ALMA MATER STUDIORUM-UNIVERSITY OF BOLOGNA

DEPARTMENT OF CHEMISTRY "G. CIAMICIAN"

Ph. D. COURSE IN CHEMICAL SCIENCE

COURSE XXIV

Afference sector: CHIM/03A1

Scientific-disciplinary sector: CHIM/01

PYROLYSIS-GAS CHROMATOGRAPHY-MASS SPECTROMETRY AND CHEMOMETRIC ANALYSIS FOR THE CHARACTERIZATION OF COMPLEX MATRICES

Presented by

SIMONA MONTALBANI

School doctorate coordinator **Advisor**

Prof. ADRIANA BIGI **Prof. DORA MELUCCI**

FINAL EXAM 2012

CONTENTS

INTRODUCTION

Analytical pyrolysis is a powerful technique for rapid analysis of complex and heterogeneous organic materials, based on controlled thermal degradation of the sample in inert atmosphere. The organic material is either degraded to lighter neutral molecules or flash vaporized when thermally stable and volatile. Identification of pyrolysis products is obtained by retention times and mass spectra after coupling pyrolysis with gas chromatography and mass spectrometry (Py-GC-MS). Even though this is a destructive method, a single analysis with a minute amount of sample (less than 1 mg), without preliminary pre-treatment, is capable to provide information on a wide range of organic materials difficult to be analyzed by other methods. Py-GC-MS is extensively applied in the field of cultural heritage, environmental science, forensics and others.

The present PhD thesis was focused on the development and application of chemical methodology and data-processing method, both by classical univariate statistics and by multivariate data analysis (chemometrics). The chromatographic and mass spectrometric data obtained with this technique are particularly suitable to be interpreted by chemometric techniques such as *principal component analysis* (PCA), and classification techniques.

As a first approach, some issues related to the field of cultural heritage were discussed with a particular attention to the differentiation of binders used in pictorial field. In particular, it was possible to determine as a marker of egg tempera the phosphoric acid esterified, a pyrolysis product of lecithin. The analyses were carried out both on phospholipids standards either on painting standard layers prepared by the Opificio delle Pietre Dure (Florence, Italy).

The best results were obtained using as a derivatizing reagent the HMDS (hexamethyldisilazane) rather than the TMAH (tetramethylammonium hydroxide).

Recently, the interest in the use of analytical pyrolysis to study biological samples greatly increased. The validity of analytical pyrolysis as tool to characterize and classify different types of bacteria was verified.

Fatty acids represent the main organic compounds present in the structure of the wall bacterial cell; their chemistry is extremely variable because of the differences in chains length, the presence or absence of unsaturated groups and branched chains and hydroxylated groups. The fatty-acids chromatographic profiles represent an important tool for the bacterial identification.

Because of the complexity of the chromatograms, it was possible to characterize the bacteria only according to their genus, while the differentiation at the species level has been achieved

by means of chemometric analysis. To perform this study, normalized areas peaks relevant to fatty acids were taken into account. Chemometric methods such as PCA (Principal Component Analysis) and SIMCA (soft independent models of class analogy) were applied to experimental datasets. The obtained results demonstrate the effectiveness of analytical pyrolysis and chemometric analysis for the rapid characterization of bacterial species.
Application to a samples of bacterial (*Pseudomonas Mendocina*), fungal (*Pleorotus ostreatus*) and mixed-biofilms was also performed. A comparison with the chromatographic profiles obtained from different biofilms was carried out and it was verified the possibility to:

- Differentiate the sample according to the fatty acid profile
- Characterize the *fungal biofilm* by means the typical pattern of pyrolytic fragments derived from saccharides present in the fungal structure

Individuate the *markers* of bacterial and fungal biofilm in the same mixed-biofilm sample

SECTION A THEORY

CHAPTER A1 ANALYTICAL PYROLYSIS

A1-1 Introduction to analytical pyrolysis

Pyrolysis is defined as a *chemical degradation reaction* caused by thermal energy alone and carried out in an inert atmosphere [1, 2]. The term chemical degradation refers to the decompositions and eliminations that occur in pyrolysis with formation of primary fragments, molecules of lower mass arising from the direct breaking of bonds of the analyte molecules, and secondary fragments of higher mass, resulting from intermolecular reactions between substrate and the primary products not yet degraded. The pyrolytic fragmentation is analogous to the processes that occur during the production of a mass spectrum: the energy supplied determines the breaking of the molecules into stable fragments. The pyrolytic reactions usually take place at temperatures between 500 ° C and 800° C; the chemical transformations taking place under the influence of heat at a temperature between 100°C and 300° C are called thermal degradations and not pyrolysis.

Analytical pyrolysis [3, 4] is by definition the *characterization* of a material by chemical degradation reactions induced by thermal energy, while the pyrolysis itself is just a process that allows the transformation of the sample into other compounds. The pyrolytic process is carried out in a pyrolysis unit (pyrolyzer) interfaced with the analytical instrumentation. It is also possible to perform "off line" pyrolysis (no direct interface analytical instruments), followed by the analytical measurement. The pyrolyzers have a source of heat where the sample is pyrolyzed and the products are usually swept by a gas flow from the pyrolyzer to the analytical instrument. Pyrolysis can be performed in different modes: flash pyrolysis (pulse mode), slow gradient heating pyrolysis (continuous mode), step pyrolysis, etc. Usually, the pyrolysis for analytical purposes is carried out in “pulse mode”, that consists in a very rapid heating of the sample from ambient temperature, targeting isothermal conditions at a temperature where the sample is completely pyrolyzed. There is a fairly wide array of commercially available instruments for performing pyrolysis; most of them are designed primarily for the use with gas chromatographs. Microfurnaces provide a constantly heated, isothermal pyrolysis zone into which samples are introduced by a liquid syringe, solid plunger syringe or in a little cup. Curie-point pyrolyzers apply the sample to a piece of ferromagnetic metal which is inserted into the pyrolyzer when cold, then heated rapidly through induction of current using a high frequency coil. Depending on the metal alloy used, when it reaches a characteristic temperature (the Curie-point of that metal) no more current can be induced, so the temperature stops at that point. Filament style pyrolyzers use a piece of resistive metal

(frequently platinum) with a wide temperature range and circuitry capable of heating the filament up to a programmed temperature at controlled rate.

