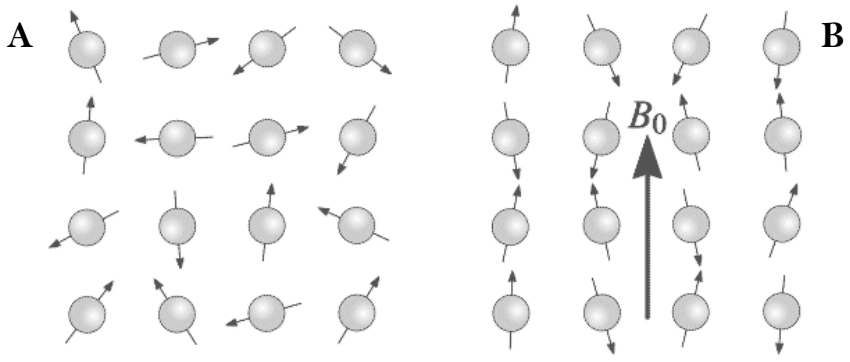

Figure 1.1 The classical model of the formation of net nuclear magnetization in a sample: (A) In the absence of a magnetic field, the individual nuclear magnetic moments (represented by vector arrows here) have random orientation so that there is no net magnetization; (B) In the presence of applied magnetic field, the nuclear magnetic moments are aligned preferentially with the applied field.

If an external magnetic field (B_0) is applied to the sample, the magnetic moments tend to line up in the direction of the field (Figure 1.1B). In this situation, a supplementary movement takes place. The magnetic moment μ precesses around the direction of the external magnetic field with a rotation frequency defined as ω_0, the *Larmor frequency:* $\omega_0 = -\gamma B_0$. This rotation can be clockwise or counterclockwise, depending on the sign of the gyromagnetic ratio.

According to quantum mechanics, for spin of I=1/2, only two orientations of the magnetic moment are allowed: *parallel* and *antiparallel* to the external magnetic field (Figure 1.2). The two allowed orientations have different energies (Figure 1.3) and the number of nuclei adopting each orientation is not equal, the lower energy state (parallel to the Bo field for the positive γ nuclei) being the most populated one. The macroscopic result of this status is that the individual magnetic moments do not cancel to zero anymore. Indeed, taken B_0 along the z axis of the laboratory frame of reference, in the equilibrium situation, the precessing magnetic moments are randomly distributed on the surface of a cone around z (Figure 1.4). The macroscopic result is that there is no net magnetization in the xy plane, the projections of magnetic moments in the xy plane canceling to zero, but only a net magnetization (M_0) on the direction of the external magnetic field.

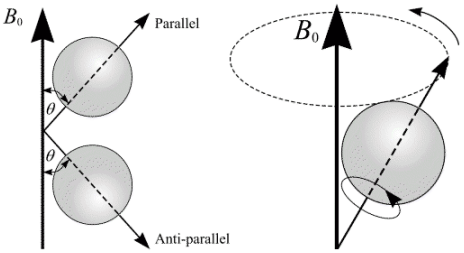

Figure 1.2 (left) In the presence of an externally applied magnetic field, B_0, nuclei are constrained to adopt one of two orientations with respect to B_0. As the nuclei possess spin, these orientations are not exactly at 0 and 180 degrees to B_0, a magnetic moment precessing around B_0. This path describes the surface of a cone (right).

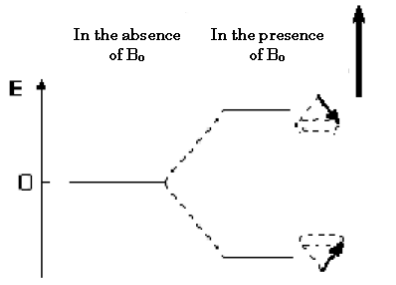

Figure 1.3 Energy levels of nuclear magnetic dipoles of I=1/2 in the absence and in the presence of magnetic field B_0

The energy gap (ΔE) between spatially quantized (allowed) orientations of the magnetic moments depends on the gyromagnetic ratio and on the external magnetic field: $\Delta E=B_0h\gamma/2\pi$, where h is the Planck's constant. If a radiofrequency field is now applied in a direction perpendicular to the external magnetic field and at a frequency that matches the precession frequency of the nucleus, absorption of energy will occur and the nucleus will "flip" from its lower energy orientation to the higher energy orientation. Thus, the net magnetization M_0 can be flipped from the z axis into the xy plane (Figure 1.5). After the radiofrequency field is switched off, the system will return to equilibrium. Consequently, the readable magnetization in the xy plane decay to zero in a certain time which is recorded by the spectrometer in the form of a free induction decay (FID). During an NMR experiment, the signal is measured in the time domain, i.e. as a function of time, and, then, Fourier transformed to obtain the spectrum in the frequency domain. Thereby, the FIDs from active NMR nuclei in a sample, which superimpose in the time domain, are sorted out according to their Larmor frequency by the Fourier transformation.

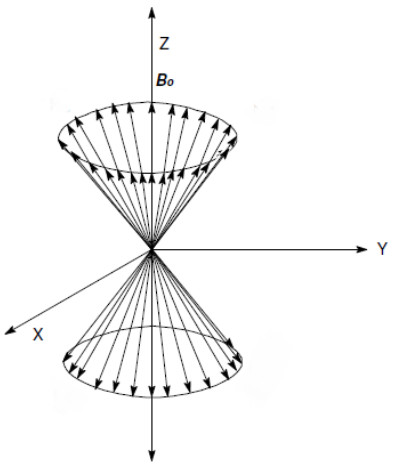

Figure 1.4 Distribution of **N** magnetic moments in equilibrium under the influence of an external magnetic field Bo.

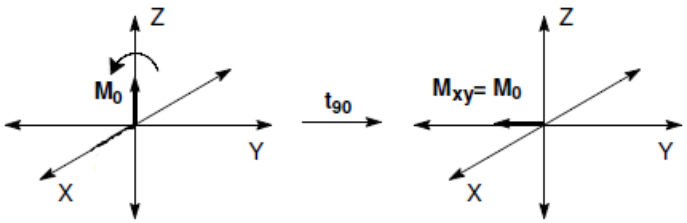

Figure 1.5 The effect of a 90° pulse on equilibrium magnetization.

1.1.1 Types of information provided by an NMR spectrum

The basic information observed in an NMR spectrum is simply a spectral line characterized by position, multiplicity, intensity, and linewidth. A brief description of the meaning of these parameters follows.

The position of an NMR signal reflects the dependence of the nuclear magnetic energy levels on the electronic environment in a molecule. Indeed, the nucleus of a particular element, which is part of a molecule, experiences an effective magnetic field, which is smaller than the external magnetic field B_0. The reason for this phenomenon is the shielding of the nucleus by electrons. The electrons shielding the nucleus could be those belonging to the same atom, those involving the atom in chemical bonds or those of the neighbor atoms. Each atom (of the same element) in a different position in a molecule will have a different electronic surrounding, thus experiencing a different effective magnetic field. The difference between the resonance frequency of a free nucleus and the same nucleus in a chemical environment is named *chemical shift*. The chemical shift is measured by the dimensionless parameter (δ) as the difference between the resonance frequency of the studied nucleus and the resonance frequency of the nucleus of the same element in a reference compound. In order to make this parameter independent of the field strength (and thus independent of the constructive characteristics of the spectrometer), the difference frequency is also divided to the resonance frequency of the standard. As the difference in frequency is very small, the resulting expression is further multiplied with 10^6, thus the chemical shift being expressed as parts per million (ppm). The most widely used reference substance for ^{1}H NMR spectroscopy is TSP-d4 (per-deuterated 3-trimethylsilyl propionate sodium salt), as it exhibits almost complete shielding. The difference between the nucleus and a reference with maximum shielding (per definition 0 ppm) will result in positive value on the ppm axis, as most compounds have a smaller shielding than TSP.

As each individual nucleus active in NMR behaves as a small magnet, it is not surprising that these local magnetic fields interact with neighboring nuclei in a molecule. When interactions of nuclei occur through chemical bonds, the observable effect, known as *spin-spin coupling*, is the splitting of the signal of a particular nucleus in direct relationship with the number and type of neighboring

nuclei. The size of the splitting, called *coupling constant* (*J*) is independent of the magnetic field. The coupling patterns can provide very valuable information for structure assignments.

The *area* under a NMR signal is proportional to the number of nuclei giving rise to that signal. Thus, the ratios of the integrals of various signals in the NMR spectrum represent the relative numbers of atoms for those signals. This property is very valuable in structure elucidation problems. If one refers the integrals of signals in a sample to the integral of a signal belonging to a compound which was added in a known concentration to the sample, then, the concentrations of compounds in complex mixtures can be determined. The accuracy of the integration of NMR signals depends dramatically on the experimental conditions. Each type of nucleus has different characteristics in terms of relaxation properties. In order to obtain fully relaxed signals, one should employ appropriate waiting (relaxation) delays.

The *linewidth* of an NMR signal, usually measured at half height of the peak, can be influenced by factors like: relaxation time, relaxation mechanism, chemical exchange, intramolecular rotations, temperature, presence of paramagnetic impurities, homogeneity of the sample, homogeneity of the spectrometer's magnetic field, interactions with neighboring nuclei. It contains information on the rate of processes, including rate of molecular motions.

1.1.2 Two-dimensional NMR spectroscopy

Besides being able to detect multiple nuclei, NMR spectroscopy has also the advantages of measuring two (or more) nuclei at the same time, by introducing a second frequency dimension. Conventional NMR spectra (one-dimensional spectra) are plots of intensity vs. frequency; in two-dimensional spectroscopy intensity is plotted as a function of two frequencies, usually called F1 and F2. There are various ways of representing such a spectrum on paper, but the one most usually used is to make a contour plot in which the intensity of the peaks is represented by contour lines drawn at suitable intervals, in the same way as a topographical map. The position of each peak is specified by two frequency coordinates corresponding to F1 and F2.

One of the advantages of 2D NMR spectroscopy is the power of structure elucidation of molecules. Types of 2D NMR experiments mostly used in food science include COSY (COrrelation SpectroscopY) and TOCSY (Total Correlation SpectroscopY). COSY shows which signals in a proton spectrum have mutual spin-spin couplings. The resulting spectrum has the conventional 1D NMR spectrum along the diagonal cross-peaks at chemical shifts corresponding to pairs of coupled nuclei. In general, protons within three bonds are well correlated in the COSY spectrum but in sp^2 spin systems such as olefinic and phenolic compounds, even the correlation beyond three bonds is readily detected. TOCSY provides information on unbroken chains of coupled protons in the same molecule. For example, if anomeric protons of carbohydrates are defined, the rest of the signals in the crowded region can be easily assigned using the correlation in the TOCSY spectrum.

1.2 Applications of NMR in food science

In the last twenty years, increasing application of high-resolution (HR) NMR spectroscopy in the study of agriculture food products has been remarked (Alberti et al., 2002; Bertocchi and Paci, 2008; Consonni and Caligiani, 2010). The attention to this technique by scientists, official control institutions and food industries can be attributed both to the high specificity and versatility of the NMR technique and to the improvement of NMR instrument performances and availability. The main characteristics that make NMR very suitable for studying food matrices are the following:

1) it is a non-destructive method, thus it is possible to perform different analyses on the same sample;
2) it is able to detect different nuclei, allowing a study of the sample under different perspectives;
3) it is structure-sensitive, i.e. capable of investigating structural features;
4) it is sensitive to dynamics, which allows differentiation between molecules or portions of molecules with different mobility.

Compared to more common optical spectroscopic techniques, NMR gives a far more detailed molecular picture of food composition, since it simultaneously detects, in a non-targeted way, signals from different compounds, such as carbohydrates, amino acids, organic and fatty acids, amines, and lipids,

without any upfront separation (Belton P.S. et al., 1996). Furthermore, NMR can give information (directly or indirectly) also on the physical status of water and fat, the starch and protein in emulsions, the internal structure of solid foods and so on. More recently, the combination of high resolution ^{1}H or ^{13}C NMR fingerprinting with advanced chemometric methods has emerged as a promising approach for exploring pattern of biomarkes of food quality and authenticity (Sacchi R. et al., 1998; Brescia M.A. et al., 2004; Brescia M.A. et al., 2005; Sacchi R. et al., 2007; Consonni R. et al., 2008; Donarski J.A. et al., 2008). This topic will be discussed in more details in Chapter 2.

However, it is worth reminding that, although NMR spectroscopy has many advantages, this technique has some major drawbacks, mainly concerning the sensitivity (this aspect is based on the small difference of the spin populations in the nuclear energy states, i.e. the population difference) (Abragham, 1961). Although sensitivity reasons rendered the ^{1}H nucleus as the most exploited, other nuclei can be studied by NMR. For instance, one of the main fields in which HR ^{13}C NMR spectroscopy is applied in food science concerns food lipids to asses fatty acid composition, distribution of acyl chain on triacylglicerols, quality and authenticity of food (Sacchi R. e al., 2007).

Besides HR-NMR, other NMR techniques have become highly informative in food applications. Among these:

- *High-resolution solid-state NMR* spectroscopy was developed in the 1970s, but it is not until the recent years that HR ^{1}H magic angle spinning (MAS) NMR spectroscopy has been recognized as an efficient analytical method for analyzing solid food such as beef (Shintu et al., 2007) and flour (Calucci L., 2004). In addition, the HR-MAS method has become an important new tool for the study of semi-solids such as flour doughs or intact fruits (Perez E.M.S., 2011; Vermathen M., 2001).
- The methods of authentication through the determination of stable isotope ratios have in recent years found a remarkably wider range of food applications. The relative deuterium concentration and specific deuterium-site locations in a molecule can be determined using *site-specific natural isotope fractionation NMR (SNIF-NMR*) (Martin G.J. et al., 2006). This technique can provide information about the chemical pathway of formation and, in some cases, information about

the geographical origin, being particularly interest for authentication and certification. SNIF-NMR has been applied to the analysis of wine (Košir I.J. et al., 2001), fishes (Aursand M. et al., 2000), honey (Cotte J.F. et al., 2007), milk and dairy product (Belloque J. et al, 2000).

- *Low-field (LF) NMR* gives information about the relaxation time, strictly correlated with inter and intramolecular motions, diffusion processes, and structural properties of liquids in porous systems or amorphous phase. Very interest and recent applications has been focused on the measurement of bound and free water in food (Bertram H.C. et al., 2002), whereas, in the particular case of aqueous solutions, the relaxation characteristic of the single water signal brings about a wealth of information about the solute molecules that could be relevant in the characterization of a specific sample.
- The past few years have seen also a rapid increase in the applications of *Magnetic Resonance Imaging* (MRI) to foods. MRI is a powerful method for the non-invasive investigation of soft tissues in fruit and vegetables; the high water contents make imaging fairly straightforward and factors such as ripeness and tissue damage appear to affect images in a clearly measurable way. A dedicated review of the application of NMR microscopy in food chemistry has been recently published in the literature (Koizumi M. et al., 2008)

References

Abragam A., *The Principles of Nuclear Magnetism*; Oxford University Press, London, **1961**

Alberti, E; Belton, P.; Gil, A. Applications of **NMR** to food science. *Annu. Rep. Nmr Spectro.* **2002,** 47, 109-148.

Aursand, M.; Mabon, F.; Martin, G.J. Characterization of farmed and wild salmon (*Salmo salar*) by a combined use of compositional and isotopic analyses**.** *J. Am. Oil Chem. Soc* **2000**, *77*, 659-666.

Belloque, J.; Ramos, M. Application of NMR spectroscopy to milk and dairy products. *Trends Food Sci. Tech.* **2000**, *10*, 313-320.

Belton, P.S.; Delgadillo, I.; Holmes, E.; Nicholls, A.; Nicholson, J.K.; Spraul, M. Use of high-field ^{1}H NMR spectroscopy for the analysis of liquid foods. *J. Agric. Food Chem.*, **1996**, *44*, 1483–1487.

Bertocchi F.; Paci M., Applications of High-Resolution Solid-State NMR Spectroscopy in Food Science. *J. Agric. Food Chem.*, **2008**, *56*, 9317–9327.

Bertram, H.C.; Purslow, P.P.; Andersen, H.J. Relationship between Meat Structure, Water Mobility, and Distribution: A Low-Field Nuclear Magnetic Resonance Study. *J. Agric. Food Chem.* **2002**, *50*, 824–829.

Brescia, M.A.; Mazzilli, V.; Sgaramella, A.; Ghelli, S.; Fanizzi, F.P.; Sacco, A. ^{1}H NMR characterization of milk lipids: A comparison between cow and buffalo milk. *J. Am. Oil Chem. Soc* **2004**, *81*, 431-436.

Brescia, M.A.; Monfreda, M.; Buccolieri, A.; Carrino, C. Characterisation of the geographical origin of buffalo milk and mozzarella cheese by means of analytical and spectroscopic determinations. *Food Chem.* **2005**, *89*, 139-147.

Calucci, L. Structure and dynamics of flour by Solid State NMR: effects of hydration and Wheat aging. *Biomacromolecules* **2004**, *5* , 1536–1544.

Consonni, R.; Cagliani, L.R. Geographical characterization of polyfloral and Acacia Honeys by Nuclear Magnetic Resonance and Chemometrics. *J. Agric. Food Chem.* **2008**, *56*, 6873–6880.

Consonni R; Cagliani L R, Nuclear magnetic resonance and chemometrics to assess geographical origin and quality of traditional food products *Advances in Food and Nutrition Research*, **2010**, *59*, 87-165.

Cotte, J.F.; Casabianca, H.; Lhéritier, J.; Perrucchietti, C.; Sanglar, C.; Waton, H.; Grenier-Loustalot, M.F. Study and validity of ^{13}C stable carbon isotopic ratio analysis by mass spectrometry and ^{2}H site-specific natural isotopic fractionation by nuclear magnetic resonance isotopic measurements to characterize and control the authenticity of honey. *An. Chim. Acta* **2007**, *582*, 125–136.

Donarski, J.A.; Jones, S.A.; Charlton, A.J. Application of cryoprobe ^{1}H Nuclear Magnetic Resonance Spectroscopy and Multivariate analysis for the verification of Corsican Honey. *J. Agric. Food Chem.* **2008**, *56*, 5451–5456.

Martin, G.J.; Akoka, S.; Martin, M.L. SNIF-NMR—Part 1: Principles. *Mod. Magn. Reson.* **2006**, *Part 7*, *Part 29*, 1651-1658.

Koizumi, M.; Naito S.; Ishida, N.; Haishi, T.; Kano, H. A dedicated MRI for Food Science and Agriculture. *Food Sci. Technol. Res.* **2008**, *14*, 74-82.

Košir, I.J.; Kocjančič, M.; Ogrinc, N.; Kidrič, J. Use of SNIF-NMR and IRMS in combination with chemometric methods for the determination of chaptalisation and geographical origin of wines (the example of Slovenian wines). *An. Chim. Acta* **2001**, *429*, 195–206.

Sacchi, R.; Mannina, L.; Fiordiponti, P.; Barone, P.; Paolillo, L. Characterization of Italian extra virgin olive oils using ^{1}H-NMR spectroscopy. *J. Agric. Food Chem.* **1998**, *46*, 3947-51.

Sacchi, R.; Paolillo, L. NMR for food quality and traceability in advances in food diagnostics. *Nollet, M.L. and Toldrá Eds, F. Blackwell*, **2007**, *5*, 101-117.

Perez, E. M. S; Lopez, J. G.; Iglesias, M.J.; Ortiz, F.L.; Toresano, F; Camacho, F, Study of the suitability of HRMAS NMR for metabolic profiling of tomatoes: Application to tissue differentiation and fruit ripening , *Food Chem.* **2010**, *122*, 877-887.

Shintu L.; Caldarelli S.; Franke B. M, Pre-selection of potential molecular markers for the geographic origin of dried beef by HR-MAS NMR spectroscopy , Meat Sci., **2007**, *76*, 700-707.

Vermathen, M.; Marzorati, M.; Baumgartner, D.; Good, C.; Vermathen, P.; Investigation of Different Apple Cultivars by High Resolution Magic Angle Spinning NMR. A Feasibility Study, *J. Agric. Food Chem.* **2011**, *59*, 12784-12793.

2. NMR-based Metabolomics

2.1 Metabolomics

The terms *metabolomics* and *metabonomics* are often used interchangeably[1] to indicate a scientific area aimed at identifying and quantifying the *metabolome*, the dynamic set of low molecular weight molecules (metabolites[2]) as found in a an organism or a biological sample. Generally, metabolites include organic species (i.e. amino and fatty acids, nucleic acids, carbohydrates, organic acids, vitamins, polyphenols, lipids), although inorganic and elemental species can also be studied.

Metabolites are often simply view as one of the end-products of gene expression and protein activity. It is increasingly understood that metabolites themselves modulate macromolecular processes through, for example, feedback inhibition and as signalizing molecules. Metabolomic studies are therefore intended to provide an integrated view of the functional status of an organism. The choice of the analytical approach to be used depends on the specific problem. Basically, four different approaches can be employed: *target analysis*, *metabolic profiling*, *metabolomics*, and *metabolic fingerprinting* (Oldiges M. et al., 2007). The *target analysis* is used when one is interested in a specific metabolite; in this case a selective extraction can be performed to concentrate the selected metabolite and to avoid interference from other compounds. The *metabolic profiling* is used when one is interested in the specific role of a selected metabolic pathway; it requires the identification and quantification of a selected number of pre-defined metabolites in a given sample.

---

[1] Historically, the term ***metabonomics*** is defined as "the quantitative measurement of the time related multi-parametric metabolic response of living systems to pathophysiological stimuli or genetic modification" (Nicholson J.K. et al., 1999). It was devised by Jeremy Nicholson, Elaine Holmes, and John Lindon of Imperial College (London) from the Greek roots "meta" (change) and "nomos" (rules or laws) in reference to chemometric models that have the ability to classify changes in metabolism (Lindon J.C. et al., 2004). While not expressly defined, the term ***metabolomics*** was indicated by Fiehn (2001) to be the "comprehensive and quantitative analysis of all metabolites. . . ."

[2] Within the context of metabolomics, a metabolite is usually defined as any molecules less than 1kDA in size. However, there are exception to this, depending on the sample and the detection methods.

Metabolomics requires a complete analysis in which all the metabolites are quantified and identified. Finally, the *metabolic fingerprinting* is used when the sample classification, without quantification of individual specific metabolites, is required

Recent advances in analytical chemistry, combined with multivariate data analysis, have brought the scientific community closer to the final goal of metabolomics, i.e. the comprehensive evaluation of all metabolites, both quantitatively and qualitatively, in living organisms. Among many different technological platforms, NMR and Mass Spectrometry (MS) have been successfully used for metabolic fingerprinting analysis. These two techniques have their respective advantages and limitations, and are often discussed as being complementary (Reily and Lindon, 2005). However, as a tool for metabolomics, NMR has some unique advantages over MS-based methods. Indeed, it can provide a detailed analysis on the bimolecular composition very quickly with relatively simple sample preparation (Reo N.V., 2002).Furthermore, it is a universal detector for all molecules containing NMR-active nuclei, unlike MS where detection of analytes is influenced by selective ionization or ultraviolet spectrometers where only chromophore-bearing compounds are detected. For all proton-bearing molecules, the intensity of all proton signals is proportional to the molar concentration of the metabolite. Thus, using a proper internal standard, the real concentration of metabolites can be easily calculated.

In metabolomics, the detection of metabolites is not the final step, but all data obtained from the analytical methods should be further analyzed by statistical methods in order to extract all possible information. The accuracy and correctness of the data to be further analyzed by the statistical methods are inevitably reliant on the robustness of the raw analytical data set. In this aspect, NMR has a unique advantage, the highest reliability in metabolomics. Unlike the retention time in chromatography-based techniques, with a few exceptions, the chemical shift, coupling constant, and integral of each signal in an NMR spectrum do not change as long as they are measured under the same conditions: applied field strength, solvent, pH, and temperature. Despite the low intrinsic sensitivity, the robustness of data and ability to cover a broad range of metabolites has enabled NMR to be the favoured overall metabolomics and fingerprinting tool. In addition to the

advantages of data robustness, the power of NMR in structure elucidation of metabolites cannot be matched by any other method.

Until recently, most of the work in metabolomics has focused primarily on clinical or pharmaceutical applications such as drug discovery (Kell D.B. et al., 2006), drug assessment (Lindon J.C. et al., 2004), clinical toxicology (Griffin J.L. et al., 2004) and clinical chemistry (Moolenaar S.H. et al., 2003; Wishart D.S. et al., 2001). However, over the past few years, metabolomics has emerged as a field of increasing interest also in food science. With the advent of metabolomics, foods are now being analyzed with considerably more chemical details, with hundreds or even thousands of distinct chemical identities (metabolites) being detected and/or identified in certain food (Moco S. et al., 2006; Ninonuevo M.R. et al., 2006). The potential to chemically "deconstruct" foods and beverages into their chemical constituents offers food chemists a unique opportunity to understand the molecular details of what gives certain foods and drinks their unique taste, texture, aroma, or colour.

2.2 Multivariate statistical analysis in metabolomics

In NMR-base metabolomics studies, many hundreds of samples are routinely analyzed and a minimum of several hundreds of signals are detected in each spectrum. It is, therefore, crucial to extract the relevant information from such a huge data set. Although detailed inspection of NMR spectra and integration of individual peaks can give valuable information on dominant biochemical changes, subtle variation in spectra may be overlooked. The complexity of NMR data makes is of prime importance to utilize data reduction techniques in order to access the latent chemical information in the data. To this aim, multivariate statistical methods (chemometrics) provide an expert means of analyzing and maximizing information recovery from complex NMR spectral data sets (Trygg J., 2007).

In chemometrics there are three basic categories of analysis:

1. *Explorative analysis* that give an overview of all of the data in order to detect trends, patterns, or clusters;
2. *Classification analysis and discriminant analysis* which classifies samples into categories or classes;

3. *Regression analysis and prediction models* used when a quantitative relationship between two block of data is sough.

In the following paragraphs, the chemometric methods most frequently employed in food science to the analysis of NMR data are briefly described

2.2.1 Principal Component Analysis (PCA)

Principal components analysis (PCA) is widely used in metabonomic studies and is an unsupervised approach in that it allows inherent clustering behaviour of samples to be ascertained with no *a priori* knowledge of sample class membership (Jackson J.E.A., 1991). PCA reduces the dimensionality of a data set as it allows multidimensional data vectors to be projected onto a hyperplane of lower dimensions (typically 2 or 3), with this projection explaining as much of the variation as possible within the data.

For instance, in the case of a NMR-based metabolomic investigation, the NMR data consists of a matrix of *N* observations (spectra) and *K* variables (spectral regions) so that a variable space of *K* dimensions is created: each variable represents a numerical value on one coordinate axis, and each observation is placed in *K*-dimensional space (Figure 2.1).

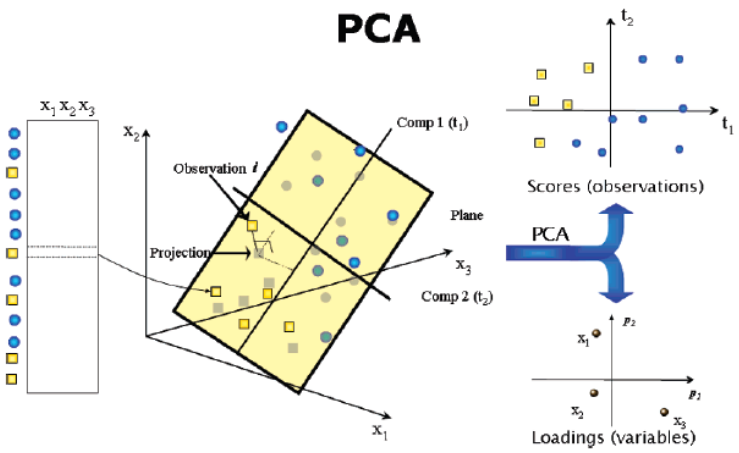

Figure 2.1 A principal component analysis (PCA) model approximates the variation in a data table by a low dimensional model plane.

The PCA model components (PCs) are calculated by a standard method: the first principal component (PC1) is a linear combination of the original input variables, and it describes the largest variation in the data set; the second PC (PC2) is then calculated, and this is orthogonal to PC1 and describes the next highest degree of variation in the data set and so on. When two PCs have been defined, they constitute a plane; hence, projection of the observation vectors in the multidimensional space onto this plane enables the data to be visualized in a two-dimensional (2D) map known as a *scores plot*. This plot reveals any inherent clustering of groups of data, trends or outlier, based purely on the closeness or similarity of their input coordinates. Thus, the analysis provides a convenient and objective means of reducing the complexity of the original data and of visualizing groups and classifying them. A *loadings plot* describes the influence of the variables in the model plane, and the relation among them. An important feature is that directions in the score plot correspond to directions in the loading plot, and vice versa.

2.2.2 Supervised data exploration

In many areas of life sciences today, classification problems constitute the most prevalent forms of intricacies, both in terms of discrimination between groups and interpretation of group differences in meaningful ways. The choice of analytical method adopted is related to whether discriminatory power is of greater importance than the ability to interpret the underlying chemical or biological changes related to class differences.

In what are known as *supervised methods*, data sets can be modelled so that the class of separate samples (a validation set) can be predicted on the basis of a series of mathematical models derived from the original data or training set. One widely used supervised method is the *Partial Least Squares* (PLS) regression (Wold, 1985). PLS is a multivariate method for assessing a relationship between a descriptor matrix X (i.e. spectral intensity values) and a response matrix Y (containing dependent variables). PLS regression has foremost been used in the field of multivariate calibration where the response matrix is quantitative, but might additionally be employed for qualitative data structures typical in discrimination analysis in the form of PLS-DA. However, as PLS-DA explains differences between

overall class properties, the interpretation becomes progressively more complicated as the number of classes increase.

Orthogonal PLS (O-PLS) consists of a new way to decompose the PLS solution into two components: (a) components orthogonal to Y and (b) components correlated to Y. This procedure allows the investigation of the possible relationship between a descriptor matrix X and a response matrix Y, enhancing the relevant information and decreasing or eliminating any structured noise in the data set. As a matter of fact, O-PLS enables the elimination of strong systematic orthogonal variation with respect to Y from a given dataset X, like the so called "structured noise". Since NMR signals could be affected by several sources of noise information, such as temperature, pH and electronic instabilities (Halouska S.M., 2006), the use of O-PLS could improve this source of "noise" information with further advantage of improved detection limit for outliers in the scores, predictions and simplification of data interpretation. OPLS can, analogously to PLS-DA, be used for discrimination (OPLS-DA) (Cloarec O. et al., 2005).

2.3 Pre-processing of NMR data in metabolomics

The initial objective in metabonomics is to classify a NMR spectrum based on identification of its inherent patterns of peaks and, second, to identify those spectral features responsible for the classification, which can be achieved via both supervised and unsupervised pattern recognition techniques. In order to achieve these goals, the NMR data must be prepared for multivariate modelling. The steps involved in analysis of metabonomics NMR data have been well described (Lindon J.C. et al., 2005) and typically involve at a minimum:

a. post-instrument processing of acquired spectroscopic data;
b. production of a data table from the analytical measurements such that there are m rows (observations, samples) and n columns (variables, frequencies, integrals);
c. normalization of the data or some related adjustment to the spectral intensities;
d. scaling of the data;
e. multivariate statistical modelling of the data.

All the necessary processing steps are described in the following.

Phase and baseline correction. Until recently, spectral variations were sought to be manually corrected which is highly time intensive for large data sets. Commercial NMR software now includes multiple methods to automatically correct the baseline inconsistency and phase adjustment. This method appears to work well for phasing and frequency shifting, but only of spectral regions that contains well resolved resonances, not multiple overlapped spectra which is the case in complex NMR spectra of biological matrices. For this reason, complex NMR spectra are often manually phased corrected. As to the baseline offset, it is most often automatically corrected. The post-instrument processing of acquired spectroscopic NMR data mainly includes polynomial baseline correction by removal of offsets.

Binning, intelligent bucketing, and chemical shift alignment. There are a number of approaches for reducing an NMR spectrum to a series of descriptors for metabolomic analysis. Historically, pattern recognition of NMR-based metabolite data was performed using either quantitative or scored integrals of specific spectral peaks. This approach does not work well in crowded regions of spectra with substantial peak overlap and is not easily automated for application to large sample sets. Calculating the peak areas within specified segments of a spectrum (*binning or bucketing*) was introduced originally to allow comparison of NMR data measured at different magnetic field strengths by minimizing, but not eliminating, the effects of second order spectral differences (Gartland K.P.R. et al., 1991; Anthony, M.L. et al., 1994). Accordingly, NMR spectra are typically reduced using bin widths between 0.01-0.04 ppm and the total area within each bin is used as an abstracted representation of the original spectrum (Figure 2.2). This operation reduces the effect of pH-induced changes in chemical shift, ensuring that the same species is always counted correctly across samples with such variation. Besides, binning also encompasses the typical width of an NMR resonance, taking into account spin-spin splittings and line widths. The binning operation results in a data matrix consisting of rows that reflect observations/samples and columns that represent variables, i.e. the spectral integrals of defined bins across the whole spectral width. However,

although the bucketing approach deals with the peak shift problem, it ruins the resolution of the acquired data, thus confounding variance contributions from small peaks with variance contributions from large peaks in the same bucket. It is worth remarking that binned data should only be used for development of chemometric classification models, but it is necessary to examine and analyze the full resolution spectra for biomarker identification.