The production of smaller molecules from some larger favored the use of pyrolysis as a sample preparation technique, extending the applicability of instrumentation designed for the analysis of gaseous species to solids, especially polymeric materials. Polymers are not volatile and some of them are hardly soluble in common solvents and some decompose easily during heating.

The direct application of powerful analytical tools such as gas chromatography-mass spectroscopy (GC-MS) to most polymers and many complex materials is not feasible. Pyrolysis of these kinds of samples (polymers, composite organic materials) generates, in most cases, smaller molecules. The characterization of the pyrolytic fragments is performed by coupling the pyrolyzer to analytical methods such as gas chromatography, (Py-GC, Pyrolysis-Gas Chromatography), mass spectrometry (Py-MS, Pyrolysis-Mass Spectrometry), or more sophisticated integrated Py-GC-MS (Pyrolysis-Gas Chromatography-Mass Spectrometry). In these systems, the sample is heated to a temperature causing the rupture of the bonds and fragments are separated and recorded by an analyzer that provides the characteristic pyrolysis profile. This profile, called *pyrogram*, under suitable and controlled experimental conditions, can be considered a sort of finger-*print* of the substrate, both as regards the appearance of characteristic fragmentation products and for the distribution and relative concentration of these. It is an indirect characterization as it determines the composition of the substrate from the analysis of the fragments.

The advantages encountered by the application of the technique are numerous:

- Reduced time analysis
- No sample pre-treatment
- Small sample amount (lower than mg)
- Analysis of complex matrices

The major disadvantage is that the pyrolysis is a destructive technique, but this negative aspect is mitigated by the fact that only small amounts of sample are necessary: 0.1 mg for standard samples and 0.5 mg for others.

Analytical pyrolysis is frequently considered specific for polymers analysis, which may at first seem fairly limited. In fact, many compounds belong to this class, such as that proteins, polysaccharides, plastics, adhesives, paints, etc... natural and synthetic polymers, in the forms of textile fibers, wood products, foods, leather, paints, varnishes, plastic bottles and bags, and paper and cardboard, are materials that are part of daily life. Consequently, the study of these materials by pyrolysis has become a very broad field, including diverse topics such as soil nutrients, plastic recycling, criminal evidence, bacteria and fungi, fuel sources, oil paintings, and computer circuit boards.

A1-2 Degradation mechanisms

The pyrolysis products reflect the molecular structure, stability, free radicals, substitution and internal rearrangements of the polymers constituting the original sample. Identical molecules that undergo the same pyrolysis conditions degrade in the same characteristic way. The thermal degradation that a polymeric material undergoes when subjected to pyrolysis is characterized by breaking of chemical bonds and formation of free radicals. The way in which a molecule degrades depends on the type of bonds involved and on the stability of smaller fragments that are formed. These products, identified by GC-MS, can provide a fingerprint of the original polymer composition and microstructure and may help to determine the mechanisms of degradation.

The three main mechanisms include *random scission*, *side-group scission*, and *monomer reversion*.

Random scission involves the random breaking of the polymer's C-C bonded backbone as all the bonds are of equal strength, resulting in the formation of products including, alkanes, alkenes and alkadienes of smaller sizes. The polyolefins are good examples of materials that behave in this manner (Figure A1-1).

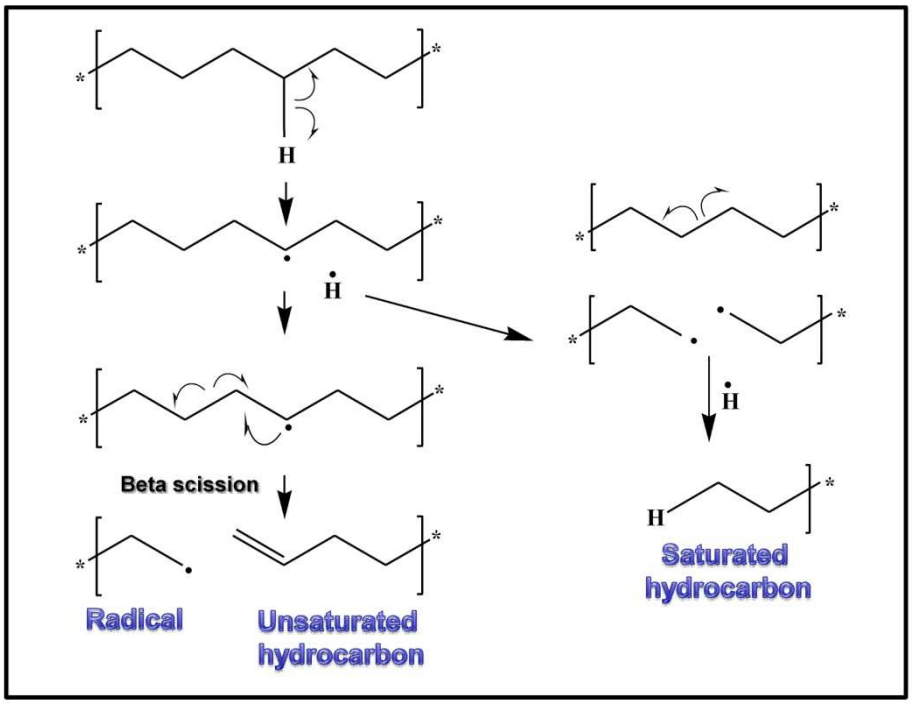

Figure A1-1 Random scission mechanism (from *Ref*. [3])

Chain scission produces hydrocarbons with terminal free radicals which may be stabilized in several ways. If the free radical extracts a hydrogen atom from a neighboring molecule, it becomes a saturated end and creates another free radical in the neighboring molecule which may be stabilized in a number of ways. The most probable way is beta scission, which accounts for most of the polymer backbone degradation by producing an unsaturated end and a new terminal free radical. This process continues producing hydrocarbon molecules that are saturated and have one terminal double bond or a double bond at each end. When analyzed by gas chromatography, the resulting pyrolytic profile presents a series of triplet of peaks corresponding to the diene, alkene, and alkane containing a specific number of carbons and eluting in that order.