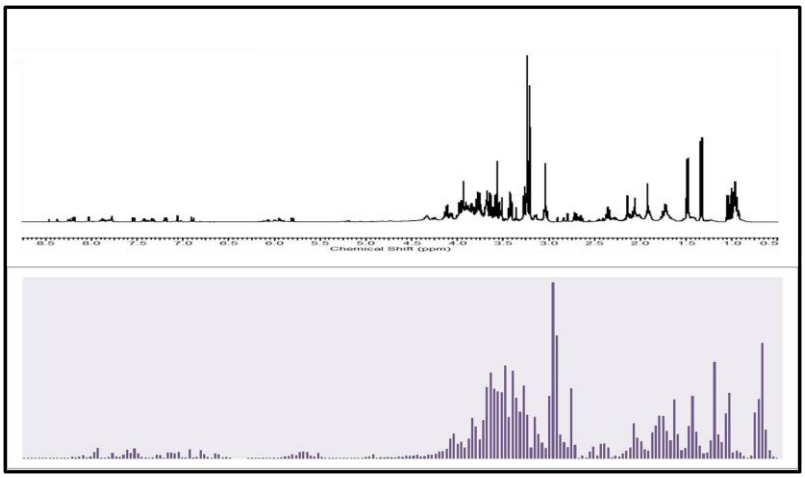

Figure 2.2 An example of segmented NMR spectrum

An alternative approach to standard, equidistant binning is the use of variable bin sizes. This approach should enhance equidistant binning in a very robust way, since a priori knowledge is not introduced and data modifying peak alignments are avoided. Several methods have been proposed, such as nonequidistant binning (Dieterle et al., 2006a) and adaptive binning (Davis R.A. et al., 2007). Both create a reference spectrum, respectively by averaging or taking maximal intensities over all spectra, followed by the determination of the smooth minima of this spectrum, the bin edges. Although the applied smoothening procedures counteract peak shift differences, they depend on rather arbitrarily chosen parameters and the creation of a reference spectrum, implicating loss of information compared with the procedure using all the spectra. Despite their drawbacks, these algorithms outperform standard binning. A similar algorithm, referred to as *intelligent bucketing*, is also

available in the commercial package ACD/Laboratories (www.acdlabs.com). Intelligent Bucketing allows smaller or larger bins within a predefined range and the bin edges are also based on local minima.

Since the use of binned data can lead to inaccuracies in peak intensities (e.g., by inclusion of variable amounts of baseline offset), more attention has recently been paid to methods that forego the need for binning of metabonomic data. Recent advances in chemometric approaches involve the utilization of *full resolution NMR data*, where each data point in an acquired spectrum is extracted as a variable for modelling. This approach has many advantages, for example, the spectral structure is retained, which enables the NMR user to identify metabolites with ease, and it also avoids searching within bins post data modelling to determine metabolites of discriminatory importance. However, in the presence of chemical shift variations due to experimental conditions (for instance, pH), in order to get significant the comparison among spectra, additional pre-processing of NMR spectra must be performed aimed at the *peak alignment* (Stoyanova R. et al., 2004; Forshed J. et al., 2005). Different algorithms for NMR alignment are available in the literature, such as *icoshift* (interval correlated shifting), *COW* (Correlation Optimized Warping), *PARS (*Peak Alignment by Reduced Set mapping*)*, and *SWA (*Segment-Wise Alignment*)*. Among these, in the present PhD work the icoshif program was used (Figure 2.3).

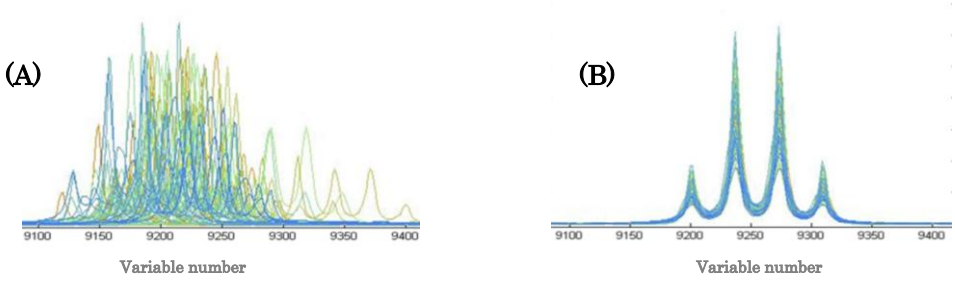

Figure 2.3 Example of peak position variation: nonaligned ^{1}H NMR lactic acid peaks from 45 spectra (A) and profiles processed by Coshift (B).

iCoshift divides spectra into segments and aligns these to the corresponding segments of a reference spectrum. The alignment is performed by shifting the segments sideways so as to maximize their correlation. In practice, this procedure involves calculating the crosscorrelation between the segments by a fast Fourier transform engine that aligns all spectra of a data set simultaneously. The segments can be user-defined or of constant length. Missing parts on the segment edges are either filled with 'missing values', or by repeating the value of the boundary point. The maximum shift correction of the segments can either be equal to a constant defined by the user, or the algorithm can search for the best value for each segment (Savorani F. et al., 2009).

Normalization. This is a row operation that is applied to the data from each sample and comprises methods to make the data from all samples directly comparable with each other. A common use is to remove or minimize the effects of variable dilution of the samples. One common method of normalization involves setting each observation (spectrum) to have unit total intensity by expressing each data point as a fraction of the total spectral integral. We refer to this method as normalization to a constant sum.

Scaling. This operation is performed on the columns of data (i.e., on each spectral intensity across all samples), aimed primary to reduce the noise in the data, and thereby enhance the information content and quality. From a biological point of view, metabolites present in high concentrations are not necessarily more important than those present in low concentration. If concentration determination is the final objective, no scaling of the NMR data should be done, since the relative intensities of the NMR resonances are proportional to the concentration of the observed nuclei. On the other hand, the dominant resonances within the NMR spectra may not necessarily be the spectral features that reveal systematic variation occurring within the analysed samples. Therefore, down weighting of those variables that are the least stable can become a necessary step.

A number of scaling methods are used for NMR data including mean-centring, autoscaling, and pareto scaling (Van den Berg R.A. et al., 2006). The choice for a pretreatment method depends on the biological question to be answered, the

properties of the data set and the data analysis method selected. Mean centering involves the subtraction of the mean value of a descriptor from all values of that descriptor so that the mean for each variable is 0. This is typically done so that all the components found by PCA have as their origin the centroid of the data, resulting in a parsimonious model. Differently, each column of the table can be scaled so that it has unit variance, by dividing each value in the column by the standard deviation of the column. If the data are mean centered, the weighting reflects the covariance of the variables, while in unit variance scaling, the weighting reflects their correlation. Pareto scaling also mean-centers the data, but uses the square root of the standard deviation as the scaling factor.

Basically, mean-centering implies that the variables are centered, but not scaled. In UV scaling, variables are centered and scaled to unit variance, which means that "long" variables are shrunk and "short" variables are stretched, so that all variables will rest on equal footing using as scaling factor the standard deviation. Pareto scaling is between no scaling and UV scaling.

References

Anthony, M.L.; Sweatman, B.C.; Beddell, C.R.; Lindon, J.C.; Nicholson, J.K. Pattern recognition classification of the site of nephrotoxicity based on metabolic data derived from proton nuclear magnetic resonance spectra of urine. *Mol. Pharmacol.* 1994, *46*, 199-211.

Cloarec, O.; Dumas, M.E.; Trygg, J.; Craig, A.; Barton, R.H.; Lindon, J.C.; Nicholson, J.K.; Holmes, E. Evaluation of the orthogonal projection on latent structure model limitations caused by chemical shift variability and improved visualization of biomarker changes in ^{1}H NMR spectroscopic metabonomic studies. *Anal. Chem.* **2005**, *77*, 517-26.

Craig, A.; Cloareo, O.; Holmes, E.; Nicholson, J.K.; Lindon, J.C. Scaling and normalization effects in NMR spectroscopic metabonomic data sets. *Anal. Chem.* **2006**, *78*, 2262-2267.

Davis, R. A.; Charlton, A. J.; Godward, J.; Jones, S. A.; Harrison, M.; Wilson, J. C. Adaptive binning: An improved binning method for metabolomics data using the undecimated wavelet transform *Chemom. Intell. Lab. Syst.* **2007**, *85*, 144–54.

Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. (a) Metabolite projection analysis for fast identification of metabolites in metabonomics. Application in an amiodarone study *Anal. Chem.* **2006**, *78*, 3551–61; (b) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR metabonomics. *Anal. Chem.* **2006**, *78*, 4281-4290.

Forshed, J.; Torgrip, R.J.; Aberg, K.M.; Karlberg, B.; Lindberg, J.; Jacobsson, S.P. A comparison of methods for alignment of NMR peaks in the context of cluster analysis. *J. Pharm. Biomed. Anal.* **2005**, *38*, 824-32.

Gartland, K.P.R.; Beddell, C.R.; Lindon, J.C.; Nicholson, J.K. Application of pattern recognition methods to the analysis and classification of toxicological data derived from proton nuclear magnetic resonance spectroscopy of urine. *Mol. Pharmacol.* **1991**, *39*, 629-642.

Griffin, J.L.; Bollard, M.E. Metabonomics: Its Potential as a Tool in Toxicology for Safety Assessment and Data Integration. *Curr. Drug Metab.* **2004**, *5*, 389-398.

Jackson, J. E. A. *A users guide to Principal Components. Wiley New York* **1991**.

Kell, D.B. Systems biology, metabolic modelling and metabolomics in drug discovery and development. *Drug Discovery Today* **2006**, *11*, 1085-1092.

Lindon, J.C.; Holmes, E.; Bollard, M.E.; Stanley, E.G.; Nicholson, J.K. Metabonomics technologies and their applications in physiological monitoring, drug safety assessment and disease diagnosis. *Biomarkes* **2004**, *9*, 1-31.

Lindon, J.C.; Nicholson, J.K.; Holmes, E.; Keun, H.C.; Craig, A.; Pearce, J.T.; Bruce, S.J.; Hardy, N.; Sansone, S.A.; Antti, H.; Jonsson, P.; Daykin, C.; Navarange, M.; Beger, R.D.; Verheij, E.R.; Amberg, A.; Baunsgaard, D.; Cantor, G.H.; Lehmann-Keeman, L.; Earll, M.; Wold, S.; Johansson, E.; Haselden, J.N.; Kramer, K.; Thomas, C.; Lindberg, J.; Schuppe-Koistinen, I.; Wilson, I.D.; Reily, M.D.; Robertson, D.G.; Senn, H.; Krotzky, A.; Kochhar, S.; Powell, J.; van der Ouderaa, F.; Plumb, R.; Schaefer, H.; Spraul, M. Standard metabolic reporting structures working group. *Nat. Biotechnol.* **2005**, *23*, 833-8.

Moco, S.; Bino, R.J.; Vorst, O.; Verhoeven, H.A.; De Groot, J.; Van Beek, T.A.; Vervoort, J.; Ric de Vos, C.H. A liquid Chromatography-Mass spectrometry-based metabolome database for tomato, *Plant Physiology* **2006**, *141*, 1205-1218.

Moolenaar, S.H.; Engelke, U.F.H.; Wevers, R.A. Proton nuclear magnetic resonance spectroscopy of body fluids in the field of inborn errors of metabolism. *Ann. Clin. Biochem.* **2003**, *40*, 16-24.

Nicholson, J.K.; Lindon, J.C.; Holmes, E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR data. Xenobiotica **1999**, *29*, 1181- 1189.

Ninonuevo, M.R.; Park, Y.; Yin, H.; Zhang, J.; Ward, R.E.; Clowers, B.H.; German, J.B.; Freeman, S.L.; Killeen, K.; Grimm, R.; Lebrilla, C.B. A strategy for annotating the human milk glycome *J. Agric. Food Chem.* **2006**, *54*, 7471- 7480.

Oldiges, M.; Lütz, S.; Pflug, S.; Schroer, K.; Stein, N.; Wiendahl, C. Metabolomics: current state and evolving methodologies and tools. *Appl. Microbiol. Biotechnol.* **2007**, *76*, 495-511.

Reily, M.D.; Lindon, J.C. NMR spectroscopy: principles and instrumentation. In: Robertson, D.G.; Lindon, J.; Nicholson, J.K.; Holmes, E. editors. Metabonomics in toxicity assessment. Boca Raton: CRC Press **2005**, 75–104.

Reo, N.V. NMR-Based metabolomics. *Drug Chem. Toxicol.* **2002**, *25*, 375-382.

Savorani, F.; Tomasi, G.; Engelsen, S.B. icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. *J. Magn. Reson.* **2010,** *202*, 190-202.

Stoyanova, R.; Nicholson, J.K.; Lindon, J.C.; Brown, T.R. Sample classification based on Bayesian spectral decomposition of metabonomic NMR data sets. *Anal. Chem.* **2004**, *76*, 3666-74.

Stoyanova, R.; Nicholls, A.W.; Nicholson, J.K.; Lindon, J.C.; Brown, T.R. Automatic alignment of individual peaks in large high-resolution spectral data sets. *J. Magn. Reson.* **2004**, *170*, 329-35.

Trygg, J.; Holmes, E.; Lundstedt, T. Chemometrics in metabonomics. *J. Proteome Res.* **2007**, *6*, 469–479.

Van den Berg, R.A.; Hoefsloot, H.C.; Westerhuis, J.A.; Smilde, A.K.; vander Werf, M.J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **2006**, *7*, 1-15.

Wishart, D.S.; Querengesser, L.M.M.; Lefebvre, B.A.; Epstein, N.A.; Greinerand, R.; Newton, J.B. Magnetic resonance diagnostics: A new technology for high-throughput clinical diagnostics *Clin. Chem.* **2001**, *47*, 1918-1921.

Wold, H. (1985) Partial Least Squares. In *Encyclopedia of Statistical Sciences* (Kotz S., and Johnson N. L., Eds.) pp 581–591, Wiley, New York.

3. ^{1}H NMR Metabolite Fingerprint and Pattern Recognition of Mullet (Mugil cephalus) Bottarga

3.1 INTRODUCTION

The eviscerated roes of striped mullet *(Mugil cephalus*), a cosmopolitan species found in coastal tropical and subtropical waters, are manufactured in several countries and the salted and dried products can be found world wide under different names and typologies. The Mediterranean island of Sardinia has a long tradition in manufacturing mullet roes to obtain a product called "bottarga". Basically, the curring procedure consists of the following steps (Figure 4.1): a) extraction of ovaries from female fish without breaking. b) washing, to remove impurities; c) salting; d) drying; e) packaging.

Figure 4.1 Typical preparation of bottarga in Sardinia: a) extraction of ovaries from female fish without breaking. b) washing, to remove impurities; c) salting; d) drying, in a well ventilated room; e) packaging.

In recent years, Sardinian bottarga, that is sold as whole ovaries ("in baffe") or grated in jars, has become so increasingly popular in the international markets that mullets of the Mediterranean sea are not enough to satisfy the request of this product. As a result, the Sardinian producers must turn their attention to other fishing areas located in different regions of the globe for roe supplies. Indeed, raw roes are purchased from distributors located mainly in FAO 31 (central-western Atlantic), 34 (central-eastern Atlantic) and 41 (south- western Atlantic) fishing areas. Even if the raw material is not necessarily original of the island, Sardinian bottarga has its own peculiar rheological and organoleptic profile thanks to the skills of the local producers, inheritors of an ancient tradition in processing this delicacy. For this reason, Sardinian manufacturers of bottarga are requesting a Protected Geographical Indication (PGI) designation for this product.

The globalisation of food markets and the relative ease with which food commodities are transported between countries increases the awareness of consumers about the origin of the foods they eat. As far as fishery products are concerned, the European legislation establishes that the FAO area in which fish was caught should be part of the information available to consumers (Commission Regulation (EC) No 2065/2001). This applies also to processed products such as bottarga. It is therefore of great importance to be able to determine the geographical origin of fish, especially when used in preparing processes to perform authentication and/or traceability studies useful to enforce labeling regulations.

Traditional methods for species authentication of fish include DNA and protein analyses (Martinez I. et al., 2003). Recently, the analysis of metabolite profiles by NMR spectroscopy has been proposed as an alternative method for the authentication of seafood (Standal I. et al., 2003; Aursand M. et al., 2009; Gribbestad I.S. et al., 2005; Mannina L. et al., 2008; Martinez I. et al., 2005; Savorani F et al., 2010). In the past two decades ^{1}H NMR has proved to be a fast and versatile technique, useful both for compositional analysis and for rapid screening of food, allowing the detection of the major metabolites in a single spectrum, and, when associated with MVA techniques, can provide a suitable tool for comparing, discriminating, or classifying samples on the basis of their metabolic profile (Martinez I. et al., 2005; Lee E.J. et al., 2009). The free metabolite pool found in the aqueous phase of animal and vegetable matrices reflects the metabolic

processes of the living organism, and it can be characteristic of individuals from a specific geographical area but can also reflect the chemical and physical transformations that can take place during storage of the raw material, manufacturing, and shelf life.

As for the majority of marine products, which are rich in health-beneficial ω-3 fatty acids, most of the investigations on salted and dried mullet roes concern the lipid components and, particularly, the lipid classes and fatty acid composition (see Scano P. et al., 2010 and literature cited therein). To our knowledge, differently from the lipid components, investigations on the low molecular weight compounds of bottarga are rare in the literature (Chiou T.K. et al., 1988).

The aim of this work was to evaluate whether the ^{1}H NMR low molecular weight metabolite profile of bottarga can be considered a valid tool to characterize bottarga samples having different geographical origins and production processing protocols (grated or "in baffe"). For this purpose, the ^{1}H NMR spectra of the aqueous extract of 25 samples of bottarga, manufactured in Sardinia from mullets of known and unknown geographical origin and commercialized either in baffe or grated in jars, were recorded. Principal component analysis (PCA) was applied to the ^{1}H NMR spectral data to explore possible grouping of samples with common characteristics in terms of origin and processing of the raw material.

3.2 MATERIAL AND METHODS

3.2.1 Samples

Twenty-five samples of bottarga manufactured in Sardinia were analysed:

(a) *12 samples in baffe* were kindly gifted by a manufacturer located in Cagliari, on the South of Sardinia. They belong to the same batch of suppliers and were processed from imported frozen roe, kept at -20 °C for not longer than 6 months.

(b) *2 samples from FAO 37.1.3* were kindly gifted by a manufacturer located in Cabras, on the Central Western Coast of Sardinia, and underwent curing procedures soon after evisceration; thus, no freezing procedures were adopted.

(c) *11 samples purchased at a local supermarket* were not labeled with geographical origin. The labels reported the ingredients as mullet roe and salt.

The typology of samples and the geographical provenience of the raw roes are summarized in Table 3.1.

Table 3.1 Summary of the studied bottarga samples (Year of Mullet Catching: 2007)

N° samples	Typology	Identification of the catch area[a]	Catch area
3	in baffe	FAO 34	central-eastern Atlantic
6	in baffe	FAO 41	south- western Atlantic
3	in baffe	FAO 31	central-western Atlantic
2	grated in jar	FAO 37.1.3	Mediterranean Sea
11	grated in jar	unknown	unknown

[a]FAO Yearbook. Fishery Statistics. Catches, 2000; Vol. 86/1

3.2.2 Chemicals

Deuterium oxide (D_2O, 99.9%) was purchased from Cambridge Isotope Laboratories Inc. (Andover, MA, USA). Sodium 3-trimethylsilyl-propionate-2,2,3,3,-d_4 (TSP, 98 atom % D), perchloric acid (70%) and potassium hydroxide (KOH) were acquired from Sigma-Aldrich (Milan, Italy).

3.2.3 Low molecular weight metabolite extraction

Water soluble metabolites were extracted using perchloric acid on the basis of the procedure previously described by Gribbestad et al. (Gribbestad I.S. et al., 2005).

Approximately 2 g of bottarga was pulverized in a mortar and transferred into a glass dish. A 4 mL solution of perchloric acid (7% in D2O) was added and the mixture continuously stirred and warmed at 50°C until a paste consistency was obtained. The homogenate was centrifuged at 4000 rpm for 10 min at 4°C. Then, the supernatant was adjusted to pH 7.8 with 9M KOH in D2O and centrifuged

again to remove the potassium perchlorate. The final extract was lyophilized and stored at -20°C until analyzed. Before NMR analysis, each sample was redissolved in 1mL of D2O, and an aliquot of 600 µL was transferred into a 5 mm tube to which 50 µL of TSP/D2O solution (0.80mM final TSP concentration) was added as internal standard. All of the extractions were performed in duplicate.

3.2.4 NMR data acquisition

^{1}H NMR experiments were carried out on a Varian Unity Inova 400 spectrometer operating at 399.94 MHz. Spectra were recorded at 298 K with a spectral width of 5624 Hz, a 90° pulse of 7.5 µs, an acquisition time of 3 s, a relaxation delay of 25 s, and 64 scans. The residual water signal was suppressed by applying a presaturation technique with low-power radiofrequency irradiation for 1.5 s. The FIDs were multiplied by an exponential weighting function equivalent to a line broadening of 0.3 Hz prior to Fourier transformation. Chemical shifts were referred to the TSP single resonance at 0.00 ppm. 2D NMR ^{1}H-^{1}H COSY spectra were acquired with a spectral width of 4423 Hz in both dimensions, 2048 data points, and 512 increments with 48 transients per increment. 2D NMR ^{1}H-^{1}H TOCSY spectra were acquired in phase sensitive mode with a size and number of data points similar to those of the COSY and a mixing time of 150 ms.

3.2.5 Data processing and Multivariate Data analysis

The ^{1}H NMR spectra were segmented in 191 spectral domains of 0.04 ppm width (bins) by selecting the regions 8.50-5.10 and 4.54-0.80 ppm. No significant resonance shifts were observed for all signals that could justify a different bucketing size. The spectral region between 4.54 and 5.10 ppm was excluded from statistical analysis to remove the effect of the presaturation of the water residual resonance. Bucketing was performed by MestReNova (version 5.2.4).

The integrated area within each bin was normalized to a constant sum of 100 for each spectrum to account for difference in volume of extracts. The final data set consisted of a 25 x 191 matrix, in which rows represented samples and columns the normalized area of each domain. The generated file was imported into SIMCA-P+ program (Version 12.0, Umetrics, Umeå, Sweden) and submitted to mean-centering

and autoscaling before statistical analysis. Principal Components Analysis (PCA) was applied.

3.3 RESULTS AND DISCUSSION

3.3.1 ^{1}H NMR analysis of the aqueous extract of bottarga

The ^{1}H NMR spectral profiles of bottarga aqueous extracts were similar among the examined samples, although changes in the relative intensities of some resonances were observed in all regions. Figure 3.2 shows a representative ^{1}H NMR spectrum and Table 3.2 lists the assignment of the peaks.

41 compounds were identified on the basis of data published in the literature (Seierstad T. et al., 2008; Fan T. et al., 1996; Gribbestad I.S. et al., 2005; Standal I.B. et al., 2007; Mannina L et al., 2008), by performing 2D conventional NMR experiments (COSY and TOCSY), and by recording spectra of standard compounds. In some cases, validation of the peak attribution was achieved by adding standard compounds directly to the sample solution and recording again the NMR spectrum under the same conditions.

The high-field region of the NMR spectrum (0.8-3.0 ppm) showed signals arising from aliphatic groups of free amino acids and organic acids. In particular, signals representing leucine (Leu), isoleucine (Ile), valine (Val), alanine (Ala), lysine (Lys), proline (Pro), and methionine (Met) were identified. The predominant organic acid identified was lactic acid (Lac). In addition, acetic (Ace), malic (Mal), and succinic (Suc) acids were also present. Small singlets at 2.91 and 2.73 ppm were ascribed to trimethylamine (TMA) and dimethylamine (DMA), respectively.

In the midfield region of the spectrum (3.0-5.5 ppm), the main contributions arose from the strongly overlapped signals of the α protons of the free amino acids and from saccharides. Moreover, the intense peaks at 3.21 and 3.24 ppm were attributed to choline (Cho) and phosphorylcholine (P-Cho). Creatine (Crt) and/or phosphocreatine (P-Crt) and creatinine (Crn) were also identified. Furthermore, the

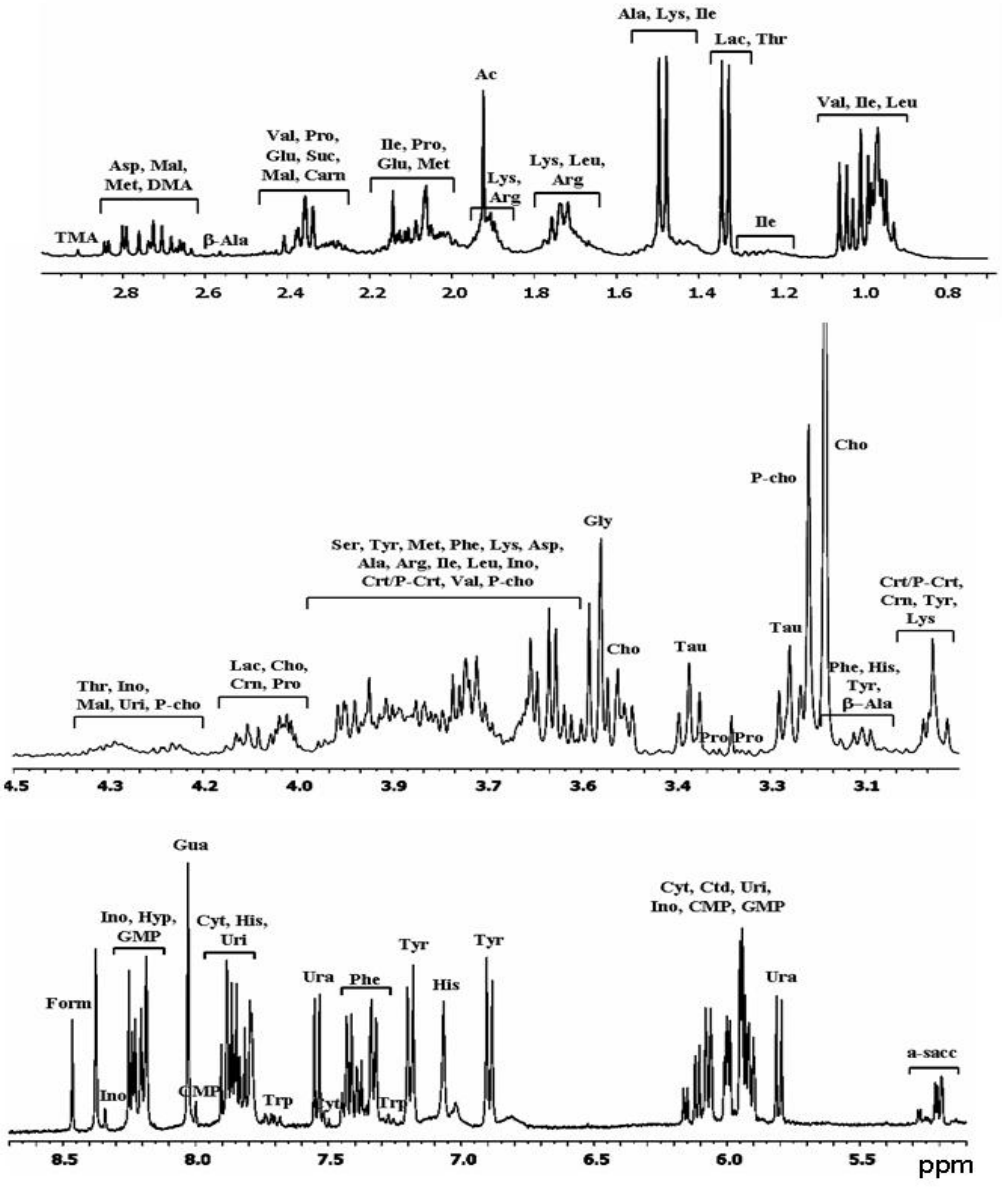

Figure 3.2 Representative ^{1}H NMR spectrum of bottarga aqueous extract in D_2O.

doublets at 5.18-5.22 ppm account for the anomeric proton of α-saccharides (a-sacc).

The signals in the low-field region (5.5-8 ppm) were assigned to the aromatic amino acids tyrosine (Tyr), phenylalanine (Phe), tryptophan (Trp), histidine (His), nucleobases (cytosine (Cyt), uracil (Ura)), nucleosides (cytidine (Ctd), uridine (Uri), inosine (Ino), guanosine (Gua)), and nucleotides. The two singlets at 8.18 and 8.21 ppm were indicative of the presence of hypoxanthine (Hyp). Formic acid (Form) resonates at 8.46 ppm.

Among the identified metabolites, there are nutrients such as taurine (Tau) and carnitine (Carn), fish taste-active amino acids (i.e., Glu, Met, Gly, Ala) (Fuke S. et al., 1991; Hayashi T. et al., 1990), nucleotide derivatives, preservatives (Lac and Mal), and biomarkers typically used for assessing fish quality. In particular, TMA accumulates in spoiling fish as a result of bacterial reduction of TMAO, and DMA, its counterpart, is diagnostic of freezing processes (Martinez I. et al., 2005), whereas the accumulation of Hyp and Ino is directly related to the degradation of adenosine triphosphate (ATP). The presence of these biomarkers in our samples can be ascribed to a series of chemical and physical transformations typically occurring in a biological matrix excised from the living organism, which can modify the original metabolic fingerprint. This is even more important when one is dealing with foodstuffs and, in particular, with fish raw materials transported across continents, stored at low temperature, as in the case of mullet roes.

Table 3.2 Proton chemical shifts (ppm) of the metabolites identified in the aqueous extract of bottarga

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlation[c] (ppm)
Acetate (Ace)	βCH_3	1.92	s	//
Alanine (Ala)	αCH	3.80	q	1.49 (C)
	βCH_3	1.49	d	3.80 (C)
β-alanine (b-ala)	αCH_2	2.56	t	3.19 (C)
	βCH_2	3.19	t	2.56 (C)
Arginine (Arg)	αCH	3.77	t	1.92 (C)
	βCH_2	1.92	m	3.77, 1.70 (C)
	γCH_2	1.70	m	1.92, 3.25 (C)
	δCH_2	3.25	t	1.70 (C)
Aspartate (Asp)	αCH	3.91	dd	2.68, 2.80 (C)
	βCH	2.68	dd	2.80, 3.91 (C)
	$\beta' CH$	2.80	dd	2.68, 3.91 (C)
Carnitine (Carn)	αCH	2.43	m	4.60 (C)
	βCH	4.60	m	2.43 (C), 3.49 (C)
	γCH	3.49	m	4.60 (C)
	$N\text{-}(CH_3)_3^+$	3.24	s	//
Choline (Cho)	$N\text{-}(CH_3)_3^+$	3.21	s	//
	$N\text{-}CH_2$	4.07	m	3.52 (C)
	$O\text{-}CH_2$	3.52		4.07 (C)
Creatine/ Phosphocreatine (Crt/P-Crt)	$N\text{-}CH_3$	3.04	s	3.93 (T)
	$N\text{-}CH_2$	3.93	s	3.04 (T)
Creatinine (Crn)	$N\text{-}CH_3$	3.06	s	//
	$N\text{-}CH_2$	4.06	s	//
Cytidine (Ctd)	C1'H, ribose	5.91	d	//
	C5H	6.07	d	7.85 (C)
	C6H	7.85	d	6.07 (C)
Cytosine (Cyt)	C5H	5.98	d	7.50 (C)
	C6H	7.50	d	5.98 (C)
Cytidine monophosphate (CMP)	C5H, ring	6.12	d	8.08 (C)
	C6H, ring	8.08	d	6.12(C)
Dimethylamine (DMA)	$N(CH_3)_2$	2.73	s	//
Formate (Form)	$HCOO^-$	8.46	s	//

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlation[c] (ppm)
Glutamate (Glu)	αCH	3.78	t	2.10 (C), 2.38(T)
	βCH	2.10	m	2.38 (C)
	βCH	2.10	m	2.38 (C)
	γCH$_2$	2.38	t	2.10 (C), 3.78 (T)
Glycine (Gly)	CH$_2$	3.57	s	//
Guanosine (Gua)	C8H, ring	8.03	s	//
Guanosine monophosphate (GMP)	C1'H, ribose	5.94	d	4.76 (C)
	C2'H, ribose	4.76	m	5.94 (C)
	C8H, ring	8.20	s	//
Histidine (His)	C2H, ring	7.79	s	7.07 (T)
	C4H, ring	7.07	s	7.79 (T)
Hypoxanthine (Hyp)	C2H, ring	8.21	s	//
	C8H, ring	8.18	s	//
Inosine (Ino)	C1'H, ribose	6.11	d	4.82 (C)
	C2'H, ribose	4.82	m	6.11 (C)
	C3'H, ribose	4.48	m	4.36 (C)
	C4'H, ribose	4.36	m	4.48, 3.88 (C)
	C5'H, ribose	3.88	m	4.36 (C)
	C2H, ring	8.34	s	//
	C8H, ring	8.23	s	//
Isoleucine (Ile)	αCH	3.68	m	1.99 (C)
	βCH	1.99	m	1.02, 1.27, 3.68 (C)
	γCH	1.48	m	0.94, 1.27 (C)
	γCH	1.27	m	0.94, 1.48, 1.99 (C)
	γ'CH$_3$	1.02	d	1.99 (C)
	δCH$_3$	0.94	t	1.27, 1.44 (C)
Lactate (Lac)	βCH$_3$	1.34	d	4.12 (C)
	γCH$_2$	4.12	q	1.34 (C)
Leucine (Leu)	αCH	3.75	t	1.70 (C)
	βCH$_2$	1.70	m	3.75 (C)
	γCH	1.72	m	0.96 (C)
	δCH$_3$, δ'CH$_3$	0.96	d	1.72 (C)

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlation[c] (ppm)
Lysine (Lys)	αCH	3.77	t	1.92 (C)
	βCH$_2$	1.92	m	1.43 (C), 3.03 (T), 3.77 (C)
	γCH$_2$	1.43	m	1.73 (C), 1.92 (C), 3.03 (T)
	δCH$_2$	1.73	m	1.43, 3.03 (C)
	εCH	3.03	t	1.43 (T), 1.73 (C),
Malate (Mal)	αCH	4.33	dd	2.38 (C)
	βCH	2.38	dd	2.70, 4.33 (C)
	β'CH	2.70	dd	2.38 (C)
Methionine (Met)	αCH	3.86	t	2.19
	βCH$_2$	2.19	m	2.65, 3.86 (C)
	γCH$_2$	2.65	t	2.19 (C)
	S-CH$_3$	2.14	s	//
Phenylalanine (Phe)	αCH	3.99	dd	3.13, 3.29 (C)
	βCH	3.29	dd	3.99 (C)
	β'CH	3.13	dd	3.99 (C)
	C2,6H, ring	7.42	m	//
	C3,5H, ring	7.42	m	7.33 (C)
	C4H, ring	7.33	m	7.42 (C)
Phosphorylcholine (P-Cho)	N-(CH$_3$)$_3^+$	3.24	s	//
	N-CH$_2$	4.32	m	3.68 (C)
	O-CH$_2$	3.68	m	4.32 (C)
Proline (Pro)	αCH	4.15	t	2.04 (C), 2.35 (C), 2.01 (T)
	βCH	2.35	m	4.15 (C)
	β'CH	2.04	m	4.15 (C)
	γCH$_2$	2.01	m	3.38 (C), 3.40 (C)
	δCH	3.38	t	2.04 (T), 2.35 (T)
	δ'CH	3.40	t	2.04 (T), 2.35 (T)
Serine (Ser)	αCH	3.86	dd	3.96 (C)
	βCH	3.96	dd	3.86 (C)
Succinate (Suc)	α,βCH$_2$	2.41	s	//

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlation[c] (ppm)
Taurine (Tau)	N-CH_2	3.27	t	3.43 (C)
	S-CH_2	3.43	t	3.27 (C)
Trimethylamine (TMA)	N-$(CH_3)_3$	2.91	s	//
Trimethylamine oxide (TMAO)	O-N-$(CH_3)_3$	3.24	s	//
Threonine (Thr)	αCH	3.60	d	4.27 (C)
	βCH	4.27	m	1.34 (C), 3.60 (T)
	γCH_3	1.34	d	4.27 (C)
Tryptophan (Trp)	C4H, ring	7.72	d	7.19 (C), 7.27 (T), 7.54 (T)
	C5H, ring	7.19	t	7.27 (C), 7.72 (C)
	C6H, ring	7.27	t	7.19 (C), 7.54 (C), 7.72 (T)
	C7H, ring	7.53	d	7.27 (C), 7.72 (T)
Tyrosine (Tyr)	αCH	3.94	dd	3.06 (C)
	βCH	3.19	dd	3.06 (C)
	β'CH	3.06	dd	3.19, 3.94 (C)
	C2,6H, ring	6.89	d	7.19 (C)
	C3,5H, ring	7.19	d	6.89 (C)
Uracil (Ura)	C5H, ring	5.81	d	7.54 (C)
	C6H, ring	7.54	d	5.81 (C)
Uridine (Urd)	CH-1' ribose	5.94	d	4.39 (C)
	CH-2' ribose	4.39	m	5.94 (C)
	C5H, ring	5.91	d	7.89 (C)
	C6H, ring	7.89	d	5.91 (C)
Valine (Val)	αCH	3.62	d	2.28 (C)
	βCH	2.28	m	0.99, 1.04, 3.62 (C)
	γ'CH_3	1.04	d	0.99, 2.28 (C)
	γCH_3	0.99	d	1.04, 2.28 (C)

[a] ^{1}H chemical shifts re reported with respect to TSP signal (0.00 ppm).