Side-group scission occurs when the side groups attached to the backbone are broken away resulting in the backbone becoming polyunsaturated. An example of a polymer which undergoes this type of degradation is poly (vinyl chloride) (PVC). A two-step degradation mechanism begins with the elimination of HCl from the polymer chain leaving a polyunsaturated backbone that, upon further heating, produces the characteristic aromatics which can be individuated in the pyrogram (Figure A1-2).

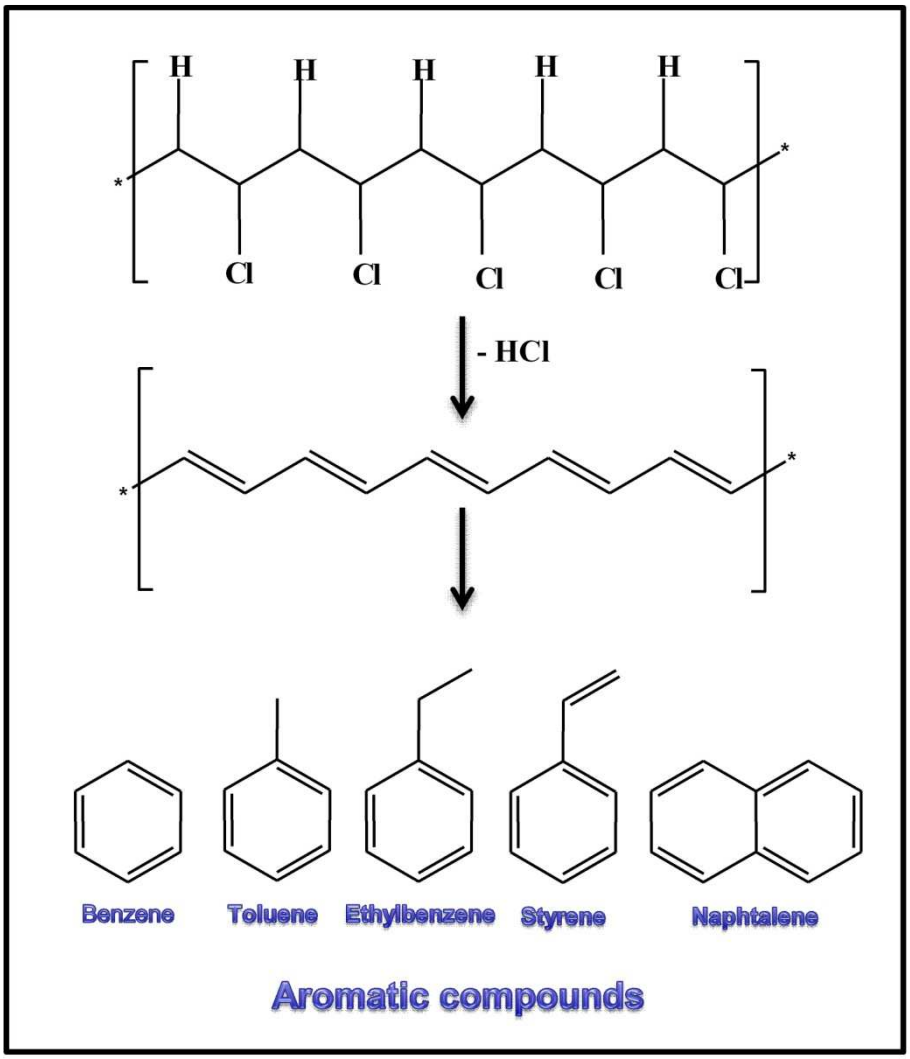

Figure A1-2 Side-group scission mechanism (from *Ref*. [3])

Monomer reversion results in unzipping and reverting back of the polymer to its original monomeric version. Polymers which are known to undergo this mechanism include polymethylmethacrylate polytetrafluoroethylene, poly-α-methylstyrene and polyoxymethylene. This proceeds in copolymers as well, with production of both monomers in roughly the original polymerization ratio (Figure A1-3).

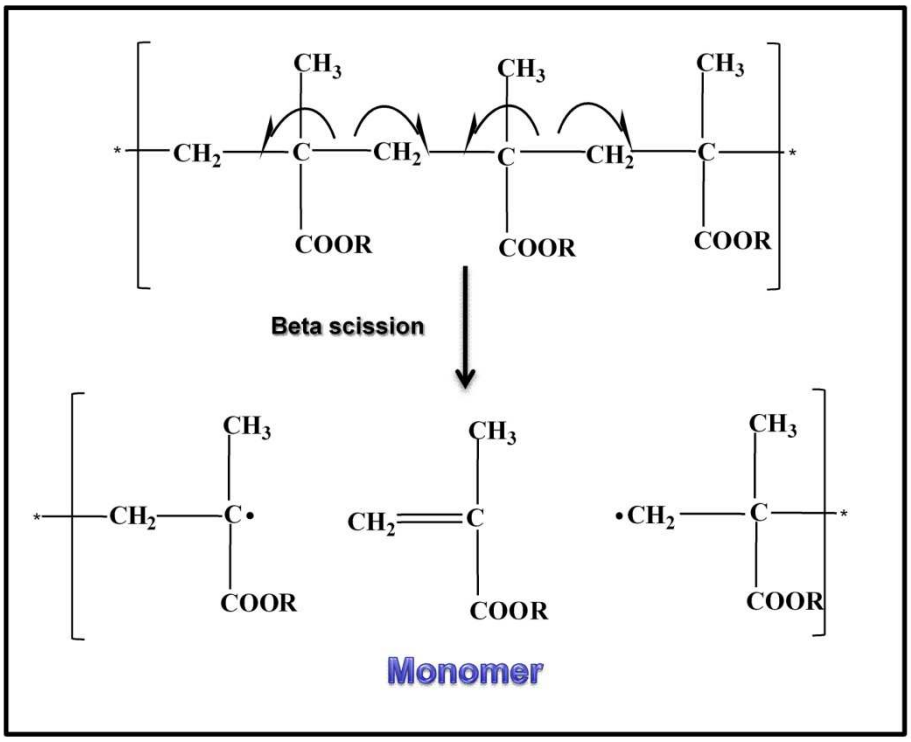

Figure A1-3 Monomer reversion mechanism (from *Ref*. [3])

Predicting the degradation mechanism of a polymer is not always immediate but is very useful. It can be achieved by three simple rules:

- Pyrolysis degradation mechanisms are free-radical processes and are initiated by breaking the weakest bonds
- The composition of the pyrolysate will be based on the stability of the free radicals involved and on the stabilities of the products.
- Free radical stability follows the usual order of order 3°>2°>1°>CH_3

A1-3 The integrated system Py-GC-MS

To perform an analysis by pyrolytic techniques, a system is required capable of heating small samples to pyrolysis temperatures in a reproducible way, interfaced to an instrument able to analyze the pyrolysis fragments produced. Pyrolysis-gas chromatography/ mass spectrometry is the integrated system mostly used to perform analytical pyrolysis.

In Figure A1-4 the integrated system Py-GC-MS used in this work is reported:

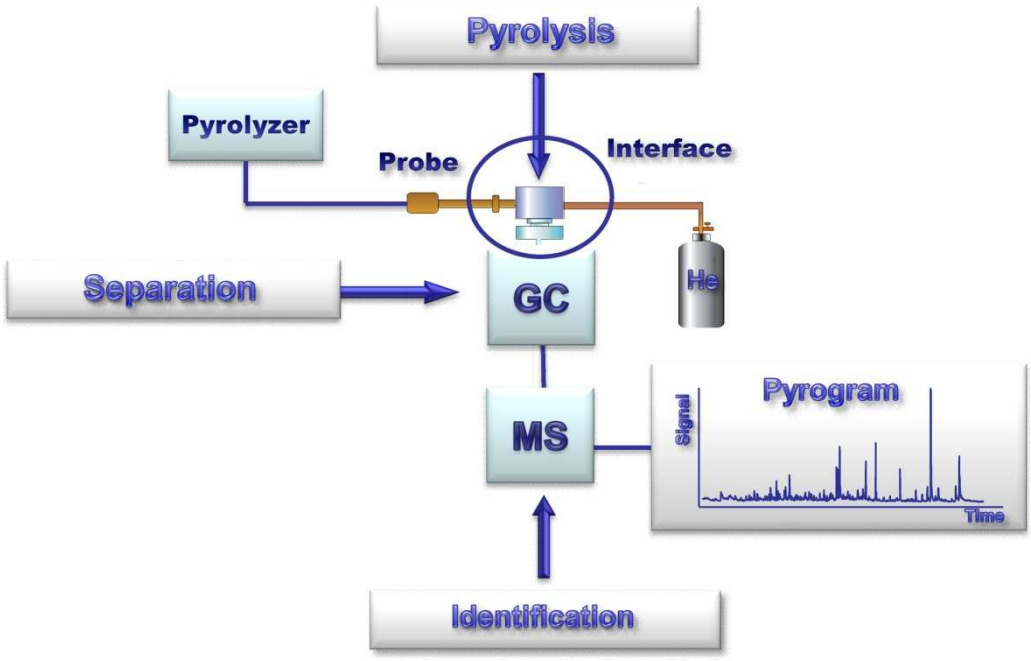

Figure A1-4 the integrated system Py-GC-MS

The pyrolyzer, a heated filament pyrolyzer, can be considered as a sample introduction device; it's directly connected to the injection port of a gas chromatograph and the pyrolysis fragments are sent with the flowing gas to the chromatographic column. Indeed, a flow of inert gas (helium) crosses the whole system ensuring an inert atmosphere in the pyrolyzer interface and acting as a carrier gas for the gas chromatographic separation of pyrolysis fragments (Figure A1-5).

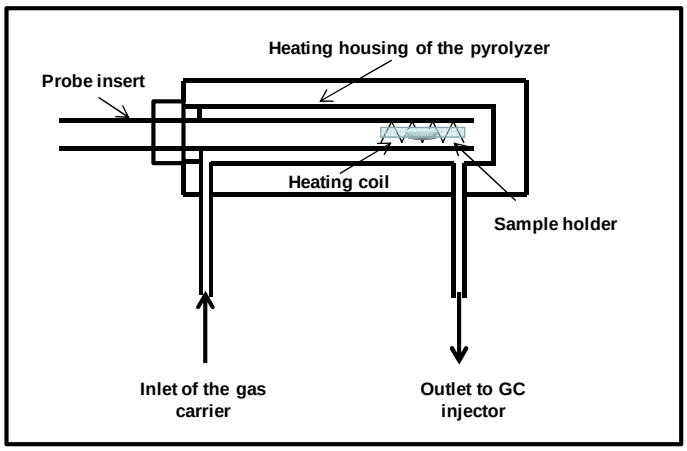

Figure A1-5 the simplified scheme of a pyrolyzer (by CDS Inc.)