[b] Multiplicity definitions: s, singlet; d, doublet; t, triplet; q, quartet; dd, doublet of doublets; m, multiplet.

[c] Experiment legend: C, COSY; T, TOCSY

3.3.2 Multivariate Data Analysis

Visual analysis of the NMR spectra did not show obvious relationships between the intensities of certain signals and the geographical origin of mullets. The use for MVA tools was therefore warranted. An exploratory analysis of the data set was carried out by applying a PCA. PCA is an unsupervised technique and requires no information about class membership; it looks just for inherent variation in the data set. Application of PCA on the 25 samples of bottarga under investigation allowed us to reduce the large ^{1}H NMR data set to three principal components with 28.5, 16.1, and 13.4% of total variance explained. As can be seen in Figure 3.3, a good separation of samples of known geographical origin is observed in the score plot of PC1 vs. PC2. Here, bottarga seems to cluster in four different groups among which samples from FAO 37.1.3. (Mediterranean sea, east Sardinia) form a cluster nicely separated from the rest of the data by the PC2. Furthermore, samples of unknown geographical origin, all showing negative PC2 values, can be grouped in two clusters: U1 on the positive side of PC1 and U2 on the negative side of PC1.

The explanation of what each PC represents in relation to the original measurements can be assessed by analyzing the coefficients by which the original variables (in our case, the spectral bins) must be multiplied to obtain the PC, that is, the "loadings". Examination of the loadings plot enables us to determine the variables with the highest impact on the variance and, thus, to identify the metabolites that contributed most to the cluster separation. However, for our data set, the latter step is not straightforward. Indeed, because the ^{1}H NMR spectra of extracts of bottarga are very crowded with signals, several variables hold contributions from more than one metabolite, and, in some cases, the same metabolite contributes to more than one variable; thus, the resulting loadings plots are difficult to analyze. For these reasons, we simplified the appearance of the PC1 versus PC2 loadings plot, underlining only those bins composed predominantly by one metabolite and lying in the extremity of the axis, that is, giving the highest contribution to the PCs (Figure 3.3b).

Here, contributions of metabolites to the sample clustering can be estimated considering the following rule: the position of a sample in a given direction in the score plot is influenced by the metabolites lying in the same direction in the

corresponding loadings plot. Grouping of samples of different origin is mainly along PC1, and metabolites that characterize this first PC are Phe, Tyr, and, on the opposite side, nucleotides and derivatives (CMP, GMP, Ctd, Uri, Ino, Gua, Hyp, Ura, and Cyt). It is interesting to note that these latter, probably from disruption of nuclei acids, are strongly correlated.

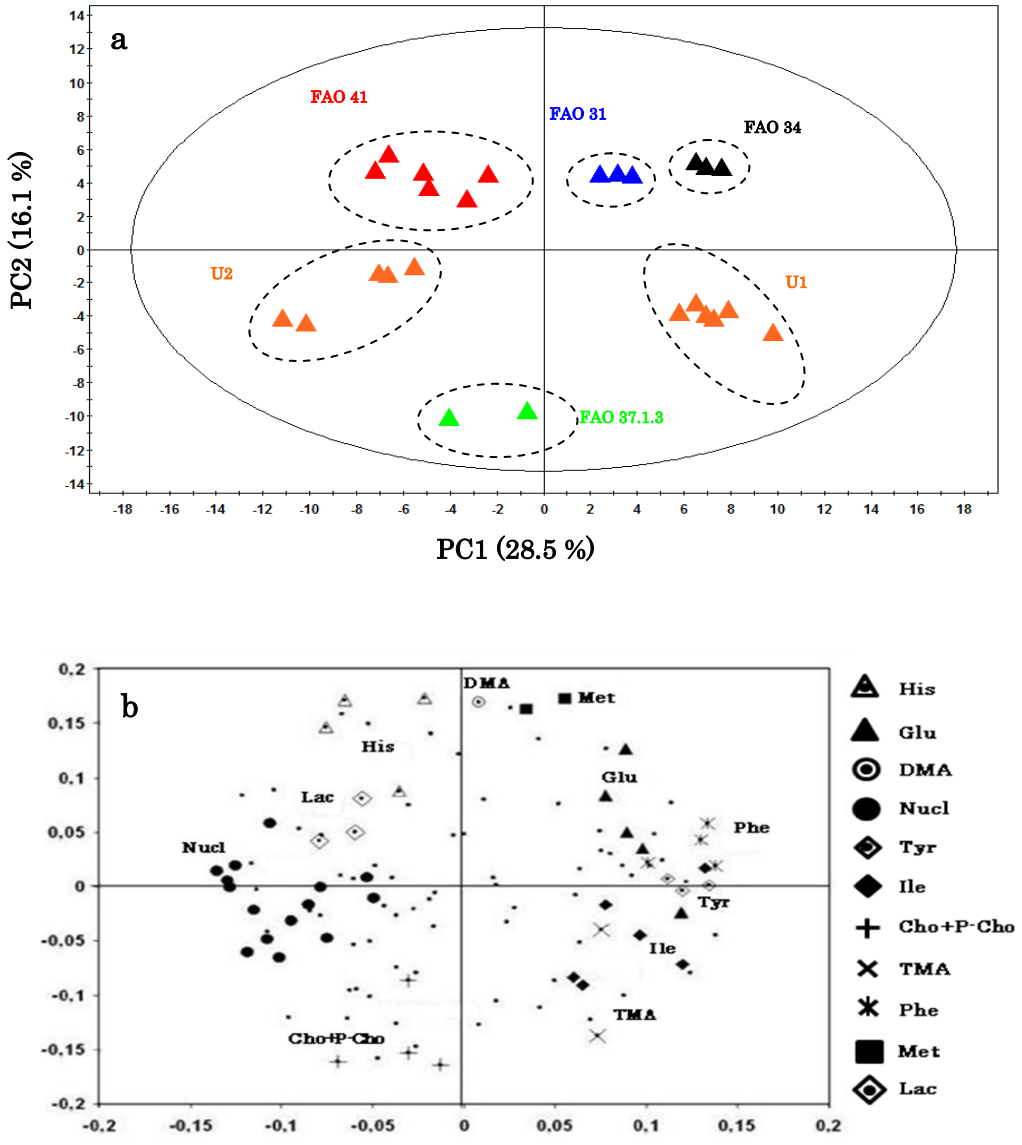

Figure 3.3 PC1 versus PC2 (a) scores plot (ellipses are arbitrarily drawn to group samples; the explained variance is given is parentheses) and (b) loadings plot (with the most significant metabolites highlighted) of PCA applied to the ^{1}H NMR spectral data of the bottarga aqueous extracts. "Nucl" includes CMP, GMP, Ctd, Uri, Ino, Gua, Hyp, Ura, andCyt.

Along PC2 we found that DMA, a biomarker of freezing, is placed on the opposite side of samples from FAO 37.1.3; this observation is in agreement with the fact that mullet roes from FAO 37.1.3 underwent curing procedures soon after evisceration. Furthermore, Cho and P-Cho, derivatives of phosphocholine (PC) in which mullet roes are rich (Scano P. et al., 2008), are placed along PC2, on the same side of the grated samples. These compounds can be considered as biomarkers of hydrolytic mechanisms on PC, caused by the salting and drying procedures and probably exacerbated by the industrial grating. Samples from FAO 34 are more characterized by Met and Glu, which are taste-active compounds in fish derivatives (Fuke S. et al., 1991; Hayashi T. et al., 1990), whereas samples from FAO 41 are characterized by His and Lac. With regard to the samples of unknown geographical origin, the U1 cluster is rich in Ile and TMA; the latter can be linked with a more marked degradation of the raw food matrix (the uncured roes), because the salinity of the final product (bottarga) does not allow extensive bacterial growth. A comparative analysis of the full-resolution spectra confirmed the above-reported observations regarding metabolites characterizing sample grouping.

Separation of bottarga according to geographical origin of mullet is still visible in the scores plot of PC1 versus PC3 in Figure 3.4a. Interestingly, here samples of cluster U2 overlap with FAO 41 and those of U1 with FAO 31. The same procedure was applied to the PC1 versus PC3 loadings plot shown in Figure 3.4b. Opposite to PC2, in the third PC, loading values of DMA, Cho and P-Cho, and His are very low, whereas Val is the metabolite that most contributes to the separation of the samples along PC3.

The results of the PCA models suggested that the application of MVA to the ^{1}H NMR data allows characterization of bottarga according to the geographical origin of the raw material and clustering of samples on the basis of their history and treatments. These encouraging results led us to investigate the possibility of classifying samples of unknown origin as belonging to a specific geographical region. To this goal, the best analytical tool would be a discriminant analysis, such as partial least-squares discriminant analysis; however, our restricted and inhomogeneous sampling, in terms of numbers of samples in each group, did not

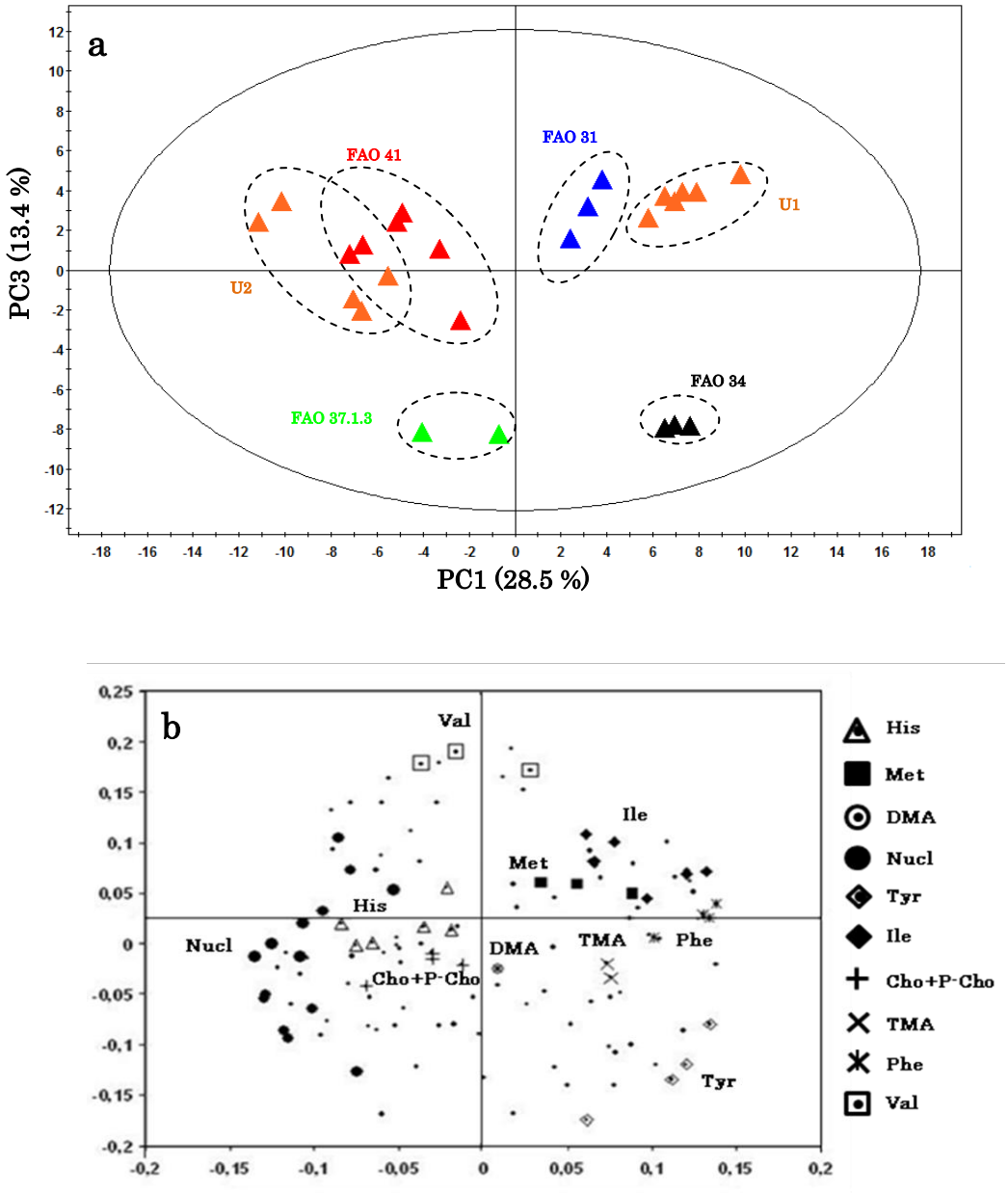

Figure 3.4 PC1 versus PC3 (a) scores plot (ellipses are arbitrarily drawn to group samples; the explained variance is given in parentheses) and (b) loadings plot (with the most significant metabolites highlighted) of PCA applied to the ^{1}H NMR spectral data of the bottarga aqueous extracts. "Nucl" includes CMP, GMP, Ctd, Uri, Ino, Gua, Hyp, Ura, and Cyt.

allow its performance. Therefore, we carried out a PCA removing samples from FAO 37.1.3 and FAO 34, because, by analysis of Figures 3.3 and 3.4, they do not seem to have any common characteristic with the unknown samples. The resulting scores plot, reported in Figure 3.5a, shows that samples called U2 of unknown geographical origin are grouped with samples from FAO 41, whereas all other

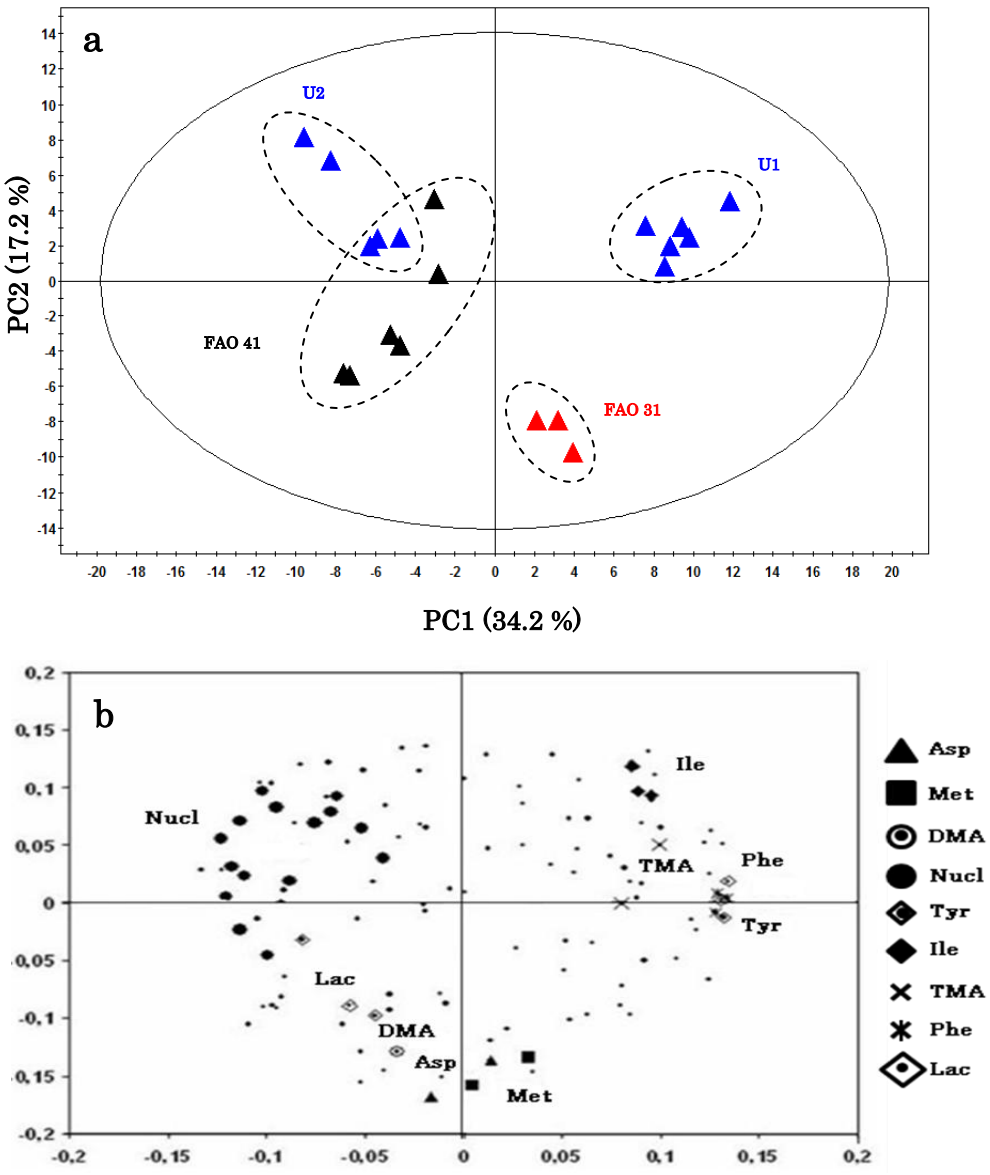

Figure 3.5 PC1 versus PC2 (a) scores plot (ellipses are arbitrarily drawn to group samples; the explained variance is given in parentheses) of PCA applied to the ^{1}H NMR spectral data of the bottarga aqueous extracts (samples from FAO 37.1.3 and FAO 34 have been removed from the original matrix) and (b) loadings plot (with the most significant metabolites highlighted). "Nucl" includes CMP, GMP, Ctd, Uri, Ino, Gua, Hyp, Ura, and Cyt; "others" refers to all the other loadings.

samples U1 show no similarities either with FAO 41 or with FAO 31. The explanation of this result might be that the unclassified samples do not belong to any of the fishing areas here studied and/or they are blends of roes of mullets caught in different geographic areas. The corresponding loadings plot, reported in

Figure 3.6b, shows a trend similar to that of Figure 3.4b; that is, nucleosides and derivatives are strongly inversely correlated to Tyr and Phe. Moreover, they mostly contributed to the overlap of U2 group with FAO 41 samples. TMA and Ile characterize U1 samples, whereas DMA, Met, and Asp characterize samples from FAO 31.

In conclusion, the results of the present study, although a larger data set is warranted, demonstrated that the application of MVA to the ^{1}H NMR spectral data allows bottarga to be characterized according to the geographical origin of the raw material and storage and manufacturing procedures. In fact, in the space spanned by the first three PCs, samples tend to cluster on the basis of their geographical origin and sample history. Among the molecular compounds unambiguously identified, Phe, Tyr, and nucleoside derivatives, followed by DMA, Cho and P-cho (FAO 37), His and Lac (FAO 41), Asp, Met, and Glu (FAO 31 and 34), and Ile and TMA (U1), are the metabolites that principally characterize the groups of samples. Among samples of unknown origin, the tendency of U2 to cluster together with samples of FAO 41 can be ascribed, in terms of metabolites, to the nucleoside derivatives and their inverse correlation with Phe and Tyr.

References

Aursand, M.; Standal, I.B.; Praël, A.; McEvoy, L.; Irvine, J.; Axelson, D.E. ^{13}C NMR Pattern Recognition Techniques for the Classification of Atlantic Salmon (*Salmo salar* L.) According to Their Wild, Farmed, and Geographical Origin. *J. Agric. Food Chem.* **2009**, *57*, 3444–3451.

Chiou, T.K.; Konosu, S. Changes in extractive components during processing of dried Mullet roe. *Nippon Sui san Gakkaishi* **1988**, *54*, 307-313.

Commission Regulation (EC) No 2065/2001, Council Regulation (EC) No 104/2000. *Official Journal L 278* , **2001**, 0006 – 0008.

Eriksson, L.; Johansson, E.; Kettaneh-Wold, N.; Trygg, J.; Wikström, C.; Wold, S. In *Multi- and Megavariate Data Analysis Part I: Basic Principles and Applications.* Umetrics, Umeå, Sweden **2006**.

Fan; T. Metabolite profiling by one- and two-dimensional analysis of complex mixtures. *Prog. NMR Spectrosc.* **1996**, *28*, 161-219.

Fuke, S.; Konosu, S. Taste-active components in some foods: A review of Japanese Research. *Physiol. Behav.* **1991**, *49*, 863-868.

Gribbestad, I.S.; Aursand, M.; Martinez, I. High-resolution ^{1}H magnetic resonance spectroscopy of whole fish, fillets and extracts of farmed Atlantic salmon (*Salmo salar*) for quality assessment and compositional analyses. *Aquaculture* **2005**, *250*, 445– 457.

Hayashi, T.; Kohata, H.; Watanabe, E.; Toyama, K. Sensory study of flavour compound in extracts of salted salmon eggs (Ikura). *J. Sci. Food Agric.* **1990**, *50*, 343-356.

Lee, E. J.; Shaykhutdinov, R.; Weljie, A.M.; Vogel, H.J.; Facchini, P.J.; Park, S.U.; Kim, Y.K.; Yang, T.J. Quality Assessment of Ginseng by 1H NMR Metabolite Fingerprinting and Profiling Analysis. *J. Agric. Food Chem.* **2009**, *57*, 7513-7522.

Mannina, L.; Sobolev, A.P.; Capitani, D.; Iaffaldano, N.; Rosato, M.P.; Ragni, P.; Reale, A.; Sorrentino, E.; D'Amico, I.; Coppola, R. NMR metabolic profiling of organic and aqueous sea bass extracts: Implications in the discrimination of wild and cultured sea bass. *Talanta* **2008**, *77*, 433–444.

Martinez, I.; Aursand, M.; Erikson, U.; Singstad, T.E; Veliyulin, E.; Van der Zwaag, C. Destructive and non-destructive analytical techniques for authentication and composition analyses of foodstuffs. *Trends Food Sci.* Tech. **2003**, *14*, 489-498.

Martinez, I.; Bathen, T.; Standal, I.B.; Halvorsen, J.; Aursand, M.; Gribbestad; I.S.; Axelson, D.E. Bioactive compounds in Cod (*Gadus morhua*) products and suitability of ^{1}H NMR metabolite profiling for classification of the products using multivariate data analyses. *J. Agric. Food Chem.* **2005**, *53*, 6889-6895.

MestReNova (Version 5.2.4) *Mestrelab R*esearch S.L.

Savorani, F.; Picone, G.; Badiani, A.; Fagioli, P.; Capozzi, F.; Engelsen, S.B. Metabolic profiling and aquaculture differentiation of gilthead sea bream by ^{1}H NMR metabonomics. *Food Chem.* **2010**, *120*, 907–914.

Seierstad, T.; Røe, K.; Sitter, B.; Halgunset, J.; Flatmark, K.; Ree, A.H.; Olsen, D.R.; Gribbestad, I.S.; Bathen, T.F. Principal component analysis for the comparison of metabolic profiles from human rectal cancer biopsies and colorectal xenografts using high-resolution magic angle spinning ^{1}H magnetic resonance spectroscopy. *Mol. Cancer* **2008**, 7-33.

Scano, P.; Rosa, A.; Cesare-Marincola, F.; Locci, E.; Melis, M.P.; Dessì, M.A.; Lai, A. ^{13}C NMR, GC and HPLC characterization of lipid components of the salted and dried mullet (*Mugil cephalus*) roe "bottarga". *Chem. Phys. Lipids.* **2008**, *151*, 69-76.

Scano, P.; Rosa, A.; Mereu, S.; Piras, C.; Atzeri, A.; Dessì, M.A. Multivariate fatty acid and fatty alcohol profile of mullet bottarga. *Eur. J. Lipid Sci. Technol.* **2010**, *112*, 1369-1374.

SIMCA-P+ (Version 12.0) Umetrics, Umeå, Sweden.

Standal, I.B.; Gribbestad, I.S.; Bathen, T.F.; Aursand, M.; Martinez, M.I.G. Low molecular weight metabolites in white muscle from cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) analyzed by high resolution ^{1}H NMR spectroscopy. *Magnetic Resonance in Food Science: from Molecules to Man* **2007**, 55-62.

4. Metabolic Fingerprinting of Fiore Sardo, a raw ewe's cheese, by ^{1}H NMR spectroscopy

4.1 INTRODUCTION

Cheese is a well-known milk product which has gained great popularity throughout the world for its health-promoting and organoleptic properties. Cheese making is essentially a dehydration process in which milk casein, fat and minerals are concentrated 6 to 12-fold, depending on the variety. The basic steps common to most cheeses are: acidification, coagulation, dehydration, and salting. Acidification is the major function of the *starter* bacteria, mainly represented by *lactic acid bacteria* (LAB). In particular, lactic acid, produced by the microbial fermentation of carbohydrates, is responsible for the fresh acidic flavour of unripened cheeses and is important in coagulation of milk casein.

During ripening, very complex biological, biochemical, and chemical processes are determined and directed by the microflora cheese curd, which are responsible of the organoleptic qualities of the end product. In fact, while the basic composition and structure of cheese are determined by the curd-producing operations, it is during ripening that the individuality and unique characteristics of each cheese variety develop. In this phase, starter *LAB* play essential roles by producing volatile flavor compounds, by releasing proteolytic and lipolytic enzymes, and by producing natural antibiotic substances that suppress growth of pathogens and other spoilage microorganisms. Although it is generally accepted that the cheese starter flora is the primary defining influence on flavour development, the contribution of *non-starter lactic acid bacteria* (NSLAB), coming from raw milk, is also considered significant, although not been well understood yet (El Soda M.A., 1993).

Even though the importance of the biodiversity of indigenous milk microflora in the manufacturing of high-value traditional cheeses has been widely accepted (De Angelis M. et al., 2001; Pisano M.B. et al., 2007), the great majority of industrialised processes rely on the addition of selected commercial LAB cultures of generic composition to the milk, before coagulation, in order to accelerate cheese ripening, inhibit the development of harmful microflora, reach the desired acidity

and, therefore, to achieve high degree of control over the fermentation process and a standardisation of the final product. Unfortunately, the use of commercial starters, in addition to or in substitution of the indigenous microflora, has also some disadvantages in cheesemaking. Primarily, since the biodiversity of commercial starters is limited, their use often leads to a loss of the uniqueness of the original product as well as a loss of some characteristics that have made the product popular.

Fiore Sardo is a typical Italian hard cheese, produced exclusively in the island of Sardinia according to ancient techniques. It is a "Protected Designation of Origin" (PDO) cheese, traditionally made from raw ewes' milk without the addition of starter cultures. Thus, the maturation is solely due to the actions of the indigenous microbial flora present in the milk and in the dairy environment. Fiore Sardo has been initially produced only at farm level, but its increasing popularity in the national and international market and the innovations in cheese-making technology have recently favoured its manufacture also in industrial plants, complying the manufacturing traditional method reported on the Fiore Sardo specification. Recently, the majority of the Fiore Sardo cheese producers have improved and optimized several technological phases and, according to the EU Hygiene Directive 853/2004, the overall hygienic quality of milk. According to this Directive, Fiore Sardo should be produced from ewe's raw milk containing poor microbial content, not sufficient to carry out a suitable acidification and a proper ripening process. As a consequence, a constant guarantee of the organoleptic and sensory features of Fiore Sardo cheese is currently lacking. To this regard, some recent findings have showed that autochthonous LAB, used as starters or adjunct cultures, can improve, to some extent, the typical sensory characteristics of Fiore Sardo (Pisano M.B. et al., 2007), thus, suggesting their use in cheese making to achieve a better management of the process and maintain the Fiore Sardo "typicality", as previously showed for similar cheeses (Macedo A.C. et al., 2004).

Among the ^{1}H NMR-based metabolomic applications in food science reported in the literature, only few studies have been performed on cheeses (Consonni R. and Cagliani L.R., 2008; Brescia M.A. et al., 2005; Rodrigues D. et al., 2011). In particular, Consonni and Cagliani (2008) investigated the water-soluble metabolite content of Parmigiano Reggiano cheese in terms both of ripening and geographical

origin, while Brescia et al. (2005) focused on the characterisation of the geographical origin of buffalo milk and mozzarella cheese. More recently, Rodrigues et al. (2011) have analysed the metabolic profiles of potential probiotic or symbiotic cheeses, made with pasteurised milk, in relation to the probiotic strain, type of cheese, and ripening time.

The main difficulties in establishing a definitive role for microflora in cheese quality are linked to the complex and dynamic nature of this system, leaving knowledge gaps on the interactions between microflora, flavour, and cheese maturation. The aim of the present wok was to perform a NMR-based metabolomic investigation to evaluate the influence of adjunct autochthonous strains on the metabolome of Fiore Sardo during ripening. In particular, the fermentative performance of these strains was evaluated by a complementary analysis of the metabolic profile of cheese and its microbiological characteristics. The effect of these wild strains was compared to that of commercial starters. Our results provide new insights on the potentiality of NMR-based metabolomics for the study of fermentation processes in dairy food matrices produced with raw milk. To the best of our knowledge, there are no ^{1}H NMR-based metabolomic investigations on raw milk cheeses in the literature.

4.2 MATERIAL AND METHODS

4.2.1 Chemicals

All solvents used, of the highest available purity, were purchased from Merck (Darmstadt, Germany). Deuterium oxide (D_2O, 99.9%) and sodium 3-trimethylsilyl-propionate-2,2,3,3,-d_4 (TSP, 98 atom % D) were acquired from Sigma-Aldrich (Milan, Italy).

4.2.2 Bacterial strains

Four different strain combinations were used for Fiore Sardo cheesemaking (Table 4.1). Combinations A1, A2, and A3 contained autochthonous strains, belonging to the Culture Collection of the Department of Public Health, Clinical and Molecular Medicine, University of Cagliari. The LAB strains of the two species

used in combination A2 were previously isolated from Fiore Sardo cheeses (Pisano et al., 2006). The other strains were isolated from raw ewes' milk used in the industrial plant for cheeses-making. The A3 combination included the same strains present in A1 with the exception of the Lactobacillus strain. All of these strains were identified by physiological and biochemical properties and by species-specific PCR, and selected on the basis of acidifying, proteolytic, lipolytic and antagonistic activity against food-borne pathogens and spoilage bacteria (data unpublished). Combination CC contained commercial LAB starter cultures.

Table 4.1 Experimental cheeses and adjunct cultures used in the manufacturing of Fiore Sardo cheese

Batches	Bacterial Species (strain code)
CC	*Commercial LAB starter cultures* (MO536)[a]
A1	*Lactococcus lactis* subsp. *lactis* (3M17LS6) *Enterococcus faecium* (1KFLS6) *Lactobacillus paracasei* (62LP39)
A2	*Lactococcus lactis* subsp. *lactis* (1FS17b) *Lactococcus lactis* subsp. *lactis* (1FS171m) *Lactobacillus plantarum* (7FS171m)
A3	*Lactococcus lactis* subsp. *lactis* (3M17LS6) *Enterococcus faecium* (1KFLS6)

[a] The commercial LAB consists of *Lactococcus lactis* subsp. *lactis* and *Lactococcus lactis* subsp. *cremoris*

4.2.3 Cultures preparation

After cultivation in appropriate medium (MRS, Oxoid, Basingstoke, UK for lactobacilli and M17, Oxoid, for cocci) for two consecutive transfers, the strains were checked for purity and inoculated (0.1 %) in MRS or M17 broth at 30 °C for 16 h. Each culture was centrifuged at 5000 × *g* at 4 °C for 20 min. Pellets were washed twice in phosphate-buffered saline (PBS; 137 mm NaCl, 2.7 mm KCl, 4.3 mm Na_2HPO_4, 1.4 mm K_2HPO_4; pH 7.3), and resuspended in pasteurised ewe'milk at a cell density of approximately 9 log cfu/ml. This cell suspension was used to inoculate cheese milk.