The pyrolyzer is constituted by a control unit connected to a cylindrical metal probe, to whose end a platinum coil resistance is mounted; inside the coil a capillary of quartz containing the sample is introduced, closed at the ends by quartz wool. The so obtained probe is inserted within the metal interface, which is mounted directly on the injector of the gas chromatograph, which is kept heated to an appropriate temperature (Figure A1-6).

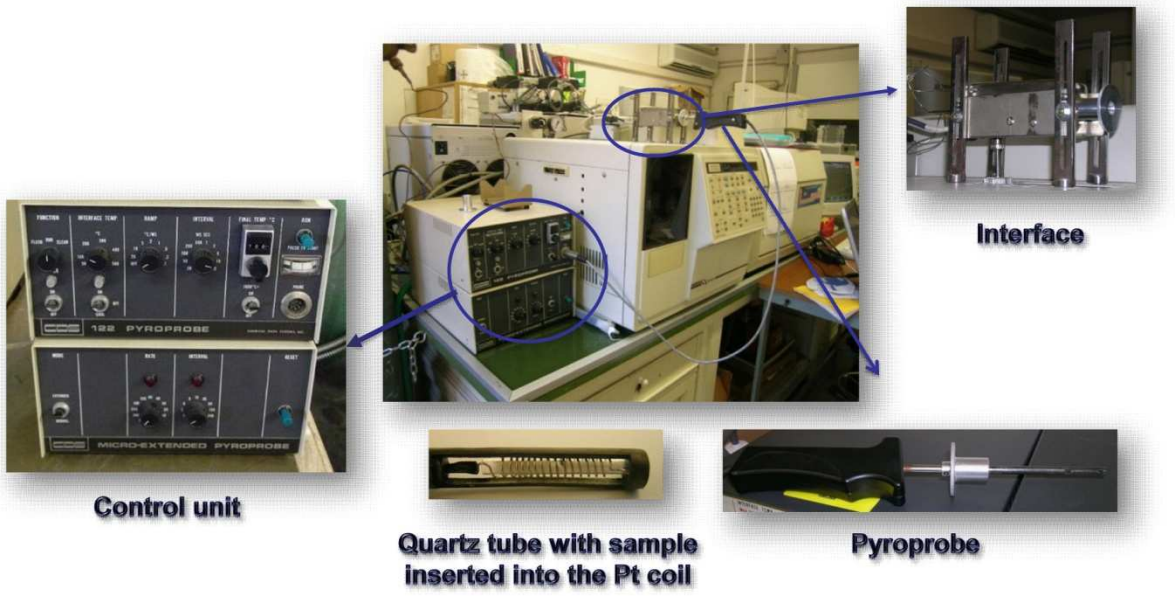

Figure A1-6 Description of pyrolyzer

The parameters used in the pyrolyzer must be set taking into account both the chromatographic performance and the proper sample degradation.

The control unit of the pyrolyzer is able to provide a strong electric pulse to the platinum resistance, which thus reaches the programmed temperature pyrolysis (TPY, Pyrolysis Temperature). The filament temperature may be monitored using the resistance of the material itself or some external measure, such as a thermocouple. The time taken by the pyrolyzer to reach the TPY set (TRT, Temperature Rise Time) must be extremely fast, with respect to the degradation temperature of the sample (*flash pyrolysis*). If the temperature raises too slowly the sample may pyrolyzes before reaching the programmed TPY and the primary fragments formed can, subjected to heating, recombine in a non-reproducible way (Figure A1-7).

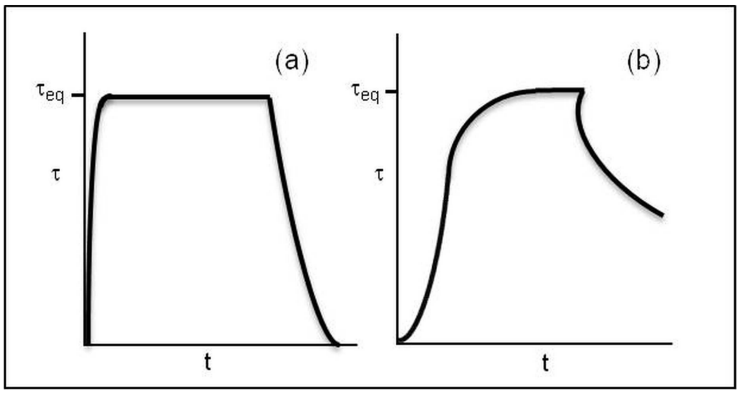

Figure A1-7 TRT, temperature rise time a) Ideal condition; b) Poor condition for instrumental performance

As for gas chromatographic conditions, the pyrolysate pass through the injector when already volatilized; if the sample is heated slowly, or to a temperature at which degradation is slow, the volatiles will travel to the column over a finite time, which may produce peaks too broad to achieve the needed chromatographic resolution.

Pyrolysis temperature is maintained for a programmable short time interval (usually 10 s). Its value must be maintained between 400 ° C and 900 ° C since, above this range of values, fragments could be formed which are too small and not significant enough to characterize the original sample. The final distribution of pyrolysis fragments, or *finger-print*, is strongly influenced by pyrolysis temperature, which strongly influences the fragmentation reactions. The parameters TPY and TTP are that the ones mostly affecting reproducibility, together with the interface temperature (T_{int}). This temperature is kept high and constant (about 250 °C) to limit as much as possible any phenomena of condensation of pyrolytic fragments and to maintain the latter in the gaseous state to allow the transport by the mobile phase gas (carrier gas) through the gas chromatographic column. The interface should have a small internal volume to avoid fragments spending too much time in the zones not covered by gas flow, to reduce the contact between analytes and hot surfaces and to transfer pyrolytic fragments away from the high temperature zone (where secondary pyrolysis could take place) and into the column efficiently. The thermal fragmentation, inside the interface, takes place in an inert atmosphere (helium) to avoid the presence of oxygen, which would cause secondary reactions of oxidation and / or combustion.