4.2.4 Manufacture of Fiore Sardo cheese

Four cheese groups were manufactured according to the technical disciplinary of Fiore Sardo cheese in an industrial plant (Argiolas Formaggi, Dolianova, Cagliari, Italy) (Figure 4.1): one type of cheese was made with the addition of commercial lactic cultures (CC), while three types of cheeses were prepared using the three different combinations of adjunct cultures (A1, A2 and A3), as reported in Table 4.1.

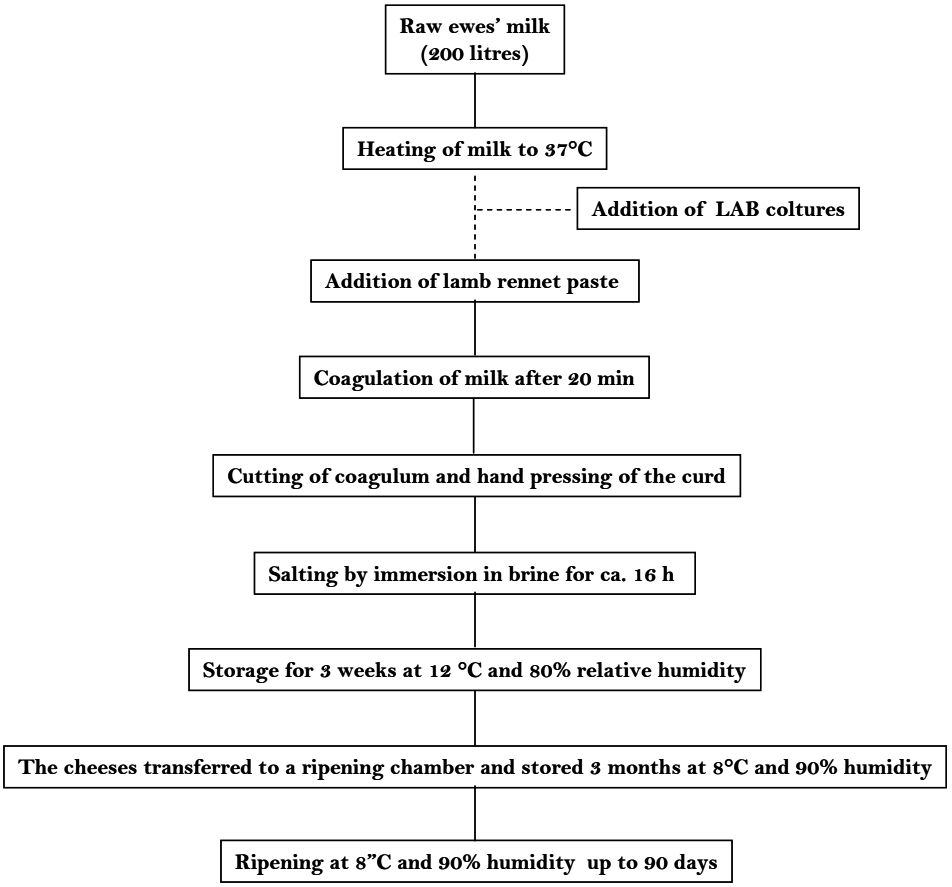

Figure 4.1 Schematic overview of the production of Fiore Sardo

A total of four cheese-making trials were performed. In each trial, two cheese batches were simultaneously produced from two vats of the same bulk raw ewes' milk for a total of eight cheeses batches The distribution of cheeses in cheeses-making was randomised.

Mean composition of raw ovine milk used for cheese-making was 6.43% fat, 5.58% protein, 4% lactose and pH 6.7. Milk in each vat (200 L) was heated to 37 °C and commercial or autochthonous adjunct cultures, grown for 18 h at 30 °C in 2 L of

pasteurized ewe's milk, were inoculated. After 20 min, liquid lamb rennet was added to milk and coagulation took place within 30 min. The curds were cut and kept in stock for about 2 min. Then, the curd pieces were thinly cut into small pieces (the size of a rice grain) that were hand pressed into moulds. After brine salting for 16 h (23 Bé, 230 g NaCl/L), the cheeses were stored for 3 weeks at 12 °C and 80% relative humidity. The cheeses were then transferred to a ripening chamber and stored 3 months at 8°C and 90% humidity. From each batch, samples at 2, 6, 15, 28, and 90 days of ripening were taken and transported to the laboratory under refrigeration.

4.2.5 Chemical Analysis

Milk samples were analysed for total solids (TS) (IDF 1987), fat (IDF 1996) and protein (IDF 1985). In cheese samples, TS (IDF 1982), moisture (calculated as 100-TS), NaCl (IDF 1988), fat (IDF 1986), total nitrogen (TN) (IDF 1964) and protein (calculated as TN × 6.38) were determined. The pH was measured with a HI8520 pH meter (Pool Bioanalysis Italiana, Milan, Italy) on milk and cheese samples (10 g aliquots) taken from at least three different places in the cheese block.

4.2.6 Microbiological and statistical analyses

Ten millilitres of milk or ten grams of cheese were transferred to a sterile tube containing 90 ml of 2% (w/v) sodium-citrate solution. Cheese samples were homogenized in a Stomacher Lab Blender (Pool Bioanalysis Italiana) for two minutes at normal speed. Decimal dilutions were prepared in sterile solution of 0.1% (w/v) peptone and plated in triplicate on specific media to enumerate microbial groups.

Total mesophilic aerobic flora was enumerated on Plate Count agar (PCA, Microbiol, Cagliari, Italy), using the pour plate method. The plates were incubated at 30 °C for 48–72 h.

Enterobacteriaceae and *Escherichia coli* were determined on Violet Red Bile Glucose agar (VRBGA, Microbiol) and Triptone Bile X-gluc agar (TBX, Microbiol), respectively, using the pour plate method. VRBG and TBX plates were covered with a layer of the same culture medium before incubation at 37 °C and 44 °C for 24 h, respectively.

Enterococci were isolated by surface plating on KF-*Streptococcus* agar (Oxoid, Basingstoke, UK) supplemented with 1% TTC, incubated for 48 h at 42 °C.

Lactic acid bacteria were counted on de Man, Rogosa Sharpe Agar (MRS, pH 5.4; Oxoid) and on M17 (Microbiol) incubated at 30 °C for 72 h, under microaerophilic and aerobic conditions, respectively.

Microbial counts were calculated as number of colony forming units (cfu) *per* gram or ml of sample and reported as $\log_{10}$ cfu/g or ml. Calculations of Standard Error were also performed. One-way ANOVA was performed on data obtained analyses for each sampling date during ripening considering one main effect, adjunct cultures type. Tukey test for multiple comparisons was used to separate treatment means. All statistics were performed using GraphPad Prism Statistics software package version 3.00 (GraphPad Prism Software Inc., San Diego, CA, USA). Statistical significance was inferred at $P < 0.05$.

4.2.7 Low molecular weight metabolite extraction

At each ripening time, a slice of cheese (ca 130 mg) was freeze-dried, after removal of 1 cm of crust around, and then pulverized in a ceramic mortar, before carrying out a methanol-chloroform-water extraction, according to the procedure previously described by Folch J. at al. (1957). Each extraction was performed in triplicate. This setup yielded a total of 120 samples for each of which a high resolution ^{1}H NMR spectrum was recorded.

Height portions (each portion ca 0.2 g) were dissolved in 12 mL of a mixture chloroform-methanol (2:1, vol/vol). Samples were kept 1 hour in darkness. After the addition of 4 ml of H_2O in each tube and centrifugation at 1000 rpm for 20 min at 4°C, the $CHCl_3$ fraction was separated from the MeOH/H_2O mixture, containing the water-soluble low-molecular-weight components. The water/methanol phase was dried and the residue was redissolved in 1 mL of D_2O. An aliquot of 600 µL was placed into a 5 mm NMR tube, adding 50 µL of TSP in D_2O solution (0.80 mM final concentration).

4.2.8 ^{1}H NMR Spectroscopy

^{1}H NMR experiments were carried out at 298 K on a Varian UNITY INOVA 400 spectrometer operating at 399.95 MHz. For each sample, 252 scans were recorded at 298 K with the following parameters: spectral width of 5623 Hz, a 90° pulse of 6.8 μs, acquisition time of 3s, and relaxation delay of 3s. The residual water signal was suppressed by applying a presaturation technique with low power radiofrequency irradiation for 1.5 s. The FIDs were multiplied by an exponential weighting function equivalent to line broadening of 0.3 Hz prior to Fourier transformation. ^{1}H NMR chemical shifts were referenced to TSP at δ 0.0 ppm.

4.2.9 NMR Data Pre-processing

The NMR spectra were phased and automatically baseline corrected by ACDlab (Advanced Chemistry Development, V. 12.01 2010). The same software was used to export data in ASCII format. The chemical shift region between 4.50 and 4.98 ppm (corresponding to the water resonance) was removed prior to statistical analysis in order to eliminate any spurious effect of variability in the suppression of water. The chemical shift region between -0.5 and 0.5 ppm, corresponding to the TSP resonance, was also removed. In this manner, spectral data points were reduced to 18945 points and, then, a spectral dataset with the dimensions 120x18945 was constructed for subsequent multivariate data analysis.

The spectral data were imported into MATLAB (R2006a; Mathworks, Inc., Natick, MA). Since several shifts in peak positions were observed, especially for those metabolites hardly dependent on pH variations, spectra were aligned using an in-house modified version of Correlation Optimised Shifting (Coshift) able to perform the Co-shift in localised regions of the spectrum (Savorani F. et al., 2010). Then, spectral data set was normalized, to avoid dilution effects of samples, and the data matrices was pareto-scaled by SIMCA-P+ version 12.0 (Umetrics, Sweden). The quality of the models was described by R^2 and Q^2 values. R^2 is defined as the proportion of variance in the data explained by the models and indicates goodness of fit. Q^2 is defined as the proportion of variance in the data predictable by the model and indicates predictability.

4.3 RESULTS AND DISCUSSION

4.3.1 Cheeses compositional and microbiological characteristics

The main physicochemical characteristics of experimental Fiore Sardo cheeses during ripening are shown in Table 4.2. A clear indication of a dynamic metabolic activity of strains in all cheeses is provided by the decrease in pH throughout ripening. The pH values in 2-days-old cheeses ranged from 5.03 to 5.15, similarly to that found for Fiore Sardo and other raw milk cheeses with longer ripening times (Macedo A.C. et al., 1995; Macedo A.C. et al., 2004; Pisano M.B. et al., 2006; Pisano M.B. et al., 2007). In particular, the pH value of A3 cheese was slightly higher ($P>0.05$), than those of cheeses made with A1, A2, and CC combinations while no significant differences were observed among these three types of cheeses ($P<0.05$). The pH profile followed a similar trend in all batches, decreasing between 2 and 15-day-ripening and, then, slightly increasing. At the end of ripening, CC and A3 cheeses exhibited pH values higher than those in A1 and A2 cheeses. This result may be related to the inclusion of lattobacilli strains (i.e. NSLAB) in the A1 and A2 combinations. Indeed, as well documented in the literature (Fox P.F. et al., 1993), while LAB growth during the early stage of ripening, NSLAB survives for longer periods, forming lactic acid by fermentation of the residual lactose.

Water activity increased during the first week of ripening, achieving values around 0.85-0.90, and, then decreased throughout ripening with no significant ($P<0.05$) differences among cheeses elaborated with different cultures. Humidity decreased during ripening in all batches. There were no statistical differences in the percentage of fat, protein, and NaCl among all types of cheeses and thought the ripening time.

The evolution of different microbial counts throughout ripening of Fiore Sardo is reported in Tables 4.3. Previous works on this cheese have revealed that its indigenous microflora is mainly constituted by homofermentative cocci and facultative heterofermentative lactobacilli[3]: *Lc. lactis* subsp. *lactis* and *E. faecium*

[3] *Homofermentative LAB* ferment glucose to lactic acid as the primary by-product; *heterofermentative LAB* ferment glucose to lactic acid, ethanol, acetic acid and carbon dioxide as by-products.

Table 4.2 Mean physicochemical characteristics* of Fiore Sardo cheeses during ripening

Parameters	Cheese	Days of ripening				
		2	**6**	**15**	**30**	**90**
pH	CC	5.03±0.02	4.92±0.04	4.41±0.37	4.93±0.12	5.07±0.08
	A1	5.09±0.01	4.99± 0.06	4.81±0.04	4.87±0.01	4.89±0.02
	A2	5.09±0.00	4.89 ±0.07	4.75±0.02	4.78±0.08	4.89±0.02
	A3	5.15±0.05	5.12 ±0.08	4.95±0.16	5.10±0.13	5.20±0.08
a_w	CC	0.81±0.01	0.85±0.06	0.81±0.05	0.82±0.06	0.72±0.05
	A1	0.84±0.02	0.85±0.05	0.82±0.05	0.81±0.05	0.71±0.01
	A2	0.87±0.01	0.90±0.01	0.88±0.03	0.88±0.02	0.76±0.02
	A3	0.85±0.05	0.84±0.03	0.82±0.06	0.80±0.06	0.71±0.03
Moisture (%)	CC	46.01±0.66	45.24±0.42	43.72±0.62	42.13±0.09	40.54±0.58
	A1	46.41±0.48	43.48±0.6	44.45±0.83	43.10±1.44	42.07±0.04
	A2	46.50±1.4	46.10±1.05	45.14±1.54	43.34±0.62	42.41±0.93
	A3	44.18±1.9	42.99±1.4	42.18±0.95	41.20±1.2	40.91±0.33
Fat (%)	CC	26.59±0.20	26.58±0.14	27.08±0.16	28.08±0.43	28.72±0.20
	A1	26.33±0.31	26.61±1.26	27.12±0.56	27.44±0.81	28.06±0.17
	A2	26.49±0.8	25.88±0.68	26.23±1.16	27.03±0.47	27.54±0.69
	A3	28.01±0.92	27.89±0.77	28.00±0.5	28.34±0.26	28.68±0.12
FDM (%)	CC	49.24±0.22	48.53±0.11	48.11±0.23	48.69±0.67	48.74±0.58
	A1	49.12±0.14	48.35±0.45	48.82±0.27	48. 22±0.2	48. 43±0.33
	A2	49.31±0.40	48. 01±0.32	47.79±0.78	47.72±0.33	47.82±0.43
	A3	50.18±0.08	48.91±0.12	48.42±0.06	48.20±0.54	48.52±0.06
Protein (%)	CC	22.93±0.39	22.94±0.38	23.26±0.44	23.94±0.14	24.48±0.3
	A1	22.73±0.27	22.76±1.06	23.01±0.6	23.45±0.77	23.94±0.05
	A2	22.97±0.39	22.23±0.39	22.54±0.73	23.08±0.2	23.62±0.52
	A3	24.17±1.4	24.09±1.32	24.25±1.21	24.81±1.2	24.65±0.02
Salt (%)	CC	0.80±0.02	1.35±0.06	1.64±0.00	1.87±0.12	2.17±0.08
	A1	0.79±0.01	1.33±0.04	1.48±0.07	1.79±0.05	2.11±0.12
	A2	0.77±0.03	1.26±0.13	1.59±0.07	1.78±0.04	2.01±0.13
	A3	0.79±0.06	1.35±0.09	1.57±0.16	1.63±0.17	2.25 ±0.07

*Values are the mean ± standard error of two independent experiments

Table 4.3 Counts* of the main microbial groups, expressed as $\log_{10}$ cfu·g^{-1} or ml^{-1}, in Fiore Sardo cheese during ripening

Microbial groups	Cheese	Days of ripening				
		2	**6**	**15**	**30**	**90**
Aerobic mesophiles	CC	8.44±0.74	9.55± 0.15	9.43± 0.07	9.53± 0.03	7.80 ±0.20
	A1	9.02±0.52	10.59±0.11	10.60± 0.09	10.54±0.06	8.60±0.03
	A2	9.51±0.95	9.93± 0.68	10.80± 0.01	10.15±0.49	8.39±0.31
	A3	8.23±0.15	9.64± 0.08	9.60 ±0.15	9.50±0.38	7.90±0.30
Lactococci	CC	8.59±0.59[a]	9.63± 0.15	9.62±0.08	8.51±0.33	6.78±0.06
	A1	8.71±0.83	9.78± 0.18	9.65±0.35	8.96±0.04	7.05±0.35
	A2	9.53±0.06	9.27± 0.31	9.75±0.06	8.61±0.03	6.56±0.32
	A3	9.53±0.06	9.27± 0.31	9.75±0.06	8.61±0.03	6.30±0.32
Enterococci	CC	6.09±0.04	5.92± 0.12	5.94±0.34	6.25±0.19	5.30±0.09
	A1	7.14±0.56	7.82± 0.34	8.03±0.61	8.52±0.25	6.57±0.03
	A2	6.54±0.05	6.39± 0.42	6.74±0.06	6.60±0.08	5.52±0.04
	A3	6.65±0.22	7.44± 0.21	7.64±0.05	6.79±0.16	6.15±0.15
Enterobacteriaceae	CC	6.82±0.82	6.24± 0.04	5.66±0.18	5.92±0.52	< 2
	A1	6.09±1,09	5.90± 0.00	5.15±0.15	3.50±0.50	< 2
	A2	6.82±0.06	6.30± 0.04	6.49±0.29	4.39±0.61	<2
	A3	6.28±0.06	5.37± 0.39	5.51±0.23	3.92±0.62	<2
E. coli	CC	6.20±0.20	5.20± 0.84	4.50±0.80	3.66±0.33	< 2
	A1	5.20±0.20	4.35± 1.35	3.30±0.70	2.50±0.50	< 2
	A2	5.30 ±0.20	5.17± 0.07	3.78±0.70	3.02±0.15	<2
	A3	4.23±0.71	3.31± 0.14	2.99±0.79	2.77±0.77	<2
Lactobacilli	CC	5.30±0.20	7.72± 0.31	7.96±0.15	8.34±0.34	6.46±0.66
	A1	7.28±0.02	9.79± 0.15	9.07±0.44	9.25±0.32	8.89±0.30
	A2	6.98±0.42	9.77± 0.34	8.37±0.33	9.17±0.09	8.99±0.14
	A3	6.66±0.09	7.70± 0.31	7.20±0.09	7.39±0.19	6.39±0.39

*Values are the mean ± standard error of two independent experiments

among cocci, *Lb. paracasei* subsp. *paracasei* and *Lb. plantarum* among lactobacilli were the predominant species isolated during ripening (Mannu L. et al., 2000; Pisano M.B. et al., 2006). The data in Table 4.3 show that aerobic mesophiles increased in all four cheese batches up to 30 days of aging and, then, declined. A1 and A2 cheeses exhibited the highest values. This result can be reasonably linked to the composition of the added strain combination, A1 and A2 including more bacterial specie compared to A3 and CC.

No marked differences were observed among the lactococci counts of cheeses, although their develop slightly differently during ripening: the maximum count in cheeses made with A2 and A3 combinations was obtained at the second week of ripening, while between the first and the second week in CC and A1 chesses. In all batches, the increase in lactococci count paralleled the drop in pH values. This behaviour is logical, if one take into consideration that lactococci are the main producers of lactic acid. The decrease of lactococci count in the later stages of ripening may be explained by inhibitor effects determined by the low pH and a_w values, and high NaCl concentrations. A further contribution may arise from the scarce capacity of lactococci to compete with other more acid-resistant microbial populations such as lactobacilli (Dasen A. et al., 2003). The progressive disappearance of the lactococci during the period of ripening has been observed also in other types of cheeses (Rodriguez Medina M.L. et al., 1995; Garcia Fontan M.C. et al., 2001).

Numbers of enterococci were similar to those observed in other cheese varieties from raw milk (Medina M. et al., 1992; Suzzi G. et al., 2000, Psoni L. et al., 2003). The presence of enterococci in high numbers is a common trait, not only in the case of raw milk cheese, but for other cheeses as well, like Manchego (Ordonez J.A. et al., 1978), Mozzarella (Coppola S. et al., 1988), Kefalotyri (Litopoulou-Tzanetaki E. et al., 1990), and Serra (Macedo A.C. et al., 1995). Whether their presence in matured cheeses contributes to the development of good and acceptable flavour or not is still debated. Indeed, on one hand, high levels of enterococci are considered to lead to deterioration of some organoleptic characteristics in certain cheeses (Thompson T.L. and Marth E.H., 1986; Lopez-Diaz T.M. et al., 1995), while, on the other hand, many reports pinpoint the desirable role that enterococci have in cheese production and quality (Centeno J.A. et al., 1996; Centeno J.A. et al., 1999). In the experimental

cheeses under investigation, the evolution that chemical and physico-chemical parameters undergo with ripening does not create unfavourable conditions to the growth of enterococci, their counts keeping relatively high during ripening. When *Enterococccus* strains were employed as a co-inoculum in experimental cheeses (i.e. in A1 and A3 combinations), higher mean counts were detected, especially in the middle stages of ripening.

The numbers of microorganisms indicative of the hygienic level, *Enterobacteriaceae* and *coliforms*, were relatively high at the beginning of ripening of all cheeses; this contamination, usual for raw milk cheeses (Hatzikamari M. et al., 1999; Macedo A.C. et al., 1995; Nikolaou E. et al., 2002; Zarate V. et al., 1997), may arises from raw milk during milking, unrefrigerated storage and transportation, and, possibly, from manufacturing. High levels of *Enterobacteriaceae* in raw milk cheeses are of great concern for the dairy industry because of their technological and public health significance. In all batches under investigation, ripening had a positive influence on the numbers of *Enterobacteriaceae* and *coliforms*, both decreasing up to undetectable levels on day 90. Various factors may contribute to this decline, including a significant increase in the concentration of NaCl (Zarate V. et al., 1997) and the inhibition of these bacteria by lactic acid bacteria (Nũnez M. et al., 1985), basically by causing a decrease in the pH and an increase in lactic acid concentration (Zarate V. et al., 1997). In particular, it is known that *Enterobacteriaceae* require values of pH inferior to 5.0–5.20 for inhibition (Medina M. et al., 1991). Since pH values below this range were recorded under our experimental conditions, it is likely that pH is an important agent regulating the survival of *Enterobacteriaceae* and *coliforms* in Fiore Sardo cheese.

Lactobacilli, which usually dominate the NSLAB microflora (Fox P.F. et al., 1993), increased significantly their numbers until day 30 of ripening. Some species of lactobacilli resist high salt concentrations and pH (Desmazeaud M. and Piard J.C., 1992) and this explains their ability to grow in cheese compare to lactococci. Their growth was more pronounced in A1 and A2 cheeses, likely due to the contribution of lactobacilli used as co-inoculum in cheesemaking. Lactobacilli decreased significantly in 3-month-old CC and A3 cheeses, while their count kept almost constant in A1 and A2 cheeses.

4.3.2 ^{1}H NMR spectra of aqueous extract of Fiore Sardo

Figure 4.3 displays representative NMR spectra of the aqueous extract of Fiore Sardo at 2, 15, and 90 days of ripening. The identification of ^{1}H NMR signals was performed using the Human Metabolome Database (HMD) (www.hmdb.ca), Biological Magnetic Resonance Data Bank (BMRB) (www.bmrb.wisc.edu), and previous literature data (Brescia M.A. et al., 2005: Consonni R. et al., 2008; Gianferri R. et al., 2007). Some minor signals or overlapped resonances were assigned by performing 2D conventional NMR experiments (COSY and TOCSY). In some cases, validation of the peak attribution was achieved by spiking with pure standard compounds. This allowed the identification of metabolites belonging to the categories of carbohydrates, organic acids, amino acids, and phenolic compounds. Chemical shifts are summarized in Table 4.4.

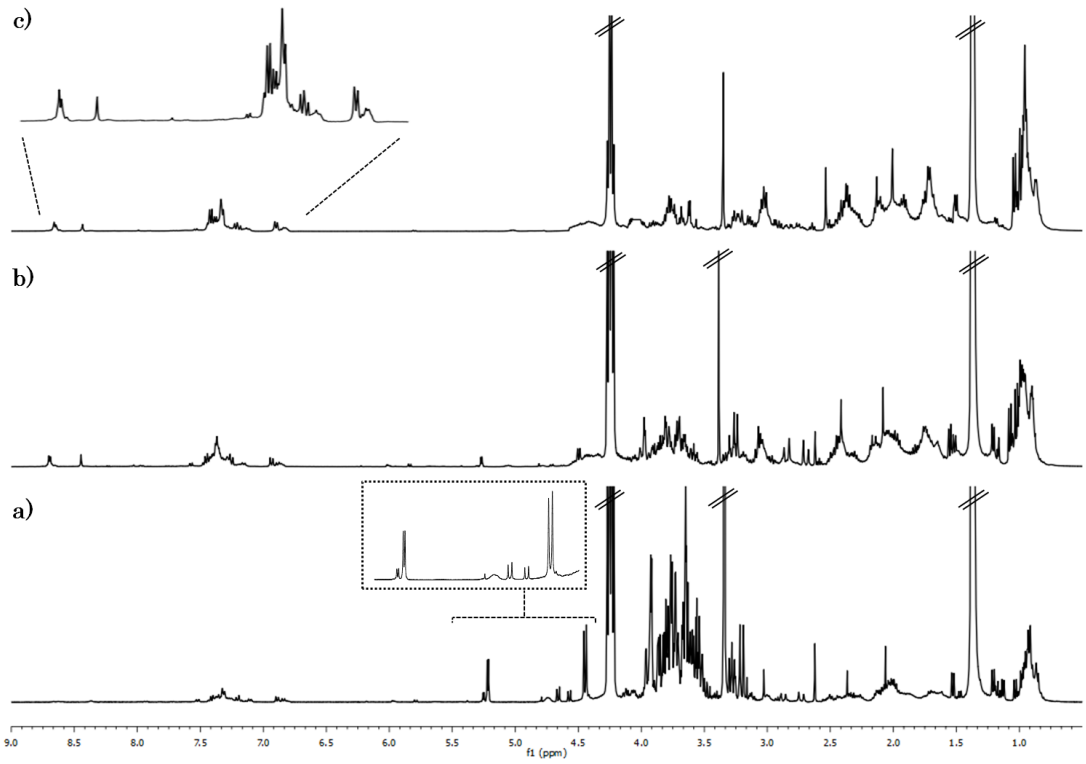

Figure 4.3 ^{1}H-NMR spectra for Fiore Sardo cheese aqueous extract at (a) 2, (b) 15, and (c) 90 days of ripening.

As expected, the most predominant signals in the NMR spectra of all cheese extracts belonged to lactic acid (1.36 and 4.23 ppm), the main product from the metabolism of lactose. Besides, citric acid was recognized by its characteristic signals at 2.70 and 2.80 ppm, while acetic, succinic, formic, and pyruvic acid were identified by their unique singlets at 2.06, 2.63, 8.44, and 2.37 ppm, respectively.

The sugar region (3.5-5.5 ppm) showed signals of the anomeric protons of α-glucose (5.22 ppm), β-glucose (4.65 ppm), α-galattose (5.25 ppm), and β-galattose (4.58 ppm). In addition, a multiplet a 5.02 ppm was observed at increasing ripening times and assigned to unidentified oligosaccharides.

Signals detected in the high-field region of the spectra (0.5-3.5 ppm) arose mainly from aliphatic amino acids. Only signals from alanine, valine, leucine, and isoleucine were clearly identified. Indeed, several barely resolved multiplets fall in this spectrum zone, making the signal assignments difficult. To this purpose, the use of 2D NMR experiments has been very useful. For instance, glutamic acid and glutamine give weak overlapped signals in the 2.05-2.15 ppm zone, which were separated in the second dimension of the TOCSY experiment. Partially overlapped signals of aspartic acid (2.73 and 2.834 ppm) and asparagine (2.88 and 2.96 ppm) were assigned on the basis of the TOCSY correlations with the α hydrogens centered at 3.893 and 4.01 ppm, respectively (Figure 4.4).

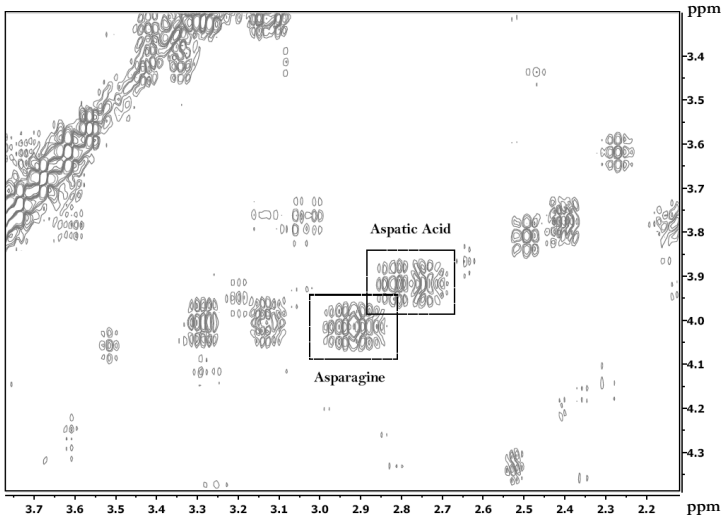

Figure 4.4 Expansion of the 2D TOCSY spectrum of Fiore Sardo aqueous extract at 90 days of ripening.

Choline (Cho) and phosphorylcholine (P-Cho) were identified in the spectra. The $-N(CH_3)_3$ group of these two compounds originates two singlets resonating at 3.19 and 3.22 ppm for Cho and P-Cho, respectively. Part of the presence of these compounds in cheese extracts may be ascribed to hydrolytic actions on the phosphatidylcholine's family by lypolisis during ripening.

Furthermore, signals from short-chain free fatty acids were also observed, such as 3-hydroxyburyrate (Figure 4.5). The presence of these compounds in the aqueous extract of cheese is mainly related to the intensity of bacterial fermentation that took place during ripening.

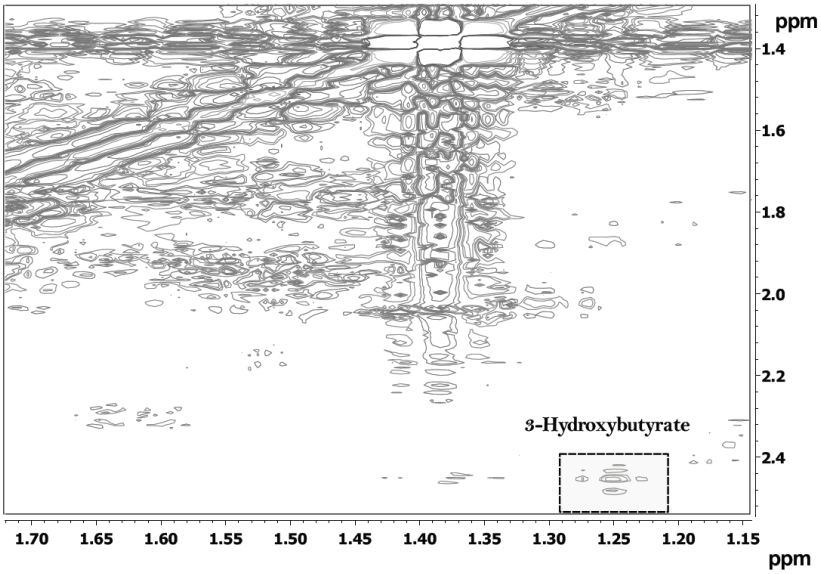

Figure 4.5 2D TOCSY expansion of Fiore Sardo aqueous extract, showing the cross-peak of 3-hydroxyburyrate

The weakest signals of the NMR spectra resonated within the low-field region (5.5–8.5 ppm) and belonged mainly to aromatic groups of amino acids. The aromatic protons of histidine, phenylalanine, and tryptophan were identified. Moreover, the nucleobases uracyl was also detected. The doublet at 7.18 ppm (J=8.1 Hz), which is coupled to another doublet at 6.89 ppm (J=8.1 Hz) in COSY and TOCSY spectra, was attributed to tyrosine. At 90 days of ripening, an other doublet appeared at 7.22 ppm (J=8.1 Hz) that was assigned to tyramine according

to the COSY correlation of this peak with the doublet at 6.90 ppm (J=8.1 Hz) (Figure 4.6).

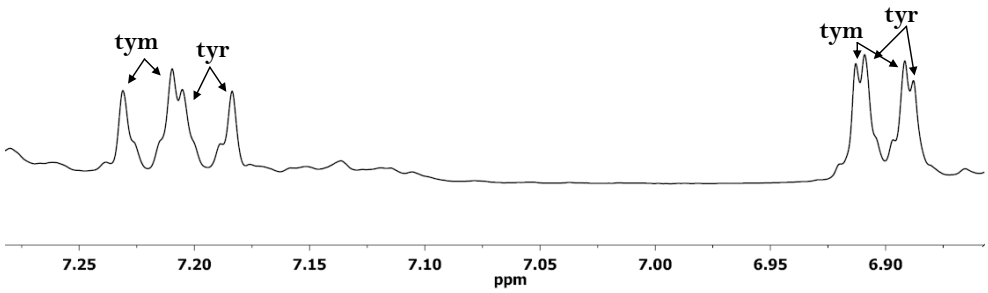

Figure 4.6 Expansion of the ^{1}H-NMR spectrum of Fiore Sardo aqueous extract. Signal assignments relative to tyrosine (tyr) and tyramine (tym) are reported.

Tyramine is a biogenic amine (BA), often produced by decarboxylation of tyrosine in food and vegetable during fermentation. The accumulation of BA in cheese is quite common throughout the ripening period and arises mainly from the enzymatic decarboxylation of amino acids by microorganisms (Joosten H. et al., 1987; Pinho O. et al., 2001), belonging largely to NSLAB (lactobacilli, pediococci, and enterococci) and to Enterobacteriaceae (Figure 4.7). In particular, Enterobacteriaceae are able to decarboxylate lysine to produce cadaverine, certain lactobacilli show also high histidine decarboxylase activity, while the most relevant decarboxylase activity from enterococci isolated from dairy products concerned tyrosine (Joosten H. et al., 1989). Besides tyramine, also histamine was detected in cheeses at 90 days of ripening.

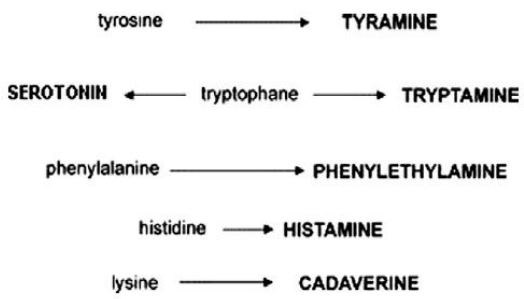

Figure 4.7 Synthesis of biogenic amine

Table 4.4 Summary of the metabolites identified in the 400 MHz ^{1}H spectrum of the aqueous extract of Fiore Sardo

Carbohydrates

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlations
α–Galactose (α-Gal)	C1H	5.25	d	3.76 (C)
	C2H	3.76	dd	5.25, 3.71 (C)
	C3H	3.71	m	3.76 (C)
β–Galactose (β-Gal)	C1H	4.58	d	3.51 (C)
	C2H	3.51	dd	4.58, 3.67 (C)
	C3H	3.67	m	3.51 (C)
α–Glucose (α-Glc)	C1H	5.22	d	3.55 (C)
	C2H	3.55	dd	5.22 (C)
	C3H	3.71	t	//
	C4H	3.40	t	//
	C5H	3.87	m	3.82 (C)
	C6H	3.82	dd	3.87(C)
β–Glucose (β-Glc)	C1H	4.65	d	3.27 (C)
	C2H	3.27	dd	4.65, 3.65 (C)
	C3H	3.65	t	3.27 (C)
	C4H	3.38	t	3.44 (C)
	C5H	3.44	m	3.38 (C)

Organic acids

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlations
Acetic acid (Ace)	CH_3	2.06	s	//
Citric acid (Cit)	$^{2,4}CH$	2.80	dd	2.71, 2.86 (C)
Formic acid (Form)	$HCOO^-$	8.44/8.35	s	//
Lattic acid (Lac)	CH_3	1.36	d	4.23 (C)
	CH	4.23	q	1.36(C)
Pyruvate (Pyr)	CH_3	2.377	s	//

Table 5.4 (Continued)

Biogenic amines

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlations
Histamine (Hys)	C7H, ring	7.34	s	//
	C5H, ring	7.42	s	//
Tyramine (Tym)	C1H	2.76	dd	2.61 (C)
	C2H	2.61	dd	2.76 (C)
	C3,4H, ring	6.69	d	7.20 (C)
	C5,6H, ring	7.20	d	6.69 (C)

Other Compunds

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlations
Ethanol (Et)	C1H	1.14	t	3.65 (C)
	C2H	3.65	q	1.14 (C)
Hypoxanthine (Hyp)	CH-2	8.22	s	//
	CH-8	8.18	s	//
Uracil (Ura)	C5H, ring	5.78	d	7.53 (C)
	C6H, ring	7.53	d	5.79 (C)
Choline (Cho)	N-$(CH_3)_3^+$	3.19	s	//
	N-CH_2	4.05	m	3.52 (C)
	O–CH_2	3.52	m	4.05 (C)
4-aminobutyrate (GABA)	αCH_2	2.35	t	1.92 (C)
	βCH_2	1.92	q	3.02, 2.35 (C)
	γCH_2	3.02	t	1.92 (C)

Table 4.4 (Continued)

Amino acids

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlation[c]
Alanine (Ala)	αCH	3.79	q	1.47 (C)
	βCH_3	1.47	d	3.79 (C)
Arginine (Arg)	αCH	3.78	t	1.93 (C)
	βCH_2	1.93	m	3.78, 1.73 (C)
	γCH_2	1.73	m	1.93, 3.26 (C)
	δCH_2	3.26	t	1.73 (C)
Aspartate (Asp)	αCH	3.93	dd	2.83 (C)
	βCH_2	2.73	dd	3.93 (C)
	$\beta' CH_2$	2.83	dd	3.93 (C)
Asparagine (Asn)	αCH	4.01	dd	2.88, 2.96 (C)
	βCH_2	2.88	dd	4.01 (C)
	$\beta' CH_2$	2.96	dd	4.01 (C)
Carnitine (Carn)	αCH	2.50	m	4.56(C)
	βCH	4.56	m	2.50 , 3.47(C)
	γCH	3.47	m	4.56 (C)
	$N\text{-}(CH_3)_3^+$	3.22	s	//
Creatine (Crt)	$N\text{-}CH_3$	3.03	s	//
	$N\text{-}CH_2$	3.93	s	//
Phosphocreatine	$N\text{-}CH_3$	3.05	s	//
	$N\text{-}CH_2$	3.98	s	//
Creatinine (Crn)	$N\text{-}CH_3$	3.12	s	//
	$N\text{-}CH_2$	4.07	s	//

Table 4.4 (Continued)

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlation[c]
Glycine (Gly)	αCH_2	3.57	s	//
Glutamate (Glu)	αCH	3.76	dd	2.09 (C)
	βCH_2	2.09	m	2.34, 3.76 (C)
	γCH_2	2.34	m	2.09 (C)
Glutamine (Gln)	αCH	3.79	d	2.12 (C)
	βCH_2	2.12	m	2.45, 3.79 (C)
	γCH_2	2.45	m	2.12(C)
Histidine (His)	C2H, ring	8.66	s	7.39 (C)
	C4H, ring	7.39	s	8.66 (C)
Isoleucine (Ile)	αCH	3.68	m	1.96 (C)
	βCH	1.98	m	1.00, 1.22, 3.68 (C)
	γCH	1.48	m	0.96, 1.22(C)
	$\gamma' CH$	1.22	m	0.96, 1.48, 1.98(C)
	$\gamma' CH_3$	1.00	d	1.98 (C)
	δCH_3	0.96	t	1.22, 1.48(C)
Leucine (Leu)	αCH	3.72	t	1.72(C)
	βCH_2	1.72	m	3.72 (C)
	γCH	1.77	m	0.94 (C)
	$\delta CH_3, \delta' CH_3$	0.94	d	1.77 (C)

Table 4.4 (Continued)

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlation[c]
Lysine (Lys)	αCH	3.77	t	1.94(C)
	βCH_2	1.94	m	3.77 (C)
	γCH_2	1.45	m	1.73 (C),
	δCH_2	1.73	m	1.45, 3.02 (C)
	εCH_2	3.02	t	1.73 (C)
Methionine (Met)	αCH	3.87	t	2.18 (C)
	βCH_2	2.18	m	2.65, 3.87 (C)
	γCH_2	2.65	t	2.18 (C)
	S-CH_3	2.13	s	//
Phenylalanine (Phe)	αCH	4.01	dd	3.12, 3.26 (C)
	βCH	3.26	dd	4.01 (C)
	β'CH	3.12	dd	4.01 (C)
	C2,6H, ring	7.42	m	//
	C3,5H, ring	7.40	m	7.32 (C)
	C4H, ring	7.32	m	7.42 (C)
Proline (Pro)	αCH	4.15	t	2.03, 2.34 (C)
	βCH	2.34	m	4.15 (C)
	β'CH	2.03	m	4.15 (C)
	γCH_2	2.02	m	3.37, 3.42 (C)
	δCH	3.37	t	//
	δ'CH	3.42	t	//
Serine (Ser)	αCH	3.85	dd	3.95 (C)
	βCH	3.95	dd	3.85 (C)

Table 4.4 (Continued)

Compound	Group	^{1}H (ppm)[a]	^{1}H Multiplicity[b]	Correlation[c]
Threonine (Thr)	αCH	3.58	d	4.18 (C)
	βCH	4.18	m	1.35 (C)
	γCH_3	1.35	d	4.18 (C)
Tryptophan (Trp)	αCH	4.07	dd	//
	βCH_2	3.50	dd	4.07 (C)
	$\beta' CH_2$	3.28	dd	//
	C2H, ring	7.31	s	//
	C4H, ring	7.71	d	7.20 (C),
	C5H, ring	7.20	t	7.27, 7.71 (C)
	C6H, ring	7.27	t	7.20, 7.54 (C)
	C7H, ring	7.54	d	7.27 (C)
Tyrosine (Tyr)	αCH	3.95	dd	3.04 (C)
	βCH	3.22	dd	3.04(C)
	β'CH	3.04	dd	3.22, 3.95 (C)
	C2,6H, ring	6.89	d	7.18 (C)
	C3,5H, ring	7.18	d	6.89 (C)
Valine (Val)	αCH	3.62	d	2.27 (C)
	βCH	2.27	m	0.99, 1.04, 3.62 (C)
	$\gamma' CH_3$	1.04	d	0.99, 2.27 (C)
	γCH_3	0.99	d	1.04, 2.27 (C)

[a] ^{1}H chemical shifts re reported with respect to TSP signal (0.00 ppm).
[b] Multiplicity definitions: s, singlet; d, doublet; t, triplet; q, quartet; dd, doublet of doublets; m, multiplet.
[c] Experiment legend: C, COSY; T, TOCSY.

4.3.3 Impact of ripening on the metabolic profile of cheeses

Two different sources of variation (i.e. strain species and ripening time) and their impact on the metabolic profile of cheese were considered in this study. Although the visual inspection of NMR spectra revealed some changes in the levels of metabolites either in term of added strain or ripening time, the analysis of the whole set of spectra was virtually impossible, thus, making obvious the need for multivariate statistical analysis to reduce the complexity of the pool of NMR data and provide a comparative interpretation of the metabolic changes.

An explorative PCA analysis of the overall data set was firstly performed. Figure 4.8 shows the PCA score plot (PC1 vs PC2), accounting for ca. 70% of the total variance. On the upper side of the Figure, samples are labelled according to the combination of the added strains, while, on the bottom side, samples are coloured according to the fermentation time.

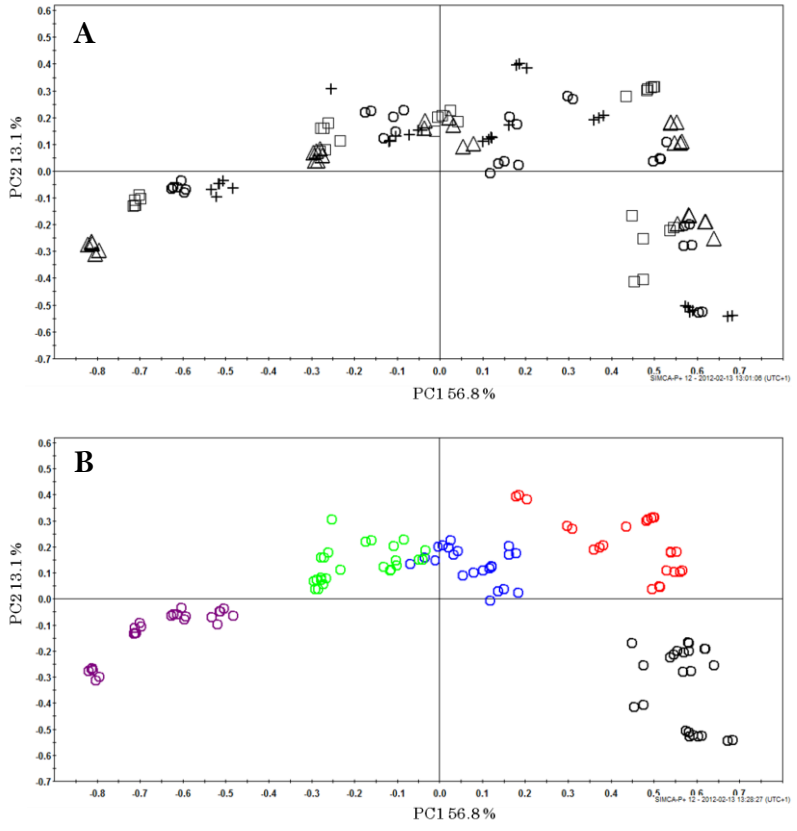

Figure 4.8 PCA scores plot derived from the ^{1}H NMR spectra of Fiore Sardo cheese aqueous extracts: (A) symbols indicate the added strain combinations: CC (triangle), A1 (cross), A2 (circle), and A3 (square); (B) colours indicates the ripening period: 2 (black), 6 (red), 15 (blue), 28 (green), and 90 (violet) days .

As it can be noted, a progressive change of the metabolic profile of cheeses occurs from the right to the left part of the plot as the ripening period increases, indicating continuous fermentation. Furthermore, it is interesting to note that, for each starter combination, a different behaviour is observed, fermentation being the fastest in CC cheeses, whereas the slowest in cheeses made with A1 combination.

For the purpose of a comparative analysis of the metabolic changes over time, a temporal pairwise comparison PCA model was calculated for each combination of added cultures. Figures 4.9-4.12 show the results relative to the cheeses manufactured with CC, A1, A2, and A3 combinations, respectively. The statistics for differentiating cheeses during maturation showed high goodness of fit and predictability with R^2 values ranging from 0.77 to 0.96 and Q^2 values from 0.69 to 0.92. Due to the high number of variables in the spectral dataset under investigation (i.e. 18945), the best way to identify the metabolites mainly affected during the course of ripening was provided by the analysis of the loading line plots (Figures 4.9-4.12 E-H): the upper section of the line plot represents metabolites that were higher in samples cluster on the positive side of PC1, whereas the lower section represents metabolites that were higher in samples cluster on the negative side of PC1. Comparison among the data indicates similar metabolic differentiations over time in all of the batches. For a clearer description of these modifications, the data discussion will carry on per molecular class.

Carbohydrates. One of the most evident modifications during ripening is the consumption of carbohydrates. Indeed, as pointed out by the comparison among the middle field spectral regions, the amount of sugars were relatively high in the early stage of ripening compared to the later stage. This result is reasonable expectable, since the primary event in cheese manufacture is the complete and rapid metabolism of lactose and its constituent monosaccharides by LAB (Leroy F. et al., 2004; Marilley L. et al., 2004). This biochemical transformation is essential for the production of good quality cheese, since the presence of fermentable carbohydrates may lead to the development of an undesirable secondary flora (Fox P.F. et al., 1990). In particular, the decrease of the carbohydrate level at the beginning of ripening is ascribable to their conversion into lactic acid via pyruvate by Homofermentative starter LAB. Pyruvate, in turn, can also lead to the generation

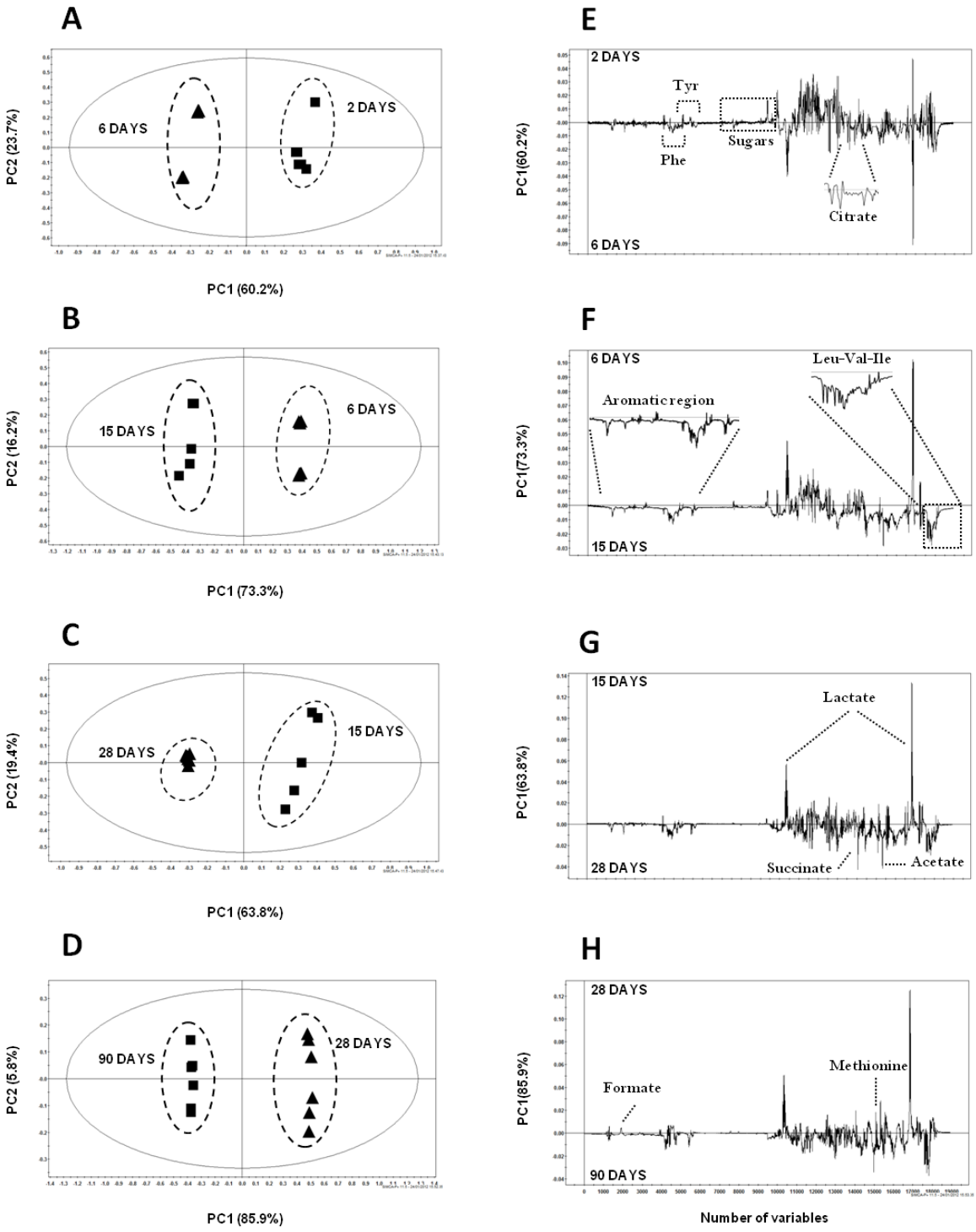

Figure 4.9 PCA scores (A-D) and loadings (E-H) plots derived from the ^{1}H NMR spectra of Fiore Sardo cheese manufactured with CC strain combination.

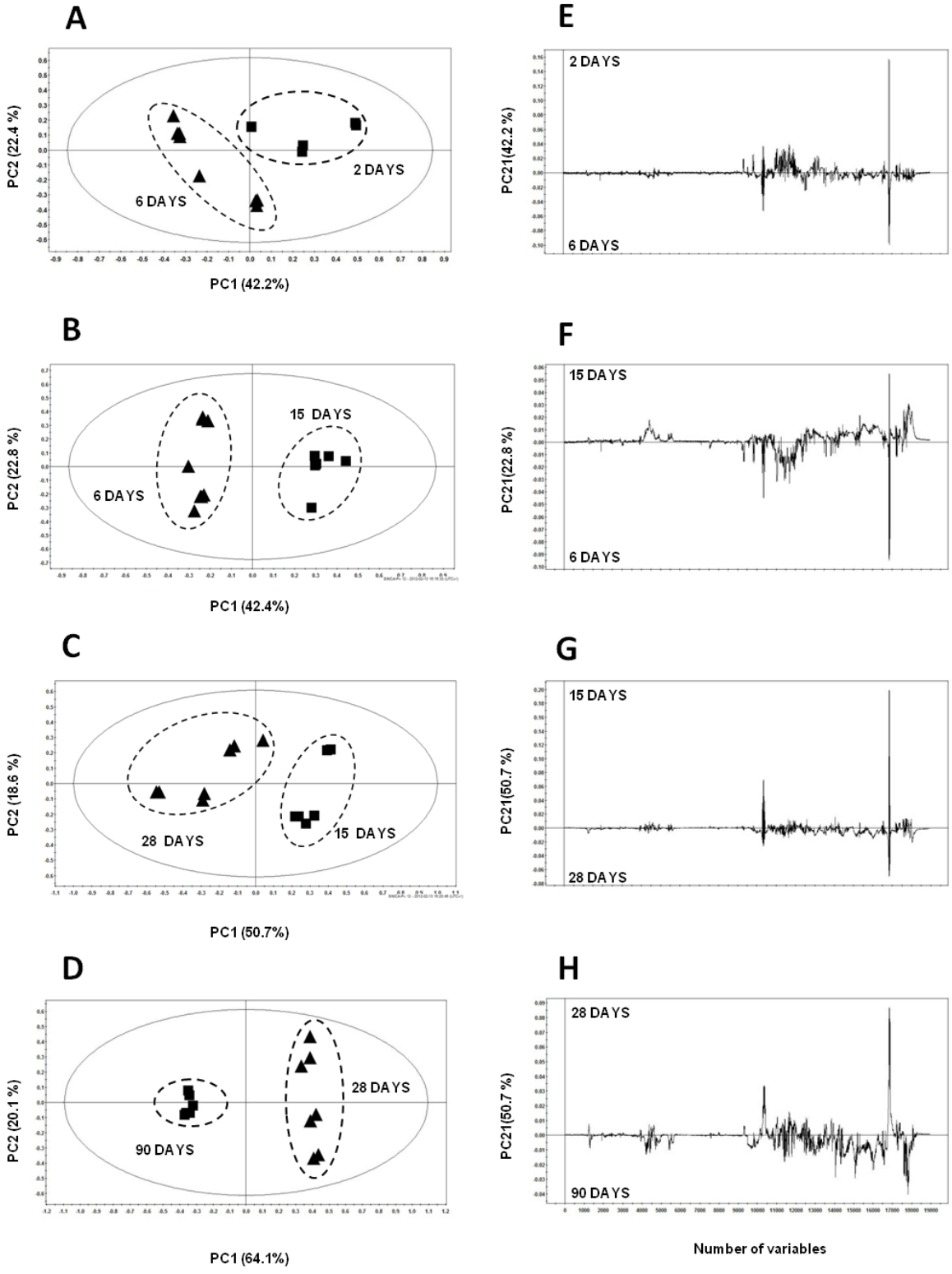

Figure 4.10 PCA scores (A-D) and loadings (E-H) plots derived from the ^{1}H NMR spectra of Fiore Sardo cheese manufactured with A1 strain combination.

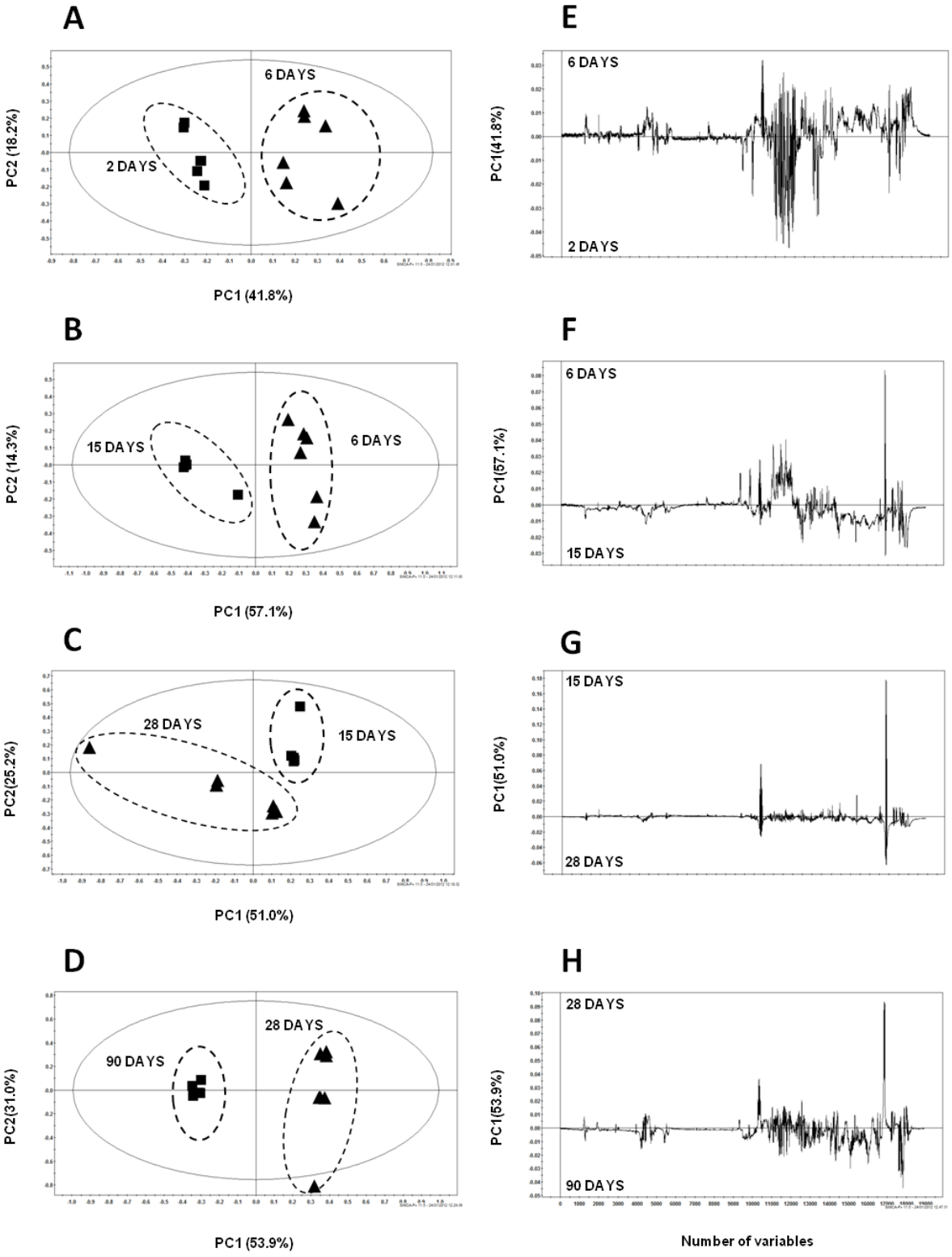

Figure 4.11 PCA scores (A-D) and loadings (E-H) plots derived from the ^{1}H NMR spectra of Fiore Sardo cheese manufactured with A2 strain combination.

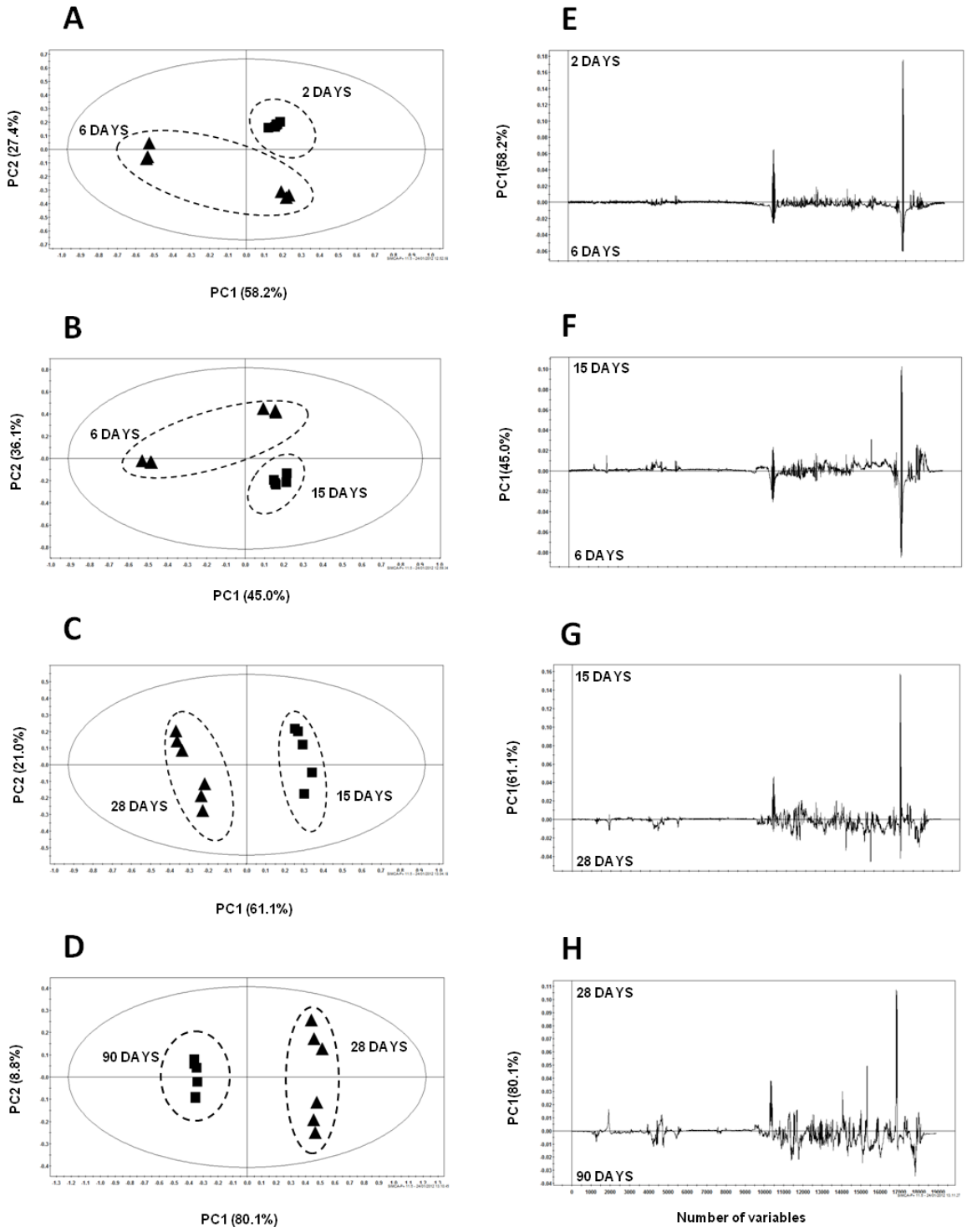

Figure 4.12 PCA scores (A-D) and loadings (E-H) plots derived from the ^{1}H NMR spectra of Fiore Sardo cheese manufactured with A3 strain combination

of other metabolites such as acetate, ethanol, diacetyl, and acetaldehyde (Cogan T.N. et al., 1993; Henriksen C.M. et al., 2001).

Organic acids. Organic acids are the major products of carbohydrate catabolism of lactic acid bacteria (Gonzalez-de Llano D. et al., 1996). Furthermore, they are important secondary carbon sources for many microorganisms and are intermediates and metabolites of a variety of biochemical processes (Bevilacqua A.E. et al., 1992; Lues J.F.R., 2000).

In all the experimental cheeses under investigation, the level of lactic acid increased in the initial stage of ripening (from 2 to 6 days), and, then, decreased. Since Lac is the most abundant organic acid detected in cheeses, the changes in its content are reasonably responsible of the observed modification in pH (see Table 4.2). In particular, the initial decrease of pH may be associated with the production of Lac from the lactose metabolism, while the following increase of pH can be related to the consumption of Lac by the nonstarter microbiota and/or by the indigenous cheese microbiota (Eliskases-Lechner F. et al., 1999; Bevilacqua A.E. et al., 1992). However, possible contributions to pH by the formation of basic compounds from proteolysis are not ruled out (Poullet B. et al., 1991).

The citric acid content progressively decreased during cheese aging in all batches. Citrate can be metabolized by citrate-positive strains of lactococci and by certain mesophilic lactobacilli in the NSLAB flora, such as strains of *Lb. casei* and *Lb. plantarum* , to produce pyruvate and acetate (Hugenholtz J. et al., 1993; Palles T. et al., 1998). It is interesting to note that the decrease of citric acid level was particularly marked in A1 and A2 cheeses, the former being inoculated with *Lb. paracei* and the latter with *Lb. plantarun.* It is therefore likely that the growth of facultative heterofermentative LAB is connected to their ability to use citrate as a source of energy.

Acetic acid (Ac) content increased as ripening progressed up to one month of maturation, then decreases. Ac is an important flavour compound in many cheeses. Besides being produced from lactose by heterofermentative LAB, it may also be formed as a result of citrate and lactate metabolism, or as a product of the catabolism of amino acids.

Pyruvic acid increased towards the end of the ripening period. This is a key compound in several metabolic pathways. It is formed both through the glycolytic pathway and through the degradation of the amino acids serina and threonine. Furthermore, pyruvate can also act as substrate of several metabolic reactions such as the formation of formic acid, ethanol, diacetyl, acetoin, and 2,3-butylene glycol (Marth E.H., 1974). The progressive increase of pyruvate level during ripening is a clear indication of its rapid production.

Amino acids. Proteolysis is the major process in cheese ripening; its degradation products, amino acids and peptides, have a considerable role on the texture, background flavour, and the availability of flavour precursors in all matured cheese varieties (Law B.A. et al., 1987; Fox P.F. et al., 1989).

The amino acids profile was similar in all of the Fiore Sardo trials: as the ripening time increases, a general increase of the amino acids level occurred, being particularly significant for leucine, isoleucine, valine, phenilalanine, and methionine. During cheese ripening, some individual free amino acids were further degraded. This is the case of the glutamic acid, tyrosine, and histamine.

Glu reached the highest content at 30 days of ripening, while in the following period decreased. The consumption of Glu was paralled by an increase in γ-amino butyric acid (GABA) level, as pointed out by TOCSY experiments performed on 90-day-old cheeses (Figure 4.13). The link between these two metabolites is the ability of some mesophilic lactobacilli to produce GABA by Glu decarboxylation. In the past, the presence of GABA in ripened cheeses was pointed out as an indication of anomalous fermentations that produce organoleptically deficient cheeses (Innocente N., 1997). More recently, the presence of GABA has been reviewed in a positive way in terms of functional properties of cheese, due to the known physiological functions of GABA such as neurotransmission, induction of hypotension, and diuretic and tranquilizer effects (Guin Ting Wong, et al, 2003).