It is appropriate to use a reasonable carrier flow (60 ml min^{-1}): the greater the time past in the high temperature zone, the greater the possibility that secondary pyrolysis reactions take place.

The transfer zone between pyrolyzer and the GC injection port should be kept hot to prevent condensation, and as short as possible to reduce surface area and volume.

The sample should be very small to prevent phenomena such as the overloading of the pyrolyzer and analytical device, causing contamination and carryover into the next analysis. Small amounts of sample are degraded more rapidly and completely, ensuring good resolution and reproducibility. When it is not possible to limit the amount of sample, the use of an injection port with a splitter and a relatively high split ratio are suggested.

Sample preparation must be carefully performed, taking into account size and shape, homogeneity and contamination. If a sample material is soluble, a few microliters of the solution may be injected directly into the quartz wool inserted in the capillary, using a syringe. Usually, analytical pyrolysis is used for direct analysis of solid samples: this is surely its strength. But a very important issue that emerges when solid samples of a few micrograms are analyzed is how they are representative of the material from which they were taken. It is very common to handle non-homogeneous materials, natural or synthetic, and some precautions have to be taken not to compromise the reproducibility of the analysis:

- reducing the sample to a fine powder, and take small amounts of this
- dissolving the sample when its components are soluble
- analyzing an amount of sample as great as possible

In the last case, the use of a splitter with a large split ratio it's suggested to limit the amount of the pyrolysate entering the analytical device. These are several ways to obtain a representative sample and ensure the reproducibility of the pyrolytic analysis.

In GC-MS, the gas flow-rate is commonly set between 0.1 to 3.0 mL/min (at near atmospheric pressure). Because the mass spectrometer operates at 10^{-5} to 10^{-7} mbar, it is equipped with efficient pumps, which maintain high vacuum inside.

The mass spectrometric analysis starts with an ionization process that takes place in the ion source of the MS instrument, where the analyte is introduced as gas phase. The ionization procedure used in this work is the electron ionization (El) that consists of an electron

bombardment, which is commonly done with electrons having energy of 70 eV. The electrons are usually generated by thermoionic effect from a heated filament and accelerated to the required energy. When the molecular ion obtains too high energy during the electron impact, there are fragments that are generated in the ion source.

The mass spectrometer used is equipped with an ion-trap where the ion separation based on their mass to charge ratio (*m/z*) takes places. After separation, ions are detected using an electron multiplier which detects the arrival of all ions sequentially at one point. The signal from the detector is amplified and interpreted by an electronic data system.

The nature and abundance of fragments is characteristic for a given compound; the fragment abundance represented versus *m/z* (mass/charge) generates a mass spectrum. Usually, the abundance is normalized to the most abundant ion (base peak expressed as 100%), and the mass peaks are shown as bar graphs (the real peaks in a mass spectrum have ideally a Gaussian shape). The mass spectrum can be used as a fingerprint, leading to the identification of the molecular species that generated it. The fragmentation (when done in standard MS conditions) generates typical patterns that allow the identification of each compound, either based on interpretation rules or by matching it with standard spectra in mass spectral libraries such NIST (National Institute of Standards and Technology).

When the mass spectrometer is used as a detector during the chromatographic process in scan mode, the chromatogram can be displayed as a total ion chromatogram (TIC) or in a single ion chromatogram (SIM). A total ion chromatogram is a plot of the total ion count (detected and processed by the data system) as a function of time. The single ion chromatogram plots the intensity of one ion (*m/z* value) as a function of time. These chromatograms have a discrete structure being made from scans (the scan number is linearly dependent of time). When the points of the chromatogram are close to each other, this gives a continuous aspect of the graph.

The information obtained is qualitative, thanks to the retention times of the chromatographic peaks and to the study of the relative mass spectra, and quantitative thanks to the proportionality between peak area and amount of the compound.

A1-4 Derivatization

In analytical pyrolysis, the use of derivatization increases the potential of the technique for the characterization of complex matrices, consisting of very polar and therefore low volatile compounds. Derivatization determines an improvement of the resolution, *i.e.* an increase of the degree of separation of peaks in the pyrogram, thanks to the best chromatographic characteristics of derivatives. The use of a derivatizing agent also allows to detect the presence of compounds in a sample that otherwise would be underestimated or completely ignored. Following the derivatization, that determines the replacement by chemical reactions of the functional groups, the chemical structure of a compound changes as well as the fragmentation pattern which is thus more significant.

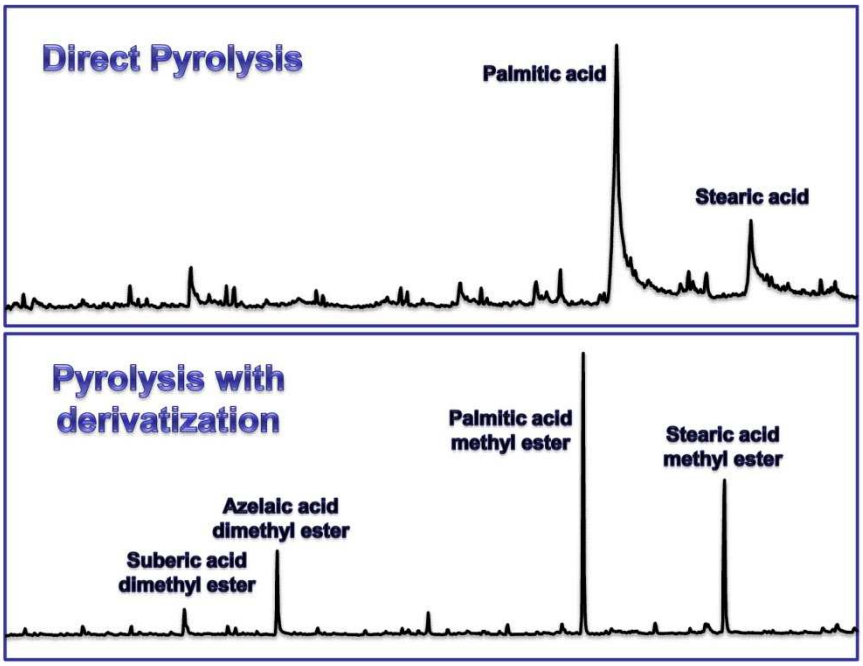

Figure A1-8 Pyrograms of a standard painting layer

In Figure A1-8 the pyrograms obtained by the analysis of a standard painting layer, containing siccative oil, are reported. Direct pyrolysis gives a poor chromatographic profile and does not allow to detect the presence of two important compounds for the characterization of a siccative oil, azelaic and suberic acid. The use of derivatizing completely changes the chromatographic profile: the resolution and the degree of separation are improved and the peaks corresponding to the siccative oil markers are well defined.

Performing the derivatization it's a simple operation: the derivatizing reagent is directly added to the sample to be analyzed, within the quartz capillary ("*in situ*"). In this way, all the

preliminary steps of chemical treatment of the sample are eliminated, with significant time saving, reduced error margin and avoiding loss of sample.
Some methods of derivatization are available:

ACYLATION: The acylation reaction reduces the polarity of the amino groups, hydroxyl and thiol; is used for highly polar compounds, such as carbohydrates and amino acids.

METHYLATION: The most used derivatization reaction is methylation of the carboxyl and hydroxyl groups, with the formation of esters and ethers [5]. The main methylating agent used in the analytical pyrolysis is the TMAH (tetramethylammonium hydroxide), that has long been used in gas chromatography. Its first application in analytical pyrolysis is due to J. M. Challinor which introduced the *Simultaneous Pyrolysis and Methylation* (SPM), a derivatization reaction, performed *in situ,* in controlled heating conditions. This technique has become one of the most important methods to easily determine the chemical composition of various types of condensation polymers and esters, without waste of time in long pre-treatment.
The mechanism provides, simultaneously with the pyrolysis, a direct methylation of the fragments by heating the reagent. The reaction mechanism suggested by Challinor includes a first thermal hydrolysis of the ester bond or ether of the original molecule by highly basic alkylating agent such as TMAH which leads to the formation of salts (STEP 1). In conditions of high temperature (STEP 2), a nucleophilic attack to carboxylate and alcoholate anions by the methyl groups of the tetramethylammonium cation leads to the formation of methyl derivatives:

STEP 1

$$[(CH_3)_4N]^+OH^- + RCOOR_1 \rightarrow RCOO^-[(CH_3)_4N]^+ + R_1OH$$

STEP 2

$$RCOO^-[(CH_3)_4N]^+ \xrightarrow{\Delta} RCOOCH_3 + N(CH_3)_3$$

$$R_1OH + [(CH_3)_4N]^+OH^- \leftrightarrows R_1O^-[(CH_3)_4N]^+ + H_2O$$

$$R_1O^-[(CH_3)_4N]^+ \xrightarrow{\Delta} R_1OCH_3 + N(CH_3)_3$$

SILYLATION: The silylation reaction is a nucleophilic SN2 substitution reaction. Silylated derivatives are formed when active hydrogen is replaced with an alkylsilyl group. The reactivity of the functional group to the silylation follows the order:

Alcohols (1°>2°>3°)>Phenols> Carboxyl> Amines> Amides

The mainly used silylating reagent in analytical pyrolysis is HMDS (hexamethyldisilazane [6-8]. This reagent, long used as a scavenger in the gas chromatographic derivatization reactions, has been fairly recently used in analytical pyrolysis. Pyrolysis with HMDS leads to formation of trimethyl-silyl-derivatives, through a mechanism of simultaneous silylation to pyrolysis which is not yet clarified. It is hypothesized that the silylation of the sample takes place thanks to the thermal fragmentation and simultaneously to it. This technique is called SPS (*Simultaneous Pyrolysis and Silylation*).

There are several advantages in performing analytical pyrolysis in presence of a derivatizing reagent:

• increase the volatility of pyrolytic fragments; this leads to greater efficiency by gas chromatography with increasing symmetry and peak resolution of pyrogram

• pyrolysis and derivatization occur at the same time; the addition of reagent is made directly in the capillary just before pyrolyze, with an evident saving of time and analyte.

• Mass spectra result more significant; this allows easy recognition of the molecular ion (usually protonated) and the characteristic peaks.

CHAPTER A2 MULTIVARIATE DATA ANALYSIS

A2-1 Introduction to chemometrics

Providing immediate responses to problems of high complexity has become a requirement in many fields, not only scientific but also political and sociological. The inherent difficulty in dealing with a complex problem resides in the fact that it depends on many factors or *variables* that must be observed, studied and simultaneously measured.

Chemometrics [9] arises from the need to solve complex problems using tools from very different fields of knowledge, such as statistics, computer science, mathematics, experimental sciences. The origin of this scientific discipline is due to the joint initiative of Svante Wold, Umea University (Sweden), and Bruce Kowalski, University of Seattle (WA, USA), who in a letter to the Journal Analytical Chemistry in June 1974 proposed *chemometrics* to define an area of science that would include all those mathematical techniques aimed at treating, processing and modeling of chemical data sets. From the beginning, chemometrics has tried to collect all the mathematical and statistical tools most appropriate to treat this type of data. It is used in all the problems that have some element of complexity, such as the analysis of environmental matrices, the study of the pharmacological activity of a molecule, the processes of industrial production and forensic investigation.

For a complex problem it is almost impossible to develop a well-defined theory allowing a direct solution (hard model): it is necessary to use a different approach (soft model). The variables are numerous, they may be not defined with good accuracy and their exact relevance to the problem may be known. Also, the experimental noise masks and confuses the true effects of the considered variables.