Decarboxilation of tyrosine to tyramine and histidine to histamine took place at the end of ripening. Indeed, the presence of these BA is clearly visible in all 90-day-old cheeses under investigation. The production of biogenic amines is an extremely complex phenomenon, dependent on several chemical and biochemical variables, such as availability of free amino acids, temperature, pH, salt concentration, the

growth kinetics of the microorganisms, and their proteolytic and decarboxylase activities. Many studies have investigated the single effects of these factors, but little information exists on their combined or interactive effects. Therefore, finding a direct correlation between the BA contents and the total bacterial count in the cheeses is not an easy task (Halàsz A. et al., 1994).

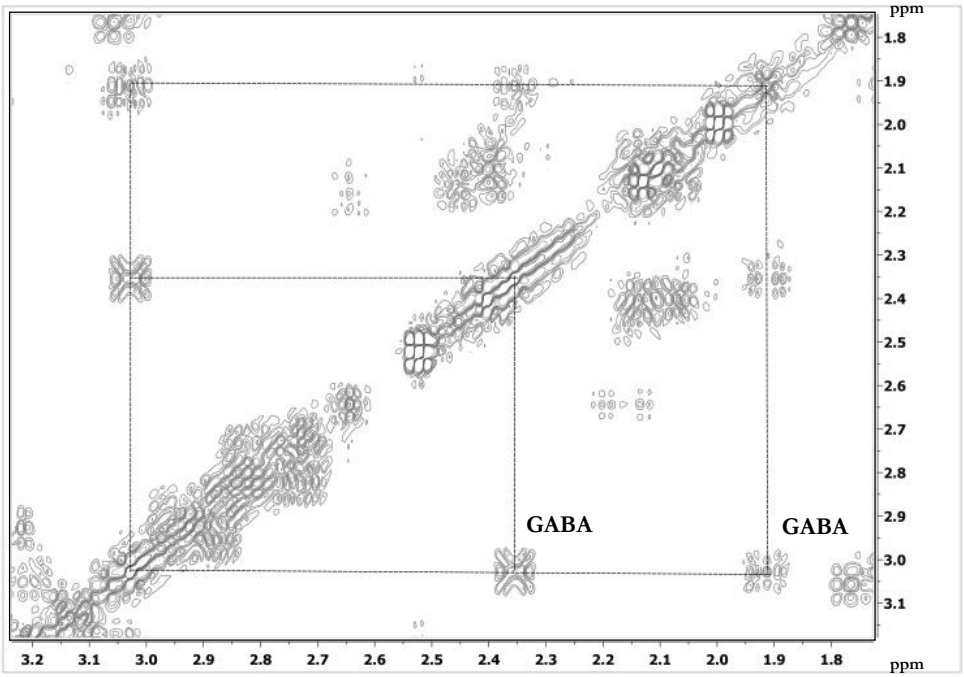

Figure 4.13 2D TOCSY expansion of the aqueous extract of Fiore Sardo at 90 days of ripening. The correlation peaks of GABA are indicated.

Osmolytes. With increasing ripening period, the level of Cho decreased. This behaviour, together with the parallel increase of betaine and glycine signals, is possibly due to the ability of some bacteria, including *E. coli*, to convert Cho into glycine-betaine and subsequently accumulated betaine during osmotic stress (Landfald B. et al., 1986). This hypothesis could explain the slower growth of glycine signal intensity in A1, A2, and A3 cheeses compared with the cheeses of type CC, the latter exhibiting the highest count in *E. coli*.

4.3.4 Impact of starter culture combination on the metabolic profile

To better compare the metabolic changes in cheeses in terms of the added strains, a PCA was applied on the 90-day-old cheeses (Figure 4.14).

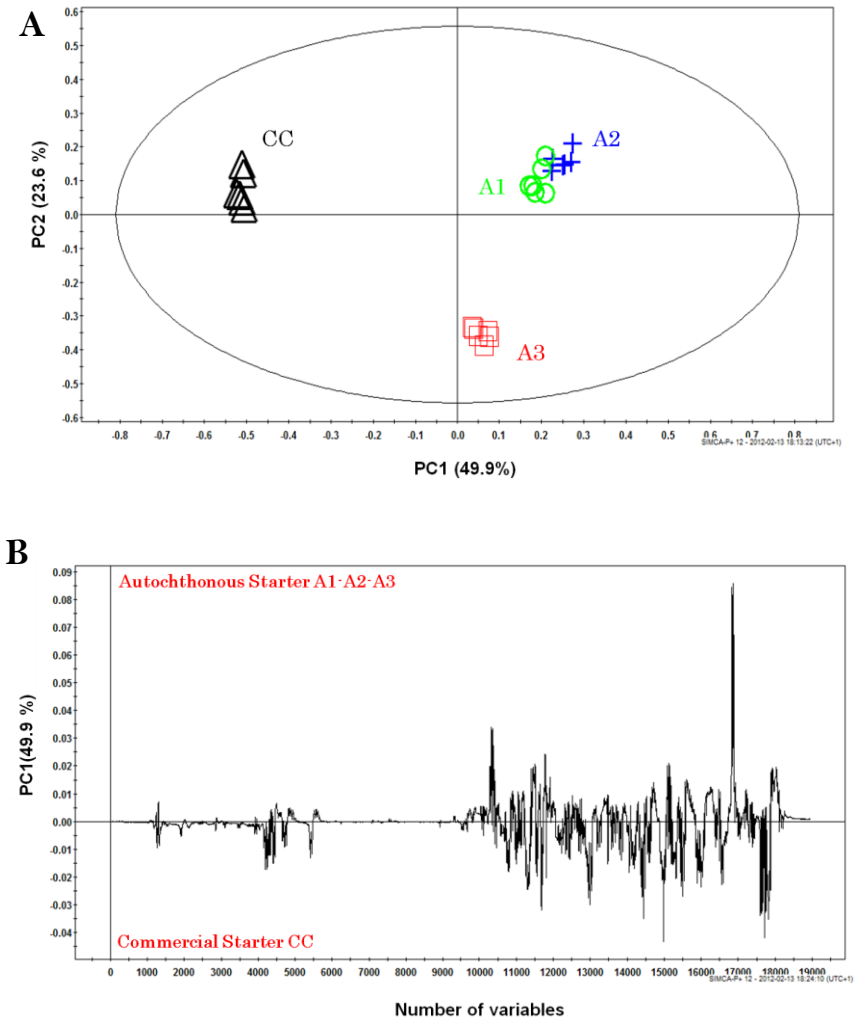

Figure 4.14 PCA scores (A) and loadings (B) plots derived from the ^{1}H NMR spectra of 90-day-old Fiore Sardo cheeses manufactured with CC (black), A1 (green), A2 (blue), and A3 (red) strain combinations.

The score plot displays a clear separation of cheeses manufactured with commercial starters from those made with the autochthonous cultures. In particular, the former cluster on the negative side of PC1 and the later on the positive side, suggesting that the first PC could be considered as representative of the origin of added strains ("autochthonous or not"). The loading plot shows that

this separation is mainly dominated by a decrease in lactic acid and an increase in amino acid and BA in CC chesees (Figure 14.14b). Futhermore, it is also worth noting that, comparing to A1, A2, and A3 cheeses, those made with commercial strains exhibited, starting from the first 48 h of ripening, overdeveloped irregularly shaped eyes and an early blowing, responsible of texture defects preserved until the end of ripening period (Figure 4.15). This kind of defects is typically due to infection of coli bacteria, which ferment the lactose and release large amounts of CO_2 and H_2 (Bachmann H.P. and Spahr U., 1995). Although this contamination is usual for raw milk cheeses, it is interesting to note that, in the cheeses under investigation, the counts of coli bacteria in CC was always higher than in A1, A2, and A3. Therefore, it seems reliable to link the observed structural defects of CC to its high content of these microorganisms.

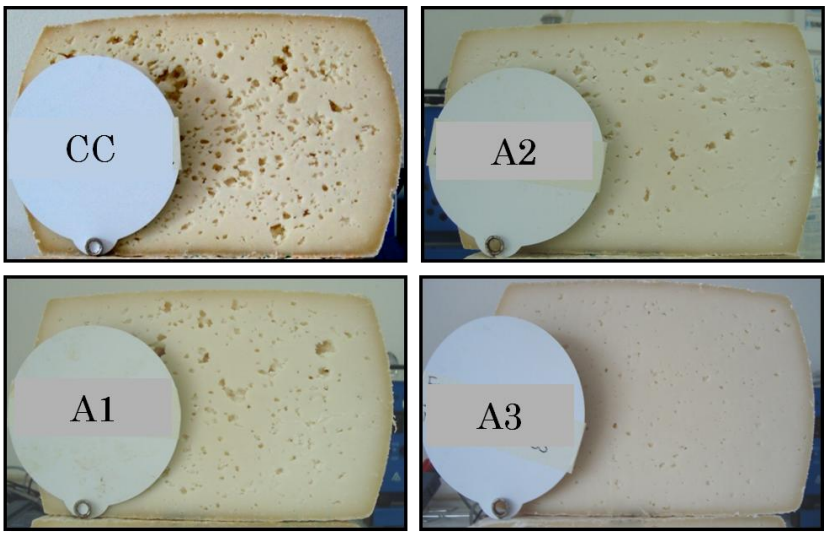

Figure 4.15 Pictures of Fiore Sardo at 90 days of ripening. Labels indicates the type of strain combination used in cheesemaking

As to the differences in the metabolic profiles of cheeses prepared with autochthnous strains, the PCA score plot in Figure 4.14 suggests the second PC as representative of the biological activity of strains (i.e LAB vs NSLAB). Indeed, the cheeses made with adjunct autochthonous LAB and NSLAB (i.e. A1 and A2) cluster near (on the positive side of PC2) and far from cheeses manufactured only with autochthonous LAB (i.e. A3). A better understanding of these differences was

achieved by performing a pair-wise PCA (Figure 4.16). On loading plots, the metabolites responsible for separation of A3 from A1 and A2 cheeses were increased level of val, leu, ile, tym, and phe, together with decreased level of form, cit, lac, and met. In addition, mainly form and phe were increased in A1 cheeses compared to those of type A2, while level of tyr, val, leu, ile and lac were higher in A2 samples.

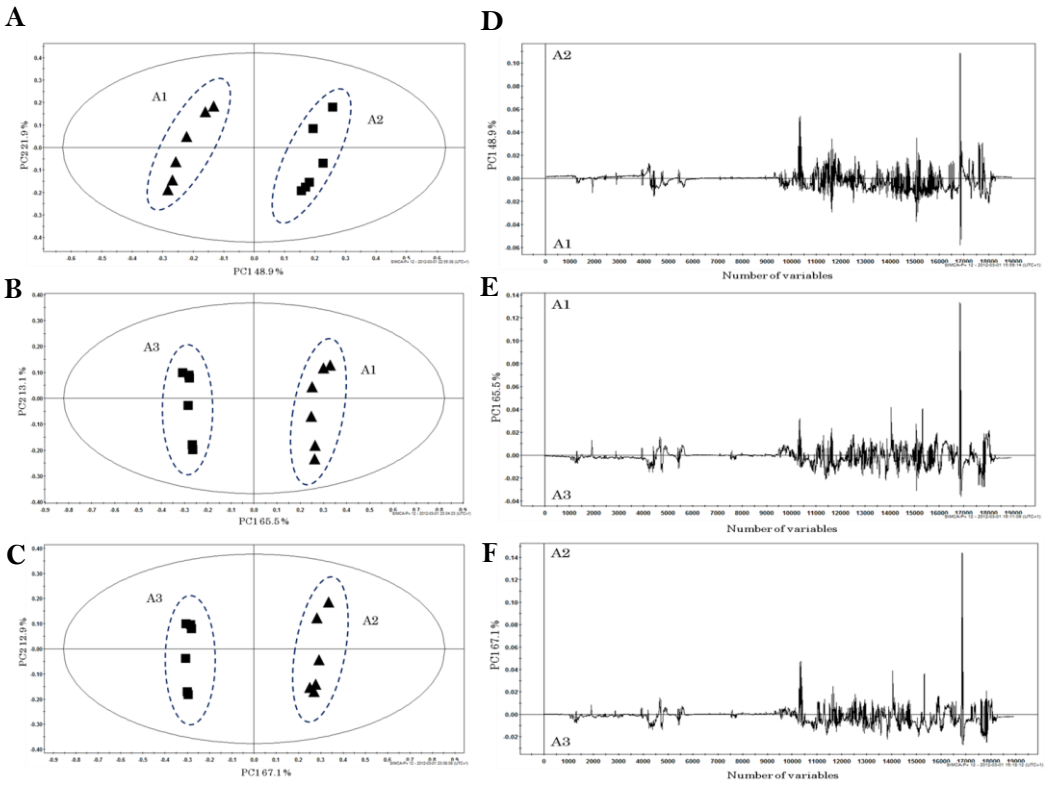

Figure 4.16 PCA scores and loadings plots derived from the ^{1}H NMR spectra of Fiore Sardo cheeses with 90 days of ripening as pair wise comparisons of cheeses with different strain combination: (A) PCA score plot showing separation between A1 (9) and A2 cheeses and (D) complementary loading plot of the first component (PC1); (B) PCA score plot showing separation between A3 (9) and A1 cheeses and (E) complementary loading plot of the PC1; (C) PCA score plot showing separation between cheeses with A3 and A2 strain combinations and (F) complementary loading plot of the PC1.

4.4 CONCLUSIONS

In the present study, the effect of adjunct autochthonous cultures on the Fiore Sardo metabolic profile was investigated by a ^{1}H NMR-based metabolomic approach. The ^{1}H NMR spectroscopy revealed to be an important tool for unbiased metabolite fingerprinting of cheese, while PCA highlighted genuine differences between strain combinations, affording clues on the nature of this differentiation. It was clear that, compared with commercial starters, blends prepared with autochthonous strains showed a higher level of lactic acid and a lower level of amino acids and, in particular, biogenic amines. Complemented with the monitored microbiological characteristics of cheeses, our results indicated a slower fermentation performance for the autochthonous strain combinations, but a more positive role with respect to the antagonist microflora and to the production of potential toxic compounds (i.e. BA).

The synergy between the NMR results and the microbiological analysis of cheeses demonstrated that the global analysis of metabolites by the combination of NMR spectroscopy and chemometrics can provided insights into cheese fermentation and the fermentative behaviours of strains. Thus, the finding of this work are promising in indicating ^{1}H NMR spectroscopy as a rapid and objectively sound technique for evaluation of the fermentative process in raw milk cheeses in any step of manufacturing. Undoubtedly, in order to better understand how to make good quality cheese, further studies are needed, complemented also by the evaluation of the influence of the native strains on the sensory characteristics of the final product.

References

Bachmann, H.P.; Spahr, U. The Fate of Potentially Pathogenic Bacteria in Swiss Hard and Semihard Cheeses Made from Raw Milk. *J. of Dairy Sci.* **1995**, *78*, 476-483.

Bevilacqua, A.E.; Califano, A.N. Changes in organic acids during ripening of Port Salut Argentino cheese. *Food Chem.* **1992**, *43*, 345–349.

Brescia, M.A.; Monfreda, M.; Buccolieri, A.; Carrino, C. Characterisation of the geographical origin of buffalo milk and mozzarella cheese by means of analytical and spectroscopic determinations. *Food Chem.* **2005**, *89*, 139-147.

Centeno, J.A.; Menéndez, S.; Rodríguez-Otero, J.L. Main microbial flora present as natural starters in Cebreiro raw cow's-milk cheese (Northwest Spain). *Int. J. Food Microbiol.* **1996**, *33*, 307–313.

Centeno, J.A.; Menendez, S.; Hermida, M.A.; Rodríguez-Otero, J.L. Effects of the addition of *Enterococcus faecalis* in Cebreiro cheese manufacture. *Int. J. Food Microbiol.* **1999**, *48*, 97–111.

Cogan, T.N.; Hill, C. Cheese starter cultures. *Cheese: chemistry, physics and microbiology* **1993**, *volume1*.

Consonni, R.; Cagliani, L.R. Ripening and geographical characterization of Parmigiano Reggiano cheese by 1H NMR spectroscopy. *Talanta* **2008**, *76*, 205-205.

Coppola, S.; Parente, E.; Dumontet, S.; La Peccerella, A. The microflora of natural whey cultures utilized as starters in the manufacture of Mozzarella cheese from water-buffalo milk *Lait dairy J.* **1988**, *68*, 295–309.

Dasen, A.; Berthier, F.; Grappin, R.; Williams, A.G.; Banks, J. Genotypic and phenotypic characterization of the dynamics of the lactic acid bacterial

population of adjunct-containing Cheddar cheese manufactured from raw and microfiltered pasteurised milk. *J. Appl. Microbiol.* **2003**, *94*, 595–607.

De Angelis, M.; Corsetti, A.; Tosti, N.; Rossi, J.; Corbo, M.R.; Gobbetti, M. Characterization of Non-Starter Lactic Acid Bacteria from Italian ewe cheeses based on phenotypic, genotypic, and cell wall protein analyses. *Appl. Environ. Microbiol.* **2001**, *67*, 2011-2020.

Desmazeaud, M.; Piard, J.C. Inhibiting factors produced by lactic acid bacteria. Bacteriocins and other antibacterial substances. *Lait* **1992**, *72*, 113-142.

Eliskases-Lechner, F.; Ginzinger, W.; Rohm, H.; Tschager, E. Raw milk flora affects composition and quality of Bergkäse. 1. Microbiology and fermentation compounds. *Le Lait* **1999**, *79*, 385 – 396.

El-Soda, M.A. The role of lactic acid bacteria in accelerated cheese ripening. *FEMS Microbial. Rev.* **1993**, *12* , 239-251.

Folch, J.; Lees, M.; Sloane-Stanley, G.H. A simple method for the isolation and purification of total lipids from animal tissues. *J. Biol. Chem* **1957**, *226*, 497-509.

Fox, P.F. Proteolysis during cheese manufacture and ripening. *J. Dairy Sci.* **1989**, *72*, 1379–1400.

Fox, P.F.; Lucey, J.A.; Cogan T.M. Glycolysis and related reactions during cheese manufacture and ripening. *Crit. Rev. Food Sci. Nutr.* **1990**, *29*, 237-253.

Fox, P.F.; Law, J.; McSweeney, P.L.H.; Wallace, J. Biochemistry of cheese ripening. In: Fox, P.F. (Ed.), Cheese: Chemistry, Physics and Microbiology, 2nd edn. General Aspects, Chapman & Hall, London, **1993**, *vol. 1*., pp. 415–417.

Garcia Fontán M.C.; Franco, I.; Prieto, B.; Tornadijoand, M.E.; Carballo, J. Microbiological changes in San Simon cheese throughout ripening and its relationship with physico-chemical parameters. *Food Microbiol.*, **2001**, *18*, 25-33.

Gianferri, R.; Maioli, M.; Delfini, M.; Brosio, E. A low-resolution and high-resolution nuclear magnetic resonance integrated approach to investigate the physical structure and metabolic profile of Mozzarella di Bufala Campana cheese. *Int. Dairy J.* **2007**, *17*, 167–176.

González de Llano, D.; Rodriguez, A.; Cuesta, P. Effect of lactic starter cultures on the organic acid composition of milk and cheese during ripening—analysis by HPLC. *J. Appl. Microbiol.* **1996**, *80*, 276-570.

Guin Ting Wong, G., T. Bottiglieri, and O. Carter Snead III. GABA, γ-hydroxybutyric acid, and neurological disease. *Ann. Neurol.* **2003**, ***6,*** 3–12.

Halász, A.; Baráth, A.; Simon-Sarkadi, L.; Holzapfel, W. Biogenic amines and their production by microorganisms in food. *Trends Food Sci. Technol.* **1994,** *5*, 42-49.

Hatzikamari, M.; Litopoulou-Tzanetaki, E.; Tzanetakis, N. Microbiological characteristics of Anevato: a traditional Greek cheese. *J. Appl. Microbiol.* **1999**, *87*, 595–601.

Henriksen, C.M.; Nilsson, D. Redirection of pyruvate catabolism in Lactococcus lactis by selection of mutants with additional growth requirements. *Appl. Microbiol. Biotechnol.* **2001**, *56*, 767-775.

Hugenholtz, J. Citrate metabolism in lactic acid bacteria. *FEMS Microbiol. Rev.* **1993,** *12*, 165-178.

Innocente, N. Free amino acids and water-soluble nitrogen as ripening indices in Montasio cheese. *Lait* **1997**, *77*, 359-369.

Joosten, H.; Northolt, M.D. Conditions allowing the formation of biogenic amines in cheese, 2: decarboxylative properties of some non-starter bacteria. *Neth. Milk Dairy J.* **1987**, *41*, 259-280.

Joosten, H.; Northolt, M.D. Detection, growth, and amine-producing capacity of lactobacilli in cheese. *Appl. Environ. Microbiol.* **1989**, *55*, 2356-2359.

Landfald, B.; Strùm, A.R, Choline-glycine betaine pathway confers a high level of osmotic tolerance in Escherichia coli. *J. Bacteriol* **1986**, *165*, 849-855.

Law, B.A Proteolysis in relation to normal and accelerated cheese ripening. In P. F. Fox (Ed.), Cheese: Chemistry, physics and microbiology. *London: Elsevier Applied Science* **1987**, *vol. 1*., pp. 365–392.

Ledda, A.; Scintu, M.F.; Pirisi, A.; Sanna, S.; Mannu, L. 1994 Technological characterization of lactococci and enterococci of Fiore Sardo sheep cheese. *Scienza e Tecnica Lattiero-casearia* **1994**, *45*, 443–456.

Leroy, F.; De Vuyst, L. Lactic acid bacteria as functional starter cultures for the food fermentation industry. *Trends Food Sci. Technol.* **2004**, *15*, 67–78.

Litopoulou-Tzanetaki, E. Changes in numbers and kinds of lactic acid bacteria during ripening of Kefalotyri cheese. *J. Food Sci.* **1990**, *55*, 111-113.

Lopez-Diaz, T.M.; Santos, J.A.; Gonzalez, C.J.; Moreno, B.; Garcia, M.L. Bacteriological quality of traditional Spanish blue cheese. *Milk Sci. Int.* **1995**, *50*, 503–505.

Lues, J.F.R. Organic acid and residual sugar variation in a South African Cheddar cheese and possible relationships with uniformity. *J. Food Compos. Anal.* **2000**, *13*, 819–825.

Macedo, A.C.; Malcata, F.X.; Hogg, T.A. Microbiological profile in Serra ewe's cheese during ripening. *J. Appl. Microbiol.* **1995**, *79*, 1–11.

Macedo, A.C.; Tavares, T.G.; Malcata, F.X. Influence of native lactic acid bacteria on the microbiological, biochemical and sensory profiles of Serra da Estrela cheese. *Food Microbiol.* **2004**, *21*, 233-240.

Mannu, L.; Comunian, R.; Scintu, M.F. Mesophilic lactobacilli in Fiore Sardo cheese: PCR-identification and evolution during cheese ripening. *Int. Dairy J.* **2000**, *10*, 383–389.

Marilley, L.; Casey, M.G. Flavours of cheese products: metabolic pathways, analytical tools and identification of producing strains. *Int. J. Food Microbiol.* **2004**, *90*, 139–159.

Marth, E.H. Fermentaions. In: Webb, B.H.; Johnson, A.H.; Alford, J.A. (Eds.), *Fundamentals of dairy chemistry.* **1974** 772–872 Westport, CT: The Avi Publishing Company.

Medina, M.; Del Pozo, B.F.; M. Rodríguez-Marin, A.; Gaya, P.; Nuñez, M. Effect of lactic starter inoculation on chemical, microbiological, rheological and sensory characteristics of La Serena cheese. *J. Dairy Res.* **1991**, *58*, 355-361.

Medina, M.; Gaya, P.; Nũnez, M. Gredos goat's milkcheese microbiological and chemical changes throughout ripening. *J. Dairy Res.* **1992**, *59*, 563–566.

Nikolaou, E.; Tzanetakis, N.; Litopoulou-Tzanetaki, E.; Robinson R.K. Changes in the microbiological and chemical characteristics of an artisanal, low-fat cheese made from raw ovine milk during ripening. *Int. J. Dairy Technol.* **2002**, *55*, 12–17.

Nũnez, M.; Gaya, P.; Medina, M. Influence of manufacturing and ripening conditions on the survival of Enterobacteriaceae in Manchego cheese. *J. Dairy Sci.* **1985**, *68*, 794–800.

Ordóñez, J.A.; Barneto, R.; Mármol, M.P. Microbial study of Casar de Cáceres cheese throughout ripening. *J. Dairy Res.* 1991, *58*, 231–238.

Palles, T.; Beresford, T.; Condon, S.; Cogan, S. Citrate metabolism in Lactobacillus casei and Lactobacillus plantarum. *J. Appl. Microbiol.* **1998**, *85*, 147–154.

Pinho, O.; Ferreira, I.M.P.L.V.O.; Mendes, E.; Oliveira, B.M.; Ferreira, M. Effect of temperature on evolution of free amino acid and biogenic amine contents during storage of Azeitão cheese. *Food Chem.* **2001**, *75*, 287–291.

Pisano, M.B.; Fadda, M.E.; Deplano, M.; Corda, A.; Cosentino, S., Microbiological and chemical characterization of Fiore Sardo, a traditonal Sardinian cheese made from ewe's milk. I*nt. J. Dairy Technol.* **2006**, *59*, 171-179.

Pisano, M.B.; Fadda, M.E.; Deplano, M.; Corda, A.; Casula, M.; Cosentino, S., Characterization of Fiore Sardo cheese manufactured with the addition of autochthonous cultures. *J. Dairy Res.* **2007**, *74*, 255–261.

Psoni, L.; Tzanetakis, N.; Litopoulou-Tzanetaki, E. Microbiological characteristics of Batzos, a traditional Greekcheese from raw goat's milk. *Food Microb.* **2003**, *20*, 575–582.

Poullet, B.; Huertas, M.; Sánchez, A.; Cáceres, P.; Larriba, G. Microbial study of Casar de Cáceres cheese throughout ripening. *J. Dairy Res.* **1991**, *58*, 231-238.

Rodrigues, D.; Santos, C.H.; Rocha-Santos, T.A.P.; Gomes, A.M.; Goodfellow, B.J.; Freitas, A.C. Metabolic profiling of potential probiotic or synbiotic cheeses by Nuclear Magnetic Resonance (NMR) Spectroscopy. *J. Agric. Food Chem.* **2011**, *59*, 4955–4961.

Rodriguez Medina, M.L.; Tornadijo, M.E.; Carballo, J.; Marten Sarmiento, R. Microbiological study of Leòn raw cow-milk cheese, a Spanish craft variety. *J. Food Prot.* **1995**, *57*, 998-1006.

Savorani, F.; Tomasi, G.; Engelsen, S.B. icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. *J. Magn. Reson.* **2010,** *202*, 190-202.

Suzzi, G.; Caruso, M.; Gardini, F.; Lombardi, A.; Vannini, L.; Guerzoni, M.E.; Andrighetto, C.; Lanorte, M.T. A survey of the enterococci isolated from an artisanal Italian goat's cheese (semicotto caprino). *J. Appl. Microbiol.* **2000**, *89*, 267–274.

Thompson, T.L.; Marth, E.H. Changes in Parmesan cheese during ripening: Microflora - coliforms, enterococci, anaerobes, propionibacteria and staphylococci. *Milchwissenschaft* **1986**, *41*, 201-205.

Zarate, V.; Belda, F.; Perez, C.; Cardell, E. Changes in the microbial flora of Tenerife goat's milkcheese during ripening. *Int. Dairy J.* **1997**, *7*, 635–641.

5. NMR metabolic profiling of the organic and aqueous extracts of *Argentina sphyraena* (Osteichthyes: Argentinidae)

5.1 INTRODUCTION

Europe's fishing grounds were once among the most productive in the world, but forty years of the Common Fisheries Policy (CFP)[4] have resulted in serious depletion of fish populations, ecosystem degradation and damage to species, habitats and sites protected by EU environmental legislation (EC, 2009). Fishing has become unsustainable, increasingly unprofitable and reliant on public subsidies. This in turn has led to deprivation in coastal communities and an ever growing reliance on imported fish.

OCEAN2012[5], a coalition of environmental organizations and associations of small-scale fishermen dedicated to stop overfishing and to bringing an end to the practice of destructive fishing, together with organizations such as Greenpeace and Living Seas, are mobilized, also at Community level, to steer the fisheries reform towards a policy that addresses the recovery of fish stock as a priority. As part of a broader, stepwise approach to returning EU fisheries to a sustainable footing, OCEAN2012 suggested replacing, or at least enhancing, relative stability with a system of allocating access to fisheries based on an explicit consideration of certain criteria. The allocation system should contribute a more equitable distribution of access to available fishing resources and a culture of compliance. In the interests of creating a principle-centered approach to fisheries management in EU waters and for the EU fleet globally, OCEAN2012 is investigating issues that may be incorporated into a reformed CFP. Some of these issues relate to giving highest priority to environmental objectives under the CFP, creating a framework to ensure decisions are made at appropriate levels, defining instruments that deliver

[4] The Common Fisheries Policy ("CFP") is the European Union's ("EU") instrument for the management of fisheries and aquaculture. European Comm'n, Fisheries, About the Common Fisheries Policy, http://ec.europa.eu/fisheries/cfp_en.htm (last visited Feb. 5, 2010).

[5] http://www.ocean2012.eu

sustainable fishing capacity, basing access to fishing on criteria that ensure a transition to, and support for, environmentally and socially sustainable fishing.

Among the proposed criteria in support of a long-term sustainability and a minimization of the adverse ecological and environmental impact of fishing activities, there is also the reduction of unwanted by-catches and progressive elimination of discards. Basically, the discarding process artificially splits the fish community into two fractions: the commercial/marketed, and the discarded/non-marketed. The incidental capture of species toward which there is no directed effort is characteristic of commercial fisheries and is termed `by-catch'. Among different fishing gears, the trawl is responsible for most fisheries discards (Stergiou K.I. et al., 1998; Hall S.J., 1999). In the Mediterranean Sea, from a total of 300 species caught, only c.a. 10% are consistently marketed and c.a. 30% are occasionally retained, depending on the sizes caught and market demands, whereas 60% are always discarded (Carbonell A. et al., 1998; Stergiou K.I. et al., 1998; Machias A. et al., 2001; D'Onghia G. et al., 2003).

In Italy, the catch of trawl fishery is highly represented by small pelagic species with low economic value, scarcely known by consumers and poorly appreciated and commercialized. Among these, *Argentina sphyraena* is an important fraction of the by-catch of the deep sea trawl fishery in western Mediterranean. In spite of its abundance, in Italy argentina has a low economic interest, finding no appreciation on the market. In fact, especially in urbanized areas, the global market trends affect mostly the consumers' preference thus privileging those species with a higher commercial value present on the market like hake (Merluccius merluccius) from fishery, or gilthead sea bream (Sparus aurata) and sea bass (Dicentrarchus labrax) from aquaculture.

The studies on fish metabolomics found in the literature cover a wide range of topics: fish physiology and development, pollutant effects on fish, fish condition and disease, and fish as foodstuffs (Samuelsson, L.M. and Larsson D.G.J., 2008). Among these, those using NMR spectroscopy in combination with pattern recognition techniques have been focused mainly to well known fish species, for the elucidation of the origin and adulteration of foodstuffs (Martinez I. et al., 2005; Aursand M. et al., 2009; Standal I.B., 2009; Savorani F. et al., 2010).

Besides being little known to the consumers, *Argentina sphyraena* is also little know in marine biology. Thus, starting from collaboration with the Department of Life Sciences and Environment, Section of Animal Biology and Ecology of the University of Cagliari, in the present study, ^{1}H NMR spectroscopy was combined with multivariate statistical analysis to investigate seasonal variations of the aqueous and lipid profile of *Argentina sphyraena*. The goals of the study were to: (i) use NMR to metabolically profile this fish for future studies, (ii) ascertain how the metabolic profiles differ according to the fishing season and (iii) provide new insights on the potential of NMR-based metabolomics as a rapid and informative screening tool in fish research.

5.2 MATERIAL AND METHODS

5.2.1 Chemicals

Deuterium oxide (D_2O, 99.9%) was purchased from Cambridge Isotope Laboratories Inc. (Andover, MA). Sodium 3-trimethylsilylpropionate- 2,2,3,3,-d4 (TSP, 98 atom % D), perchloric acid ($HClO_4$, 70%), and potassium hydroxide (KOH) and some standard amino acids, organic acids and nucleobases were acquired from Sigma-Aldrich (Milan, Italy).

5.2.2 Samples

50 individuals of *Argentina sphyraena* were caught from December to April between 2010 and 2011 in the waters around the island of Sardinia at a depth of 150m during experimental campaigns of trawling and commercial fishing (Table 5.1). Fishes were frozen at -20 °C and transported by boat to our laboratory. The fish were thawed and gutted. Muscle was chopped and mixed, to get a representative sample, and further stored at -80 °C before extracting the water-soluble component.

5.2.3 The lipid fraction extract

Lipid extraction was performed according to a modified Bligh and Dyer (Martinez I. et al., 2009) procedure: 10 g of fish muscle were homogenized for 2 min with a

Table 5.1 Fishing season, area and depth of capture, number of individuals caught of the samples studied sphyraena Argentina.

SEASONS	FISHING AREA/ DEPTH	N° SAMPLES
DECEMBER	TEULADA 150M	13
JANUARY	TEULADA 150M	13
FEBRUARY	TEULADA 150M	6
APRIL	TEULADA 150M	18

mixture of 16 ml of H_2O, 40 ml of methanol and 20 ml of chloroform. Then, 20 ml of chloroform were added to the mixture and homogenized for 40 s, prior the addition of 20 ml of H_2O and a further homogenization for 40 s. The homogenate was centrifuged for 10 min at 4000 rpm and the chloroform phase containing the lipids was recuperated for NMR analyses. The lipid phase was transferred into a round-bottomed flask and dried with a rotary vacuum evaporator. The lipids were then gravimetrically determined by weighting the flask before and after evaporation and converted to percent lipids. The extract was redissolved in 800 µL of $CDCl_3$ and transferred into a 5 mm NMR tube.

5.2.4 The water-soluble extract

Water-soluble metabolites were extracted using perchloric acid on the basis of the procedure previously described by Gribbestad et al. (Gribbestad I.S. et. al., 2005). This has prevented the enzymatic degradation of the muscle during the extraction and allowed to remove the acidic proteins and macromolecules. On portion of ca. 1g of muscle was accurately sampled from each fish adding 2 ml of $HClO_4$ (7% in D_2O). Subsequently, the sample was homogenized for about 20 seconds with the homogenizer *Ultra-Turrax T25 basic* and continuously stirred and warmed at 50 °C until a paste consistency was obtained. The homogenate was centrifuged at 4000 rpm for 10 min at 4 °C. Then, the supernatant was adjusted to pH 7.8 with 9M KOH in D_2O and centrifuged again to remove the potassium

perchlorate. The final extract was lyophilized and stored at -20 °C until analyzed. The solutions were thawed just before NMR analysis.

For NMR analysis, each sample was redissolved in 1 mL of D_2O, and an aliquot of 800 µL was transferred into a 5 mm tube to which 66.6 µl of TSP/D_2O solution was added (0.78 mM final concentration of TSP) as a frequency standard. The solution pH was accurately adjusted to 7.80 and, subsequently, the sample was centrifuged at 10000 rpm for 5 min at 4°C in order to remove precipitates.

5.2.5 ^{1}H NMR spectroscopy

^{1}H NMR experiments on aqueous extract were carried out on a Varian Unity 500 spectrometer operating at 499.83 MHz. Spectra were recorded at 300 K with a spectral width of 5624 HZ, a 45° pulse of 7.5 µs, an acquisition time of 3s, a relaxation delay of 7s, and 256 scans. The residual water signal was suppressed by applying a presaturation technique with low-power radiofrequency irradiation for 1.5 s. The FIDs were multiplied by an exponential weighting function equivalent to a line broadening of 0.3 Hz prior to Fourier transformation. Chemical shifts were referred to the TSP single resonance at 0.00 ppm. 2D NMR ^{1}H-^{1}H COSY spectra were acquired with a spectral width of 53 KHz in both dimensions, 512 increments in f_1 and 2K data points in f_2 with 96 transients. 2D NMR ^{1}H-^{1}H TOCSY spectra were acquired in phase sensitive mode with a size and number of data points similar to those of the COSY and a mixing time of 80 ms.

The ^{1}H NMR spectra of the lipid fraction were recorded at 298 K on a Varian UNITY INOVA 400 spectrometer operating at 399.95 MHz. The ^{1}H spectra of the $CDCl_3$ extracts were acquired by co-adding 40 transients, a 45° pulse of 3.4 µs, a recycle delay of 16.5 s and the chemical shifts were referenced the proton signal of the solvent (7.26 ppm).

5.2.6 Pre-processing of NMR spectra

Spectra were processed using ACD Labs 1D Spectrum Manager, version 12.0. All spectra were zero filled to 64k points and multiplied by an exponential factor of 0.3 Hz. Each FID was then Fourier transformed, phased using the simple method, baseline corrected using a 4th order polynomial and referenced to the appropriate

reference signal. Prior to integration, the spectral region of the water aqueous extracts containing the residual HOD (4.65-4.85 ppm), and those at both edges, containing only noise (between 10.00 and 9.00 ppm and between 0.50 and -0.50 ppm) were removed.

The intelligent bucketing method developed by ACD/Labs was used to integrate the spectra. A bucket width of 0.04 ppm was selected with the intelligent bucket looseness set to 50%. This resulted in bucket widths that ranged between 0.02 and 0.06 ppm. Subsequently, the matrix was normalized to minimize the differences due to the sample extract dilutions.

5.2.7 Chemometric analysis of the data

Multivariate data analysis was carried out using SIMCA-P+ 12.0 (Umetrics, Umea, Sweden). Data were Pareto scaled before analysis. Pareto scaling is a compromise between mean centering, which may fail to pick out small changes in metabolite concentrations, and scaling to unit variance, which gives equal weight to baseline imperfections, noise, and defined signals in the NMR spectrum. Principal components analysis (PCA) and OPLS-DA were used to examine inherent clustering and correlations within the data.

5.3 RESULTS AND DISCUSSION

In the first part of this section, some preliminary information on the physiological state of *Argentina sphyraena*, provided by the Section of Animal Biology and Ecology of the University of Cagliari, are reported, together with the mean muscle lipid content of the fish. In the second part, the ^{1}H NMR spectra of the water soluble and lipid extracts from muscle are shown, while the multivariate statistical analysis results are discussed in the third part.

5.3.1 Physiological data and muscle lipid content

Since the level of lipids and low molecular weight metabolites in fish muscle can be affected by both the sexual maturation and the diet of fish (Ozyurt G. and Polat A., 2006), preliminary data on the maturity stage and the feeding activity of individuals under investigation were collected.

A macroscopic observation carried out on ovary indicated that fishes caught in December were in a *post-spawning* period, those caught in January were in the *pre-spawning* state, while those of April were in the *spawning* period. Furthermore, the results for the hepatosomatic index (HSI), shown in Figure 5.1, revealed that the fish feeding activity increased from December to April, that is from the immature to the mature stage. This result indicates that *Argentina sphyraena* feed also during the spawning period, differently from other species of fish that stop feeding in this stage (Solansky K.S. et al., 2005).

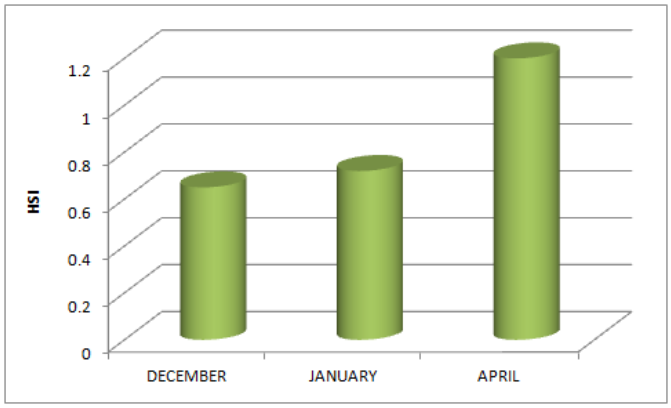

Figure 5.1 HSI measured as the liver weight to body weight.

The above information was complemented by the estimation of the lipid content in fish muscle by the standard macro-gravimetric method of Bligh–Dyer (Martinez I. et al., 2009). The data, reported in Figure 5.2, shows that the the mean fat content in April (1.1%) was 56% less than the December value (2.48%), pointing out a progressive decrease of the fat level from the immature to the mature stage. This result, together with the above mentioned physiological data, suggests that, during the *pre-spawning* period, *Argentina sphyraena* accumulate and storage fat in the muscle that are likely used as fuel during the gonadal development in the maturity stage (Hendry A.P. and Berg O.K., 2000).

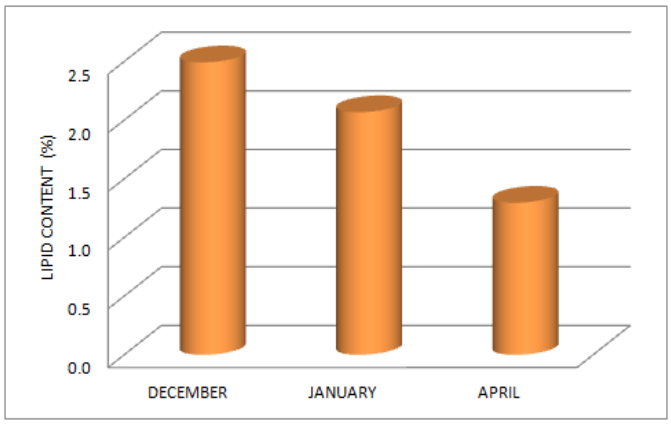

Figure 5.2 Average lipid content in argentina fish muscle estimated by the standard macro-gravimetric method of Bligh–Dyer (1959).

5.3.2 ^{1}H NMR spectrum of the aqueous extract of *Argentina sphyraena* muscle

Figure 5.3 shows a typical ^{1}H NMR spectrum of the aqueous extract of muscle of *Argentina sphyraena*, while in Figure 5.4 some spectral regions are shown in more detail. The main assignments are reported in Table 5.1

The compounds were identified on the basis of data published in the literature (Mannina L. et al., 2008; Standal I.B. et al., 2007; Fann T., 1996), by performing 2D NMR experiments (COSY and TOCSY), and by recording spectra of standard compounds. In some cases, validation of the peak attribution was achieved by adding standard compounds directly to the sample solution and recording the NMR spectrum again under the same conditions. Finally, for a limited number of cases, the software Chenomx NMR Suite (version 7.1), a powerful platform allowing easy identification and quantification of metabolites in NMR spectra or processing NMR spectra, was used. A total of 41 molecular compounds were identified which can be distinguished in the following classes: carbohydrates, organic acids, amino acids, osmolytes, and derivatives of nucleosides.

Some organic acids such as fumaric (Fu), malic (Mal), formic (Form), succinic (Suc), acetic (Ace), citric (Cit), and lactic acid (Lac) were identified. Among these, the most abundant was lactic acid (1.36 and 4.23 ppm) whose signal at 1.36 ppm is a partly overlaps with that of threonine at 1.33 ppm. Threonine presence was

highlighted by the cross speak signals at 3.58 and 4.27 ppm in the TOCSY spectrum (Figure 5.5).

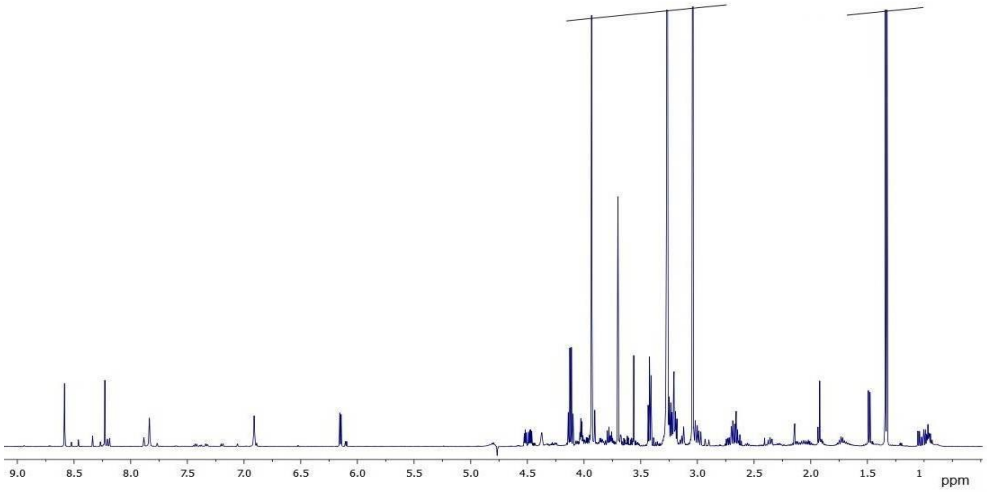

Figure 5.3 ^{1}H NMR spectrum of the aqueous extract of Argentina sphyrenae muscle

A small amount of glucose was present in the muscle aqueous extract. It was identified by means of the diagnostic anomeric doublets at 4.66 ppm (β-form) and 5.24 ppm (α-form) and by 2D NMR experiments.

The characteristic singlets at 3.04 and 3.93 ppm are indicative of the presence of creatine and/or phosphocreatine. Both compounds represent an important energy store in skeletal muscle (Walliman T., 2007). Phosphocreatine is used to anaerobically generate ATP from ADP, forming creatine.

The signals at 6.91 and 7.83 ppm were attributed to two dipeptides with antioxidant activity: carnosine, N(α)-(β-alanine)-L-1-histidine, and its methylated analogue, anserine, N(α)-(β-alanine)-L-1-methylhistidine.) (Boldyrev A.A. et al., 1990). Moreover, other amino acids, such as arginine (Arg), aspartate (Asp), phenylalanine (Phe), glycine (Gly), glutamate (Glu), histidine (His), isoleucine (Ile),

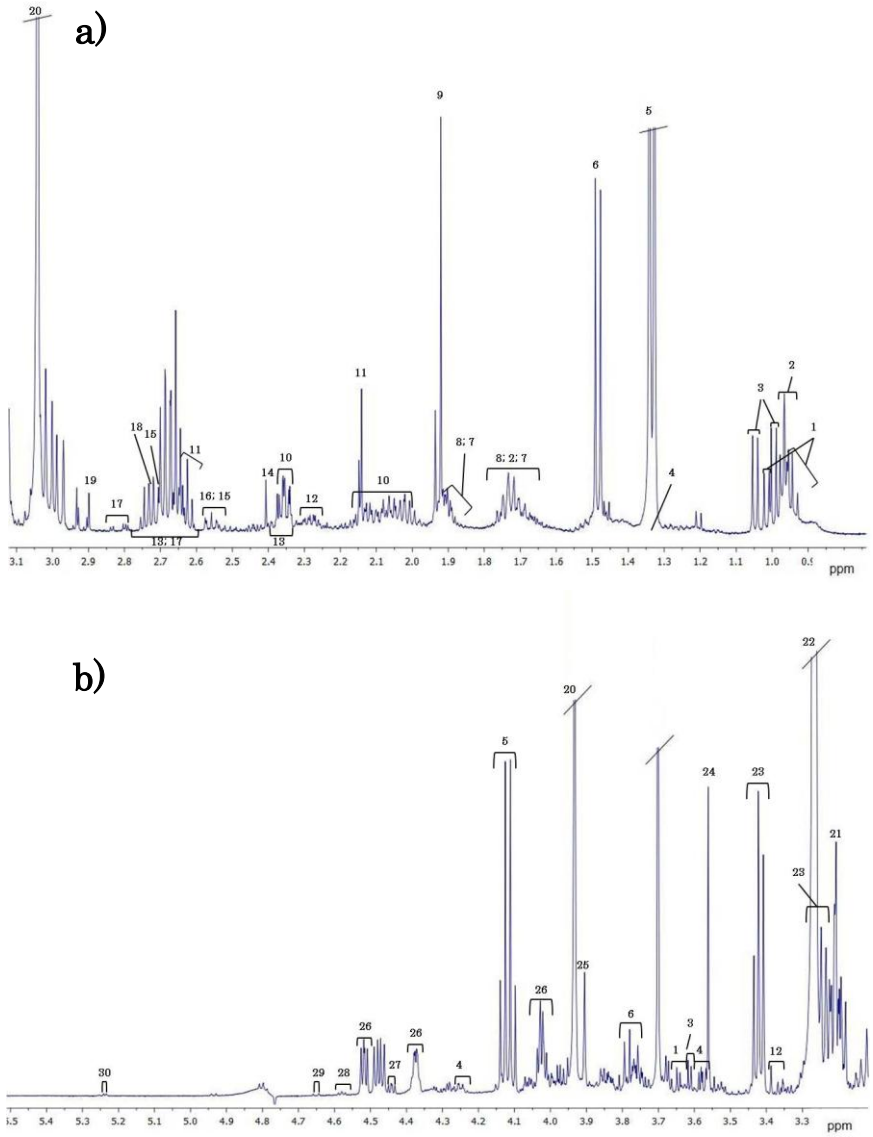

Figure 5.4 Expansions (a-d) of the ^{1}H NMR spectrum of the aqueous extract of muscle of Argentina. Assignments for the numbered resonances are given in Table 5.1.

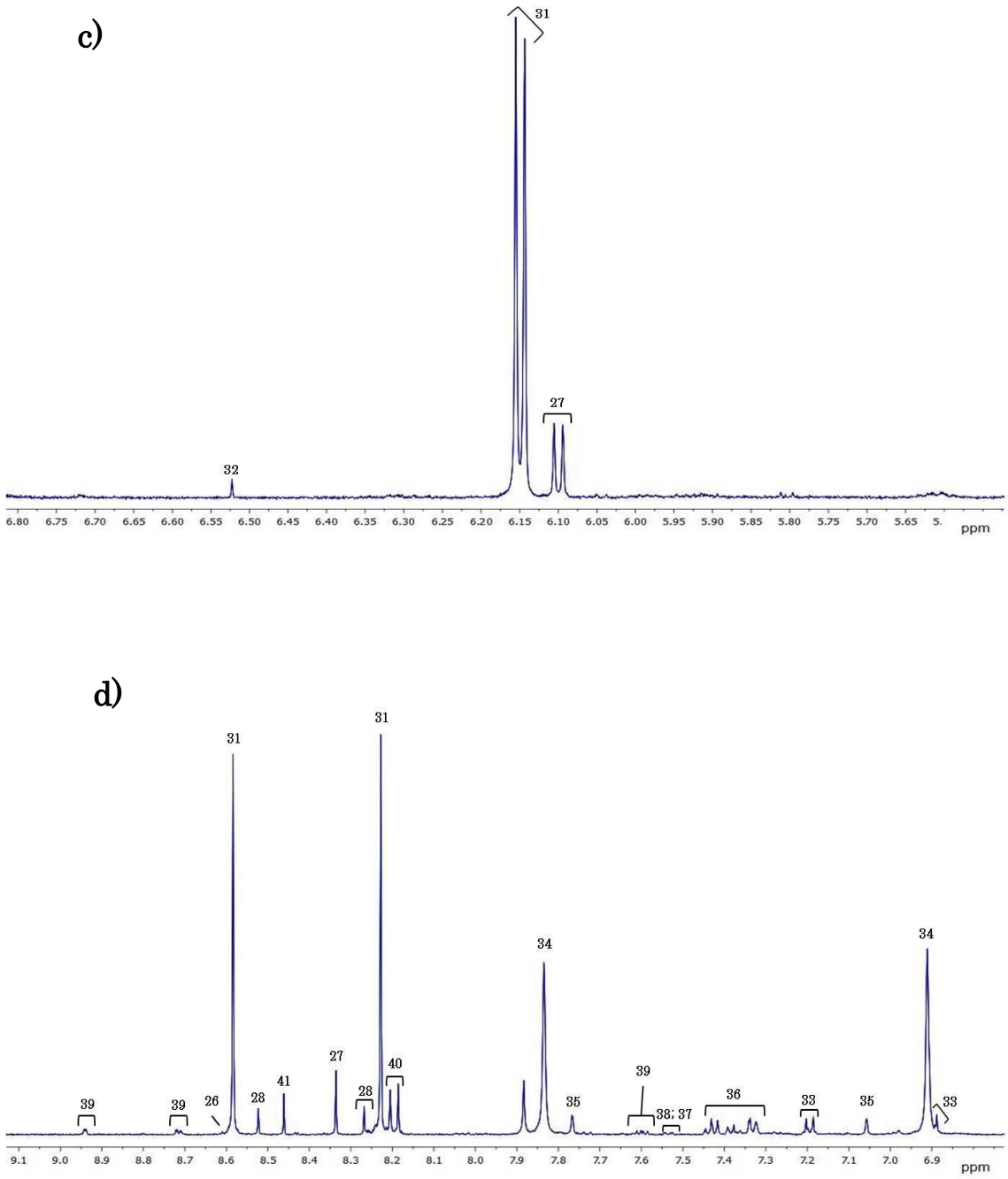

Figure 5.4 continued

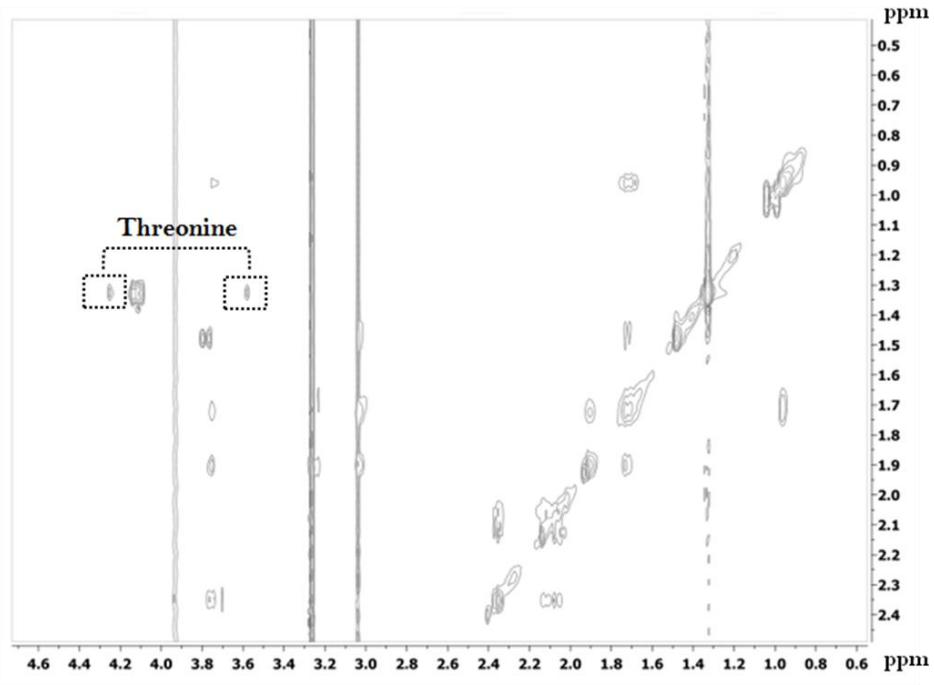

Figure 5.5 2D TOCSY expansion of the aqueous extract of muscle of Argentina, showing the cross-peak of threonine.

leucine (Leu), lisine (Lys), methionine (Met), proline (Pro), tirosine (Tyr), valine (Val) and tryptophan (Trp) were identified. Gly is known to participate at a number of metabolic reactions like the formation of porphyrins, purines, and many other metabolites (Neuberger A., 1961). Furthermore, Gly, Ala, and Ser are interconvertible and their presence in muscle can be justified (Kaushik S.J. and Luquet P., 1979). Finally, taurine, an important antioxidant (Aruoma O.I. et al., 1988), was identified by the characteristic resonances at 3.28 and 3.44 ppm. The deficiency of this metabolite results in a deterioration of eyesight (Hayes K.C. et al., 1975).

The low-field region (Figure 5.d) was characterized by singlets of hypoxanthine (Hyp) (8.19 e 8.21 ppm) and inosine (Ino) (8.34 e 8.23 ppm). The values of Ino and Hyp increase during the storage period by the action of autolytic and microbiological enzymes that determine the spoilage of the food (Surette M.E. et al., 1988). In particular, Hyp, can be considered a quite accurate indicator of freshness in many fish species (Kyrana V.R. et al., 1997; Jacober L.F. et al., 1982; Zhang H.Z. et al., 1997).

The correlation peaks COSY between the signals at 8.27, 7.60 and 8.62 ppm showed the presence of a water-soluble vitamins as niacinamide (or nicotinamide). The niacinamide is part of two coenzymes, nicotinamide adenine dinucleotide (NAD) and nicotinamide adenine dinucleotide phosphate (NADP). The compounds involved in many redox reactions and in the synthesis of fatty acids and amino acids.

Trimethylamine-N-oxide (TMAO), an important osmolyte in fishes, was identified by the intense singlet at 3.27 ppm, coming from protons in the $-N(CH_3)_4$ group. The importance of osmolytes in diet is due to their relevant roles in preserving human health which make them really bioactive compounds. TMAO has recently been shown to prevent the misfolding of the prion protein (Bennion B.J. et al., 2004). The presence of TMAO in the aqueous extract of fish suggested also a good preservation state. In fact, during the storage, TMAO is degraded to the volatile trimethylamine (TMA) responsible for the unpleasant fish odor. Another singlet at 2.73 ppm was identified as dimethylamine (DMA). DMA is diagnostic of freezing processes.

Table 5.2 ^{1}H-NMR chemical shifts of metabolites identified in the aqueous extract of muscle of *Argentina sphyraena*

Free amino acids and dipeptides

Compound	Group	(ppm)[a]	Multiplicity[b]	Correlation[c]
Alanine (Ala) **6**	αCH	3.79	q	1.48 (C)
	βCH$_3$	1.48	d	3.79 (C)
β-Alanine(β-Ala) **16**	αCH$_2$	2.57	t	3.18 (C)
	βCH$_3$	3.18	t	2.57 (C)
Arginine (Arg) **8**	αCH	3.77	t	1.92 (C)
	βCH$_2$	1.92	m	3.77 (C), 1.70 (C)
	γCH$_2$	1.70	m	1.92 (C), 3.25 (C)
	δCH$_2$	3.25	t	1.70 (C)
Aspartate (Asp) **17**	αCH	3.91	dd	2.69 (C), 2.80 (C)
	βCH	2.69	dd	2.80 (C), 3.91 (C)
	β'CH	2.80	dd	2.69 (C), 3.91 (C)
Creatine (Crt) / Phosphocreatine(PCrt) **20**	N-CH$_3$	3.04	s	3.93 (T)
	N-CH$_2$	3.93	s	//
Glycina (Gly) **24**	αCH$_2$	3.56	s	//
Glycina-Betaine (GB) **25**	N-CH$_3$	3.26	s	//
	αCH$_2$	3.90	s	
Glutamate (Glu) **10**	αCH	3.76	t	2.10 (C), 2.38(T)
	β,β'CH	2.10	m	2.38 (C), 3.76 (C)
	γCH$_2$	2.38	t	2.10 (C), 3.76 (T)
Histidine (His) **35**	C2H, ring	7.77	s	//
	C4H, ring	7,06	s	//
Histidine in Anserine (Ans) **34**	C2H, ring	6.91	s	//
	C4H, ring	7.83	s	
Isoleucine (Ile) **1**	αCH	3.68	m	1.98 (C)
	βCH	1.98	m	1.02 (C), 1.26(C), 3.68(C)
	γCH	1.47	m	0.93 (C), 1.26(C)
	γ'CH	1.26	m	0.93(C), 1.47(C), 1.98 (C)
	γ'CH3	1.02	d	1.98 (C)
	δCH3	0.93	t	1.26 (C), 1.47 (C)

Table 5.2 (Continued)

Compound	Group	(ppm)[a]	Multiplicity[b]	Correlation[c]
Leucine (Leu) **2**	αCH	3.75	t	1.70 (C)
	βCH_2	1.70	m	3.75 (C)
	γCH	1.73	m	0.97 (C)
	δCH_3,$\delta' CH_3$	0.97	d	1.73 (C)
Lysine (Lys) **7**	αCH	3.78	t	1.92 (C)
	βCH_2	1.92	m	1.45 (C), 3.04 (T), 3.78 (C)
	γCH_2	1.45	m	1.74 (C), 1.92 (C), 3.04 (T)
	δCH_2	1.74	m	1.45 (C), 3.04 (C)
	εCH_2	3.04	t	1.45 (T), 1.74 (C), 1.92 (T)
Methionine (Met) **11**	αCH	3.85	t	2.19 (C)
	βCH_2	2.19	m	2.64 (C), 3.85 (C)
	γCH_2	2.64	t	2.20 (C)
	S-CH_3	2.14	s	//
Phenylalanine(Phe) ***36***	αCH	3.99	dd	3.13(C), 3.29 (C)
	βCH	3.29	dd	3.99 (C)
	β'CH	3.13	dd	3.99 (C)
	C2,6H, ring	7.42	m	//
	C3,5H, ring	7.42	m	7.33 (C)
	C4H, ring	7.33	m	7.42 (C)
Proline (Pro) **12**	αCH	4.14	t	2.05 (C), 2.38 (C), 2.00 (T)
	βCH	2.36	m	4.14 (C)
	β'CH	2.05	m	4.14 (C)
	γCH_2	2.00	m	3.38 (C), 3.40 (C), 4.14 (C)
	δCH	3.38	t	2.05 (T), 2.38 (T)
	δ'CH	3.40	t	2.05 (T), 2.38 (T)
Taurine (Tau) ***23***	N-CH_2	3.28	t	3.44 (C)
	S-CH_2	3.44	t	3.28 (C)
Threonine(Thr) **4**	αCH	3.58	d	4.27 (C)
	βCH	4.27	m	1.33 (C), 3.58 (T)
	γCH_3	1.33	d	4.27 (C)

Table 5.2 (Continued)

Compound	Group	(ppm)[a]	Multiplicity[b]	Correlation[c]
Triptophane (Trp) **38**	C4H, ring	7.74	d	7.19 (C), 7.27 (T), 7.54 (T)
	C5H, ring	7.19	t	7.54 (T), 7.74 (C)
	C6H, ring	7.27	t	7.54 (C), 7.74 (T)
	C7H, ring	7.54	d	7.19 (T), 7.27 (C), 7.74 (T)
Tyrosine (Tyr) **33**	αCH	3.94	dd	3.06 (C)
	βCH	3.19	dd	3.06 (C)
	β'CH	3.06	dd	3.19 (C), 3.94 (C)
	C2,6H, ring	6.89	d	7.19 (C)
	C3,5H, ring	7.19	d	6.89 (C)
Valine (Val) **3**	αCH	3.62	d	2.29 (C)
	βCH	2.29	m	1.00 (C), 1.05 (C), 3.62 (C)
	γ'CH_3	1.05	d	1.00 (C), 2.29 (C)
	γCH_3	1.00	d	1.05 (C), 2.29 (C)

Nucleotides and relative compounds

Compound	Group	(ppm)[a]	Multiplicity[b]	Correlation[c]
AMP **26**	C1'H, ribosio	6.14	d	//
	C2'H, ribose	4.79	//	//
	C3'H, ribose	4.52	//	//
	C4'H, ribose	4.37	//	//
	C5'H, ribose	4.03	//	//
	CH-2	8.60	s	//
	CH-8	8.27	s	
ADP/ ATP **28**	CH-2	8.27	s	//
	NH, ring	8.52	s	//
	C1'H, ribose	6.15	d	//
Inosine (Ino) **27**	C1'H, ribose	6.11	d	4.30 (T), 4.46 (T), 4.78 (C)
	C2'H, ribose	4.78	m	4.30 (T), 4.46 (T), 6.11 (C)
	C3'H, ribose	4.46	m	4.30 (T), 4.78 (C), 6.11 (T)
	C4'H, ribose	4.30	m	4.46 (T), 4.78 (T)

Table 5.2 (Continued)

Compound	Group	(ppm)[a]	Multiplicity[b]	Correlation[c]
	C5'H, ribose	4.02	m	4.36 (C)
	C2H, ring	8.34	s	//
	C8H, ring	8.23	s	//
Inosine monophosphate (IMP) **31**	C1'H, ribose	6,15	d	4.04 (T), 4.38 (T), 4.52 (T), 4,81 (C)
	C2'H, ribose	4,81	m	4.38 (T), 4.53 (T), 6,15 (C)
	C3'H, ribose	4.52	dd	4.04 (T), 4.38 (T), 4.81 (T)
	C4'H, ribose	4.38	m	4.04 (T), 4.52 (T), 4.81 (T)
	C5'H, ribose	4.04	m	4.38 (T),4.52 (T), 4.81 (T)
	C2H, ring	8.59	s	//
	C8H, ring	8,23	s	//
Hypoxantine (Hyp) **40**	C2H, ring	8.21	s	//
	C8H, ring	8.19	s	//
Niacinamide **39**	N5, ring	7.60	q	8.27 (C), 8.62 (C)
	N4, ring	8.27	d	7.60 (C), 8.62 (T)
	N2, ring	8.95	s	//
	N6, ring	8.62	d	7.60 (C), 8.27 (T)

Sugars

Compound	Group	(ppm)[a]	Multiplicity[b]	Correlation[c]
α-Glucose **30**	C1H	5.24	d	3.52 (C)
	C2H	3.52	dd	5.24 (C)
β-Glucose **29**	C1H	4.66	d	3.27 (C)
	C2H	3.27	dd	4.66 (C)

Organic acid

Compound	Group	(ppm)[a]	Multiplicity[b]	Correlation[c]
Acetic acid (Ace) **9**	βCH_3	1.92	s	//
Citric acid (Cit) **15**	C2H	2.56	dd	2.73 (C)
	C4H	2.73	dd	2.56 (C)
Formic acid (Form) **41**	$HCOO^-$	8.46	s	//
Fumaric acid (Fu)**32**	$\alpha,\beta CH$	4.33	dd	2.38 (C)
Lactic acid (Lac) **5**	CH_3	1.33	d	4.12 (C)
	CH_2	4.12	q	1.33 (C)
Malic acid (Mal) **13**	αCH	4.33	dd	2.38 (C)
	βCH_2	2.38	dd	2.70 (C) , 4.33 (C)
	$\beta' CH_2$	2.70	dd	2.38 (C)
Succinic acid (Suc) **14**	$\alpha,\beta CH_2$	2.41	s	//

Osmolytes and other compounds

Compound	Group	(ppm)[a]	Multiplicity[b]	Correlation[c]
Dimethylamine (DMA) **18**	$N(CH_3)_2$	2.73	s	//
Uracil (Ura) **37**	C5H, ring	5.81	d	7.55 (C)
	C6H, ring	7.55	d	5.81 (C)
Trimethylamine(TMA) **19**	$N-(CH_3)_3$	2.90	s	//
Trimethylamine O (TMAO) **22**	$O-N-(CH_3)_3$	3.27	s	//
Choline (Cho) 21	$N-(CH3)_3^+$	3.21	s	//
	N–CH2	4.07	m	3.52 (C)
	O–CH2	3.52	m	4.07 (C)

5.3.3 ^{1}H NMR spectrum of muscle lipid extract of *Argentina sphyraena*

The representative ^{1}H NMR spectrum of *Argentina sphyraena* chloroform extract is shown in Figure 5.6. The assignment of ^{1}H NMR spectrum was obtained by performing 2D-NMR experiments and by using literature data (Mannina L. et al., 2008; Aursand M. et al., 1993; Sacchi R. et al., 1993; Scano P. et al., 2006) (Table 5.3). Detailed chemical shift attributions are shown in Figure 5.7.