Correlations between variables and the effects due to the combined role of two or more variables are difficult to define [10, 11]. In many cases the complexity of the problem can cause non–linear effects that can't be predicted *a priori*. It is therefore clear that system complexity greatly influences the inherent complexity of the system data that can be thought as a sum of two parts: data structure and noise (Figure A2-1).

Figure A2-1

Data structure is the signal part correlated to the property of interest while the *noise* is everything else; every observation is always a sum of both of these parts, and the structure part will be hidden in the raw data. It's not immediate to understand what should be kept and what discarded.

Chemometric methods aim to extract useful information from data. First, it is essential that measured data contain meaningful information about the interested property. The amount of information in a dataset depends on how well the problem has been defined. Also, there must be a quantitative relationship between the set of measured variables and the property of interest.

The study of information contained in experimental data may be achieved by:

- Data description (explorative data structure modeling)
- Discrimination and classification
- Regression and quantitative calibration

Assuming to have a data set consisting of *n* objects, each of them described by *p* variables, a data exploration can be performed to detect some information about statistical parameters relative to each variable, the correlations between the variables, and the presence of outliers. In particular, the *principal component analysis* (PCA) is one of the best methods for data exploration in multivariate systems.

Discrimination deals with the separation of groups of data while classification requires *a priori* class description. In this case the aim of the data analysis is to assign, to classify, new objects to the defined classes. Here also PCA can be used but in the experimental problem here faced the most useful chemometric classification approach is the Soft Independent Modeling of Class Analogy (SIMCA).

The *methods of regression* calculate mathematical models allowing to predict quantitative values of a variable (*response*) from the values known for *n* samples and taken from a set of independent variables (*predictors*). If it is considered that the model can assume a character of general validity, it becomes independent from the specific set of data used. The most used regression methods are: Principal Component Regression (PCR) and the Partial Least Square Regression (PLS-R).

A2-2 Multivariate structure data

Chemometric techniques normally apply to data structures represented by a table of numbers (*data matrix*) consisting of a number of *observations*, each of which is represented by *variables* that describe the observations (Figure A2-2). The data matrix is indicated with the symbol **X** and represented in this form:

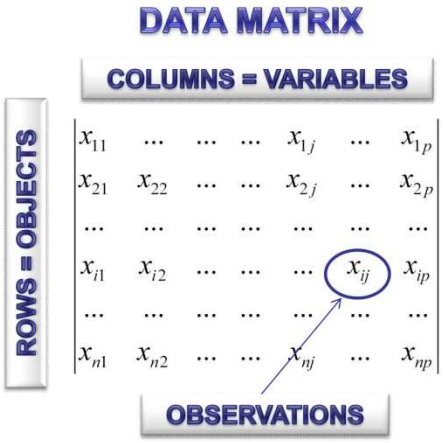

Figure A2-2 Data matrix

The *n* rows of **X** are *objects* and can be *samples, experiments*.

The *p* columns of **X** are the *variables* and can be: *properties*.

An element of X is written$x_{ij,}$ with $i = 1, ... n$ and $j = 1, ..., p$.

The *variables* or *descriptors* are the quantities used to study a given phenomenon and to describe the overall experimental observations and can be measured or calculated. The variables may be: independent variables or *predictors* and dependent variables or *responses*. It is essential to highlight the possible existence of variables that, although having different meanings, carry the same information or have the same ability to quantitatively represent the system under study. To highlight the roles of variables, we rewrite the matrix data in the form (Figure A2-3):

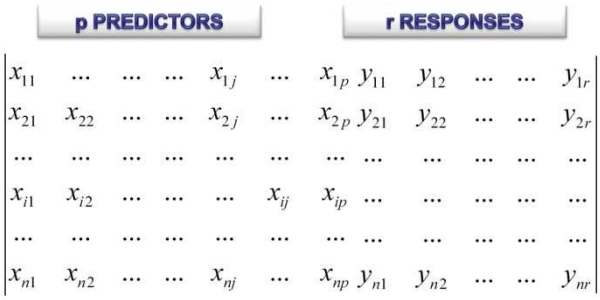

Figure A2-3 Data matrix

The *objects* are examples or samples available to understand the studied phenomenon, to construct mathematical models, to confirm the assumptions made.

An object can be described in full if, for all the variables selected to describe it, there are values that apply to it. In other cases, as frequently happens, no data are available for all variables selected and the object description is incomplete: it is said that there are *missing values* .

A2-3 Data pretreatment

Before performing any kind of processing a preliminary analysis of the data is always necessary, aiming to control and prepare the data for subsequent processing. In checking data, it is suggested to ensure absence of transcription errors, occurrence of missing data and choice of correct orders of magnitude.

The fundamental pretreatments are: *transformation of variables*, *data scaling*, and *replacement of missing data*.

A preliminary check on the variables consists of the following tests:

• there not should be *constant variables*: they would have no variance

• there not should be *degenerate variables:* they vary little with changes of objects and give variance tending to zero

• there not should be *discrete variables*: they require a particular statistical

We proceed to the transformation of variables to control abnormalities such as:

• lack of linearity

• lack of normality

• lack of additivity

• the variables have variances of very different orders of magnitude

• the system is not very strong

The most common transformations are reported in Table A2-I.

Table A2-I

LOGARITHMIC TRANSFORMATION	$x'_{ij} = \log(1 + x_{ij})$
POWER TRANSFORMATION	$x'_{ij} = x_{ij}^{\lambda}$
INVERSE TRANSFORMATION	$x'_{ij} = x_{ij}^{-1}$
SQUARE ROOT TRANSFORMATION	$x'_{ij} = x_{ij}^{0.5}$

When searching for information about relationships between variables, it is important to maximize the comparability between the variables. A difference of several orders of magnitude on average values result in several orders of magnitude of difference on the variances and therefore a statistical comparison is not possible. In these cases it is appropriate to apply a *data scaling* data can be performed in different ways. The most common are:

• Centering (CS)
• Maximum scaling (MS)
• Range scaling (RS)
• Autoscaling (AS)

Table