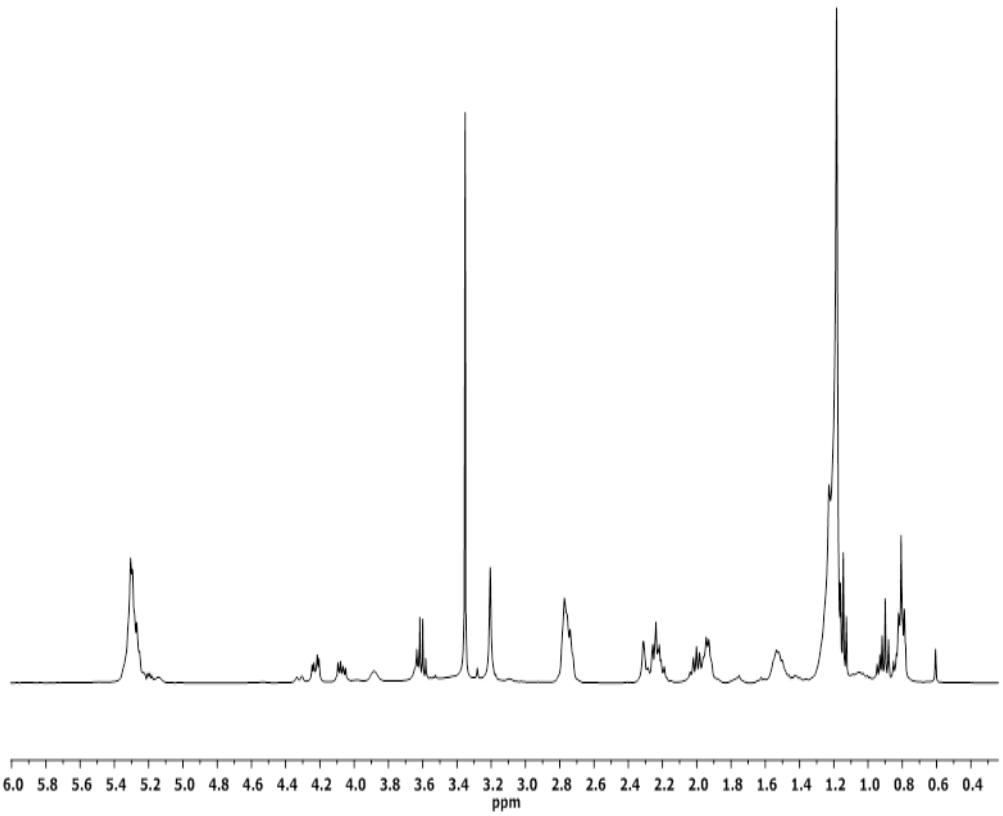

Figure 5.6 A representative ^{1}H NMR spectrum of the lipid extract of *Argentina sphyraena* muscle

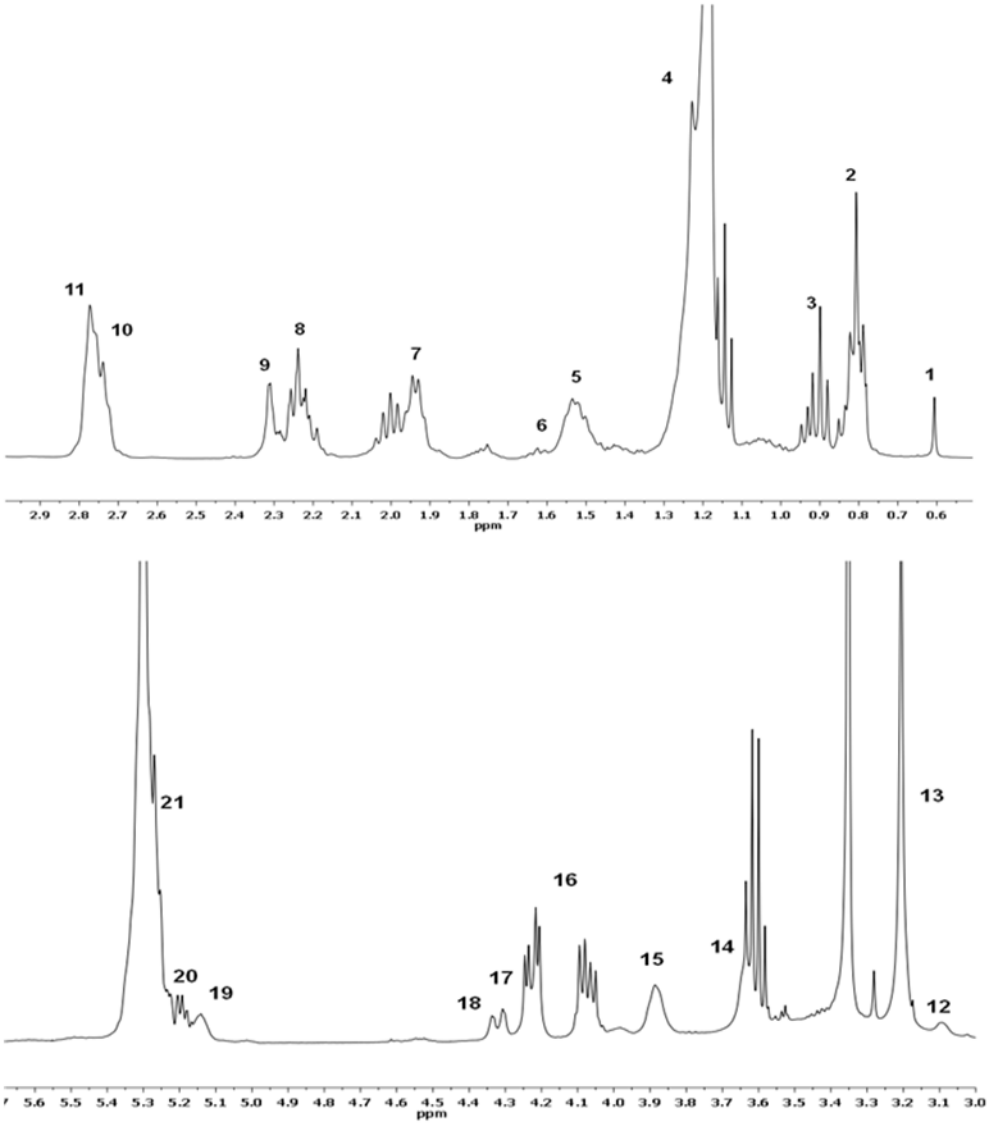

Figure 5.7 Expansions of the ^{1}H NMR spectrum of the lipid extract of muscle of argentina. Assignments for the numbered resonances are given in Table 5.3.

Table 5.3 ^{1}H NMR chemical shift assignment for the lipid extract of muscle of argentina. Numbering of peaks references to Figure 5.7.

Peak	Carbon	Compound[a]	ppm[a]
1	$-CH_3$	CHO	0.64
2	$-CH_3$	All fatty acids except ω3	0.84
3	$-CH_3$	Fatty acids ω3	0.94
4	$-(CH_2)_n$	All fatty acids except PUFA	1.22
5	$-\underline{C}H_2CH_2$	All fatty acids except DHA e EPA	1.57
6	$-\underline{C}H_2CH_2$	EPA	1.66
7	$-\underline{C}H_2CH=$	All fatty acids except DHA	1.95-2.15
8	$-\underline{C}H_2COOH$	All fatty acids except DHA	2.23-2.33
9	$-\underline{C}H_2CH_2$	DHA	2.35
10	CH_2 diallilici	DUFA	2.78
11	$-CH_2-$	PUFA	2.81
12	CH_2N-	PE	3.15
13	$(CH_3)_3N-$	PC	3.24
14	CH_2N	PC	3.68
15	CH_2 *sn3*	PC e PE	3.93
16	CH_2 *sn1,3*	TG	4.15-4.30
17	CH_2 *sn1*	PC e PE	4.35
18	CH_2OP	PC	4.38
19	CH *sn2*	PC e PE	5.18
20	CH_2*sn2*	TG	5.24
21	CH=CH	All fatty acids	5.33

[a] The chemical shifts are referenced to the peak of $CHCl_3$ (δ=7.26 ppm)

b Abbreviations: TG, triglycerides; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PUFA, polyunsaturated fatty acids; DUFA, diunsaturated fatty acids; EPA, eicosapentaenoic acid. DHA docosahexaenoic acid.

5.3.4 Multivariate statistical analysis

A preliminary analysis of NMR data relative to the aqueous and organic extracts of argentina was performed to test the potentiality of NMR to detect seasonal variations.

Metabolomic changes in the water soluble extract. Figure 5.8 shows the PCA model applied on the full ^{1}H NMR data set of water-soluble metabolites. No significant seasonal differentiation was found, noticeable overlaps being observed among samples.

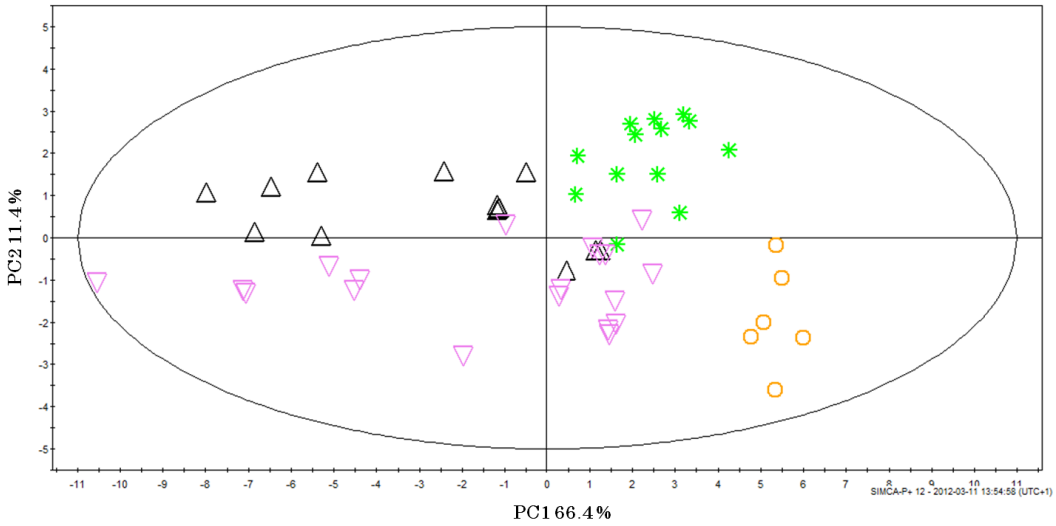

Figure 5.8 PCA score plot derived from the ^{1}H NMR spectra of the aqueous extract of Argentina muscle: **December; January; February; April.**

A more distinct differentiation according to the fishing period was achieved by OPLS-DA (Figure 5.9a), as reflected by a higher goodness of fit and predictability (R_x^2=0.94; R_y^2=0.86; Q^2=0.71), compared with the PCA model (R_x^2=0.66; Q^2=0.61). The improvement of model by OPLS-DA suggests that the structured noise was the principal source of data variation in the NMR spectra. The corresponding loadings (Figure 5.9b) indicate that the samples caught in December were characterized by the largest content of Tau, α and β Glucose, Asp, Ans, and Pro; samples fished in January by Cho, TMAO, ADP, and ATP; Crt/P-Crt, IMP, AMP, and Lac were the main representative metabolites of February; finally, April's

samples have the highest level of free aminoacids (Ala, Leu, Ile, Val, Lys, Arg, Met, Glu, Phe, Thr, and Gly) and acetic and formic acid.

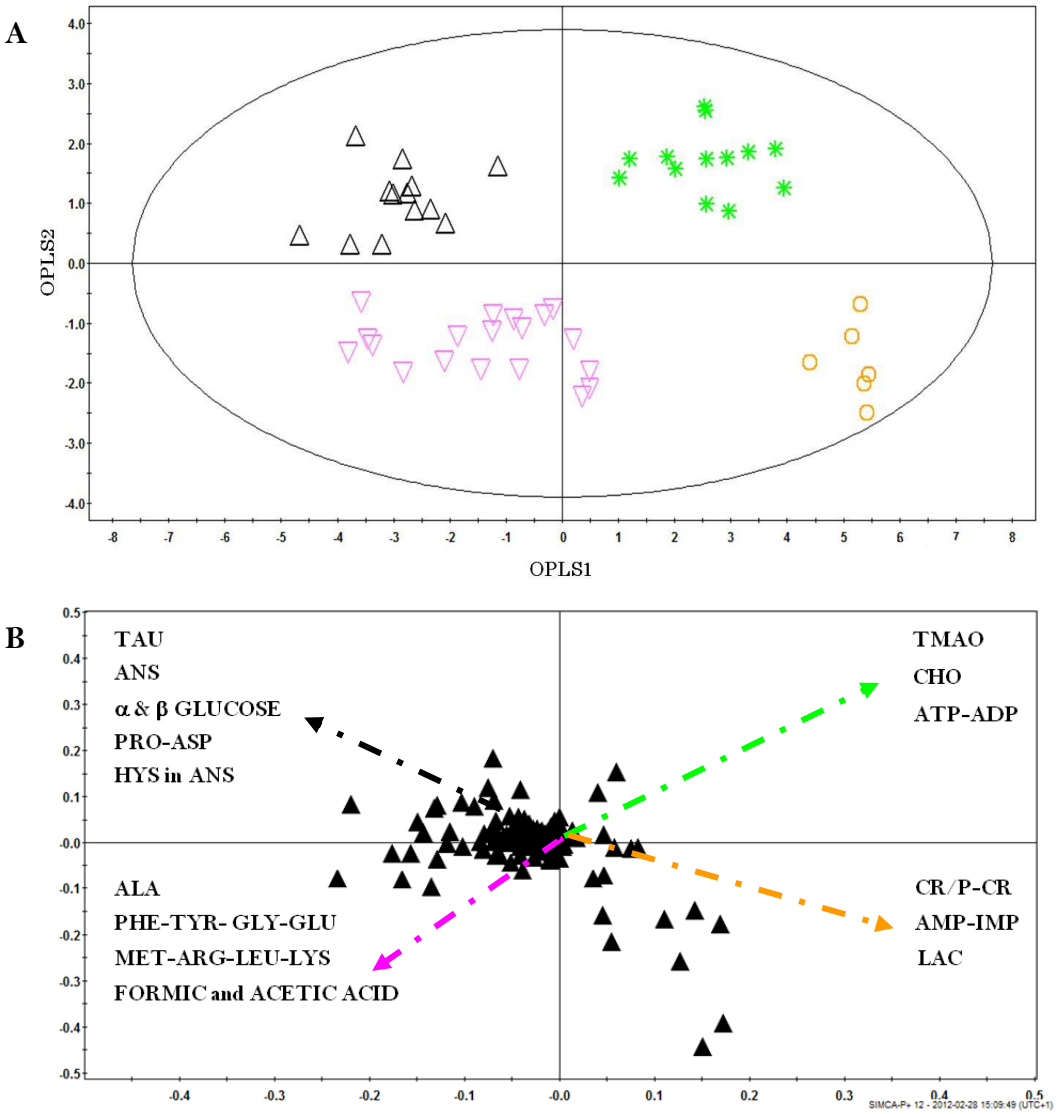

Figure 5.9 OPLS-DA score (a) and loading (b) plot derived from the ^{1}H NMR spectra of the aqueous extract of argentina caught in Teulada at a depth of 150 m: **December; January; February; April.**

It is interesting to note that, among the metabolites most representative of each fishing period, some can be reliable biomarkers for providing information on the seasonal variations of these components, while other can be involved also in *post-mortem* processes which, in turn, can be accelerated in case of acute stress during fishing. Indeed, for instance, free amino acids in fish play essential roles in metabolism, such as adjustment of osmotic pressure and as energy source. In particular, the imidazole compounds, such as histidine, make a large contribution to the buffering capacity in fish muscle and reflect its biochemical condition (Suyama M. et al., 1986). Differently, the accumulation of inosine and hypoxanthine in fish species is related to the *post-mortem* ATP depletion, due to both autolytic and microbial action, although the former seems more important. On account of this, nucleotide degradation products have been widely used as indicators of storage age or freshness. Indeed, a high content of inosine monophosphate in the flesh indicates a high level of freshness whereas a high content of hypoxanthine indicates that spoilage phase will soon start, if not started yet (Spinelli J. et al., 1964). In the light of these considerations, it is therefore evident that the analysis of the whole NMR profiling of the water soluble metabolite extract of fish requires significant development efforts in order to get reliable information on the biology of this species.

Metabolomic changes in the lipid extract. PCA on the ^{1}H NMR lipid extract data set failed to highlight any metabolic difference among samples based on seasonal variations (data not shown). Thus, in order to emphasize the different among samples, an OPLS-DA was used. The scores plot in Figure 5.10a showed a clear patter of arrangement of specimens according to the fishing period. In particular, samples caught in April showed group homogeneity slightly higher than samples caught in December and January. By examining the corresponding loadings plot (Figure 5.10b), the important contribution to the discrimination among samples was found to be the level of polyunsaturated fatty acids (PUFA), higher in samples caught in April, that is in the spawning period, and lower in December, i.e. in the post-spawning stage. It is known that the level of PUFA in fishes is highly dependent on the food availability. Thus, the maximum level of PUFA found in April may be related to increasing feeding. This hypothesis is in

line with the increase with HSI found in spring samples (Figure 5.1). However, it can not be ruled out also an importance influence by the source of dietary lipids (Shirai N. et al., 2001).

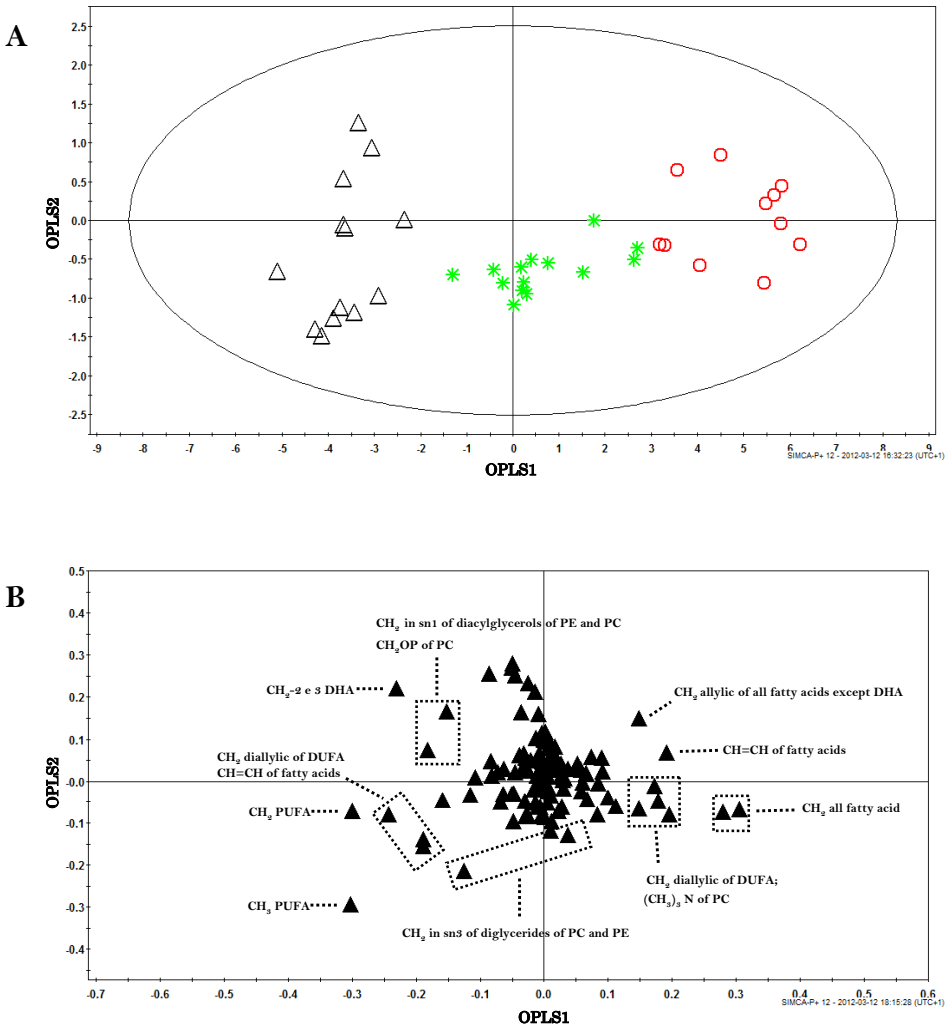

Figure 5.10 OPLS-DA score (a) and loading (b) plot derived from the ^{1}H NMR spectra of the lipid extract of argentina caught in Teulada at a depth of 150 m: **December;** **January;** **April.**

5.4 CONCLUSION

This work represents a first attempt to combine NMR spectroscopy and multivariate statistical analysis to investigate seasonal variations of the metabolic profile of fish muscle. Although this is a preliminary study, the potential of this approach is clear. On one side, the water soluble fraction of muscle is riched of metabolites playing important role in physiological functions and, thus, useful to investigate biochemical conditions of muscle. However, due to the concomitant presence of metabolites indicative of *post-mortem* metabolic processes, the results must be interpreted carefully in view of using such a model to investigate seasonal variations. On the other hand, the lipid profile allows extracting information that could be correlated with the diet of fish and its feeding activity. By taking more samples and planning a better experimental design, the desired task can be made. Toward this direction, the present work has been evolving.

References

Aruoma, O.I.; Halliwell, B.; Hoey, B.M.; Butler, J. The antioxidant action of taurine, hypotaurine and their metabolic precursors. *J. Biochem.* **1988**, *256*, 251-255.

Aursand, M.; Rainuzzo, J.R.; Grasdalen, H. Quantitative high-resolution ^{13}C and ^{1}H nuclear magnetic resonance of w-3 fatty acids from white muscle of Atlantic Salmon (*Salmo salar*). *Jaocs* **1993**, *70*, 971–981.

Aursand, M.; Standal, I.B.; Praël, A.; McEvoy, L.; Irvine, J.; Axelson, D.E.; ^{13}C NMR Pattern Recognition Techniques for the Classification of Atlantic Salmon (*Salmo salar* L.) According to Their Wild, Farmed, and Geographical Origin. *J. Agric. Food Chem.* **2009**, *57*, 3444-3451.

Bennion, B.J.; DeMarco, M.L.; Daggett, V. Preventing misfolding of the prion protein by trimethylamine N-oxide. *Biochem.* **2004**, *43*, 12955–12963.

Boldyrev, A.A.; Severin, S.E. The histidine-containing dipeptides, carnosine and anserine: Distribution, properties and biological significance. *Adv. Enzyme Regul.* **2005**, *30*, 175-188.

Carbonell A.; Martin P.; De Ranieri S. Discards of the Western Mediterranean Trawl Fleets. *Rapp. Comm. Int. Mer Médit.* **1998**, *35*.

Commission of the European Communities (EC), 2009: Green paper: reform of the common fisheries policy [COM(2009) 163 final]. The European Commission, Brussels.

D'onghia, G.; Mastrototaro, F.; Matarrese, A.; Politou, C.J.; Mytilinenou, C. Biodiversity of the upper slope demersal community in the eastern Mediterranean: preliminary comparison between two areas with and without trawl fishing. *J.Northw.Atl.Fish.Sci.* **2003**, *31*, 263-273.

Fan, T. Metabolite profiling by one- and two-dimensional analysis of complex mixtures. *Prog. NMR Spectrosc.* **1996**, *28*, 161–219.

Gribbestad, I.S.; Aursand, M.; Martinez, I. High-resolution 1H magnetic resonance spectroscopy of whole fish, fillets and extracts of farmed Atlantic salmon (Salmo salar) for quality assessment and compositional analyses. *Aquaculture* **2005**, *250*, 445–457.

Halliday, R.G. Population parameters of Argentina Sphyraena [Isospondyli] from west of Britain. *J. mar. biol. Ass. U.K.* **1969**, *49*, 407-431.

Hall S.J. The Effects of Fishing on Marine Ecosystems and Communities. *Fish Bio. and Aquatic Res.* **1999**, *1*, 296.

Jacober, L.F.; Rand, J.A.G. Biochemical evaluation of seafood. In: Martin, R.E.; Flick, G.J.; Hebard, C.E.; Ward, D.R. Chemistry & biochemistry of marine food products. Wesport CT: AVI Publishing Company **1982**, 347-366.

Kaushik, S.J.; Luquet P. Influence of dietary amino acid patterns on the free amino acid contents of blood and muscle of rainbow trout (*Salmo gairdnerii* R.). *Comp. Biochem. Physiol. Part B: Comp. Biochem.* **1979**, *64*, 175-180.

Kyrana, V.R.; Lougovois, P.; Valsamis, D.S. Assessment of shelf life of maricultured gilthead sea bream (Sparus aurata) stored in ice. *Int. J. Food Sci Tech.* **1997**, *32*, 339-347.

Ozyurt G.; Polat, A. Amino acid and fatty acid composition of wild sea bass (Dicentrarchus labrax): a seasonal differentiation. Eur. Food Res. Technol. **2006**, *222*, 316–320.

Machias, A.; Vassilopoulou, V.; Vatsos, D.; Bekas, P.; Kallianiotis, A.; Papaconstantinou, C.; Tsimenides, N. Bottom trawl discards in the northeastern Mediterranean *Sea. Fish. Res.* **2001**, *53*, 181-195.

Mannina, L.; Sobolev, A.P.; Capitani, D.; Iaffaldano, N.; Rosato, M.P.; Ragni, P.; Reale, A.; Sorrentino, E.; D'Amico, I.; Coppola, R. NMR metabolic profiling of organic and aqueous sea bass extracts: Implications in the discrimination of wild and cultured sea bass. *Talanta* **2008**, *77*, 433–444.

Martinez, I.; Bathen, T.; Standal, I.B.; Halvorsen, J.; Aursand, M.; Gribbestad, I.S.; Axelson, D.E. Bioactive Compounds in Cod (*Gadus morhua*) Products and Suitability of ^{1}H NMR Metabolite Profiling for Classification of the Products Using Multivariate Data Analyses. *Agric. Food Chem.* **2005**, *53*, 6889–6895.

Martinez, I.; Standal, I.B.; Axelson, D.E.; Finstad, B.; Aursand, M. Identification of the farm origin of salmon by fatty acid and HR 13C NMR profiling. *Food Chem.* **2009**, *116*, 766–773.

Neuberger, A. Aspects of the metabolism of glycine and of porphyrins. The second Hopkins Memorial Lecture. *Biochem.* **1961**, *78*, 1-10.

Sacchi, R.; Medina, I.; Aubourg, S.P.; Giudicianni, I.; Paolillo, L.; Addeo, F. Quantitative high-resolution 13C-NMR analysis of lipids extracted from the white Muscle of Atlantic Tuna (*Thunnus alalunga*). *J. Agric. Food Chem.* **1993**, *41*, 1247–1253.

Samuelsson, L.M.; Joakim Larsson, D.G. Contributions from metabolomics to fish research. *Mol. BioSyst.* **2008**, *4*, 974–979.

Savorani, F.; Picone, G.; Badiani, A.; Fagioli, P.; Capozzi, F.; Engelsen, S. B. Metabolic profiling and aquaculture differentiation of gilthead sea bream by ^{1}H NMR metabonomics. *Food Chem.* **2010**, *120*, 907–914.

Scano, P.; Cesare Marincola, F.; Locci, E.; Lai, A. ^{1}H and ^{13}CNMRstudies of melon and head blubber of the striped Dolphin (*Stenella coeruleoalba*). *Lipids* **2006**, *41*, 1039–1048.

Shirai, N.; Terayama, M.; Takeda, H. Effect of season on the fatty acid composition and free amino acid content of the sardine Sardinops melanostictus. *Comp. Biochem. Physiol.* **2001**, *131*, 387-393.

Standal, I. B.; Gribbestad, I. S.; Bathen, T. F.; Aursand, M.; Martinez, M. I. G. Low molecular weight metabolites in white musclefrom cod (Gadus morhua) and haddock (Melanogrammus aeglefinus) analyzed by high resolution ^{1}H NMR spectroscopy. *Magn. Reson. Food Sci.* **2007**, 55–62.

Standal I.B. Use of NMR spectroscopy in combination with pattern recognition techniques for elucidation of origin and adulteration of foodstuffs. *PhD Thesys. Norwegian University of Science and Technology* **2009.**

Solansky, K.S.; Burton, W.; MacKinnon, S.L.; Walter, J.A.; Dacanay, A. Metabolic changes in Atlantic salmon exposed to Aeromonas salmonicida detected by 1H-nuclear magnetic resonance spectroscopy of plasma. *Dis. Aquat. Org.* **2005**, *65*, 107–114.

Spinelli, J.; Eklund, M.; Miyauchi, D. Measurement of Hypoxanthine in Fish as a Method of Assessing Freshness. J. of Food Sci. **1964**, *29*, 710-714.

Suyama, M.; Hirano, T.; Suzuky, T. Buffering capacity of free histidine and its related dipeptides in white and dark muscles of yellowfin tuna. Bull. Japan. Soc. Sci. Fish. **1986**, *52*, 2171-2175.

Stergiou, K.I. Variability of fish catches in different ecosystems. In: Durand M.E.; Cury P.; Mendelssohn R.; Roy C.; Bakun A.; Pauly D. (Eds.), Global Versus Local Changes in Upwelling Systems. ORSTOM Editions, Paris, **1998**, 359–370.

Surette, M.E.; Gill, T.A.; LeBlanc, P.J. Biochemical basis of postmortem nucleotide catabolism in cod (Gadus morhua) and its relationship to spoilage. J. Agr Food Chem. **1998**, *36*, 19- 22.

Walliman, T. The phosphocreatine circuit: Molecular and cellular physiology of creatine kinases, sensitivity to free radicals, and enhancement by creatine supplementation. In S. Valdur (Ed.), Molecular system bioenergetics: Energy for life **2007**,195–264. Wiley-VCH.

Yamori, Y.; Liu, L.; Mori, M.; Sagara, M.; Murakami, S.; Nara, Y.; Mizushima, S. Taurine as the Nutritional Factor for the Longevity of the Japanese Revealed by a World-Wide Epidemiological Survey. Adv. Exp. Med. Biol. **2009**, *643*, 13-25.

Zhang, H.Z.; Lee, T.C. Gas chromatography-mass spectrometry analysis of volatile flavour compounds in mackerel for assessment of fish quality. In: Shahidi, F.;

Cadwallader, K.R. Flavour and lipid chemistry of seafoods, ACS Symposium Series 674. Washington DC: American Chemical Society **1997**, 55-63.

6. General Conclusions

In the recent years, metabolomics has been gaining considerable attention in food science, proving to be a valuable tool for the biochemical analysis of foodstuff. Indeed, metabolomics has the potential to generate a comprehensive set of biomarkers which can be used to understand and monitor food properties such as shelf life and quality. Among the analytical techniques used in metabolomics for assessing food constituents, high-field ^{1}H NMR spectroscopy is unique in its ability to describe the chemical profile of the sample and provide very quickly information about a large number of compounds (i.e. amino acids, carbohydrates, and organic and fatty acids). However, ^{1}H NMR spectra of even rather simple single-phase foods often result in complex spectra. For this reason, it is advantageous to analyze the spectra by multivariate methods like those developed in the field of chemometrics.

The present Ph.D. work showed some applications on the NMR-based metabolomic approach. The investigated food matrices were largely different, from a manufactured product that underwent only physical treatments (*bottarga*), to a manufactured product where biochemical transformations took place (*Fiore Sardo* cheese), and a raw food (*Argentina sphyraena*). All of these food matrices were not chosen by chance, but they represent an important piece of economy of the island of Sardinia, or might be further valorized, gaining more importance in the near future. Indeed, bottarga and Fiore Sardo are typical products exported all over the world, while *Argentina sphyraena* is a fish a low economic interest, finding no appreciation, at the moment, on the market.

The results of this PhD study have contributed with new insights and deeper understanding of the potential perspective of the combined NMR/multivariate methods approach in food science, showing the great versatility of NMR spectroscopy and the strong synergetic relation between NMR and chemometrics. NMR revealed its extraordinary potential, when applied to natural samples and products, while chemiometric analytical technique proved to be an essential tool to

get information on the properties of interest (e.g., geographical origin for bottarga) based on the knowledge of other properties easily obtained (i.e. NMR spectra).

The results obtained by the investigation of bottarga demonstrated that a NMR-based metabolomics technique can be a powerful tool for the detection of novel biomarkers and establishing quality control parameters for bottarga. The work presented in this study evidenced the effectiveness of metabolite fingerprinting as a tool to distinguish samples according both to the geographical origin of fish and the manufacturing process.

The results relative to the Fiore Sardo showed the potential of the combination of NMR spectroscopy and chemometrics as a promising partnership for detailed cheese analysis, providing knowledge that can facilitate better monitoring of the food production chain and create new opportunities for targeted strategies for processing. Such analysis may be performed in any stage of the cheese manufacturing, allowing for thorough evaluation of every step in the process.

Finally, the preliminary results relative to the metabolomic investigation of *Argentina sphyraena* should certainly serve as a basis for implement a research tool able to provide deeper insights on the biology of this fish species with all advantages offered by the metabolomics approach.

ACKNOWLEDGEMENTS

First of all I would like to express my sincere gratitude to my supervisor Dott.ssa Flaminia Cesare Marincola for her encouragement and precious suggestions during this work and for her invaluable help and guidance.

MIUR is also gratefully acknowledged for the financial support.

I also greatly thank Prof. Engelsen and Dr. Savorani of the Life Science Faculty, University of Copenhagen, for their ospitality during three months I spent in their group, and for their valuable support in chemometrics tools application.

I also wish to thank Dott.ssa Barbara Pisano for the microbiological analysis performed on Fiore Sardo and Dott.ssa. Serenella Cabiddu for providing us the individuals of *Argentina sphyraena*. The motivating discussions with them gave an important contribution to the success of my work.

I would like to thank Francesca, Daniela, Maura, Elisa, Shaji, Claudia and Luca, the best colleagues a PhD student could wish for.

I would also like to thank: Marzia, Andrea, Carla, Giorgia, Roberta for their support, enthusiasm and friendship.

Last but not least, I would like to thank Andrea and my family for their huge support. THANKS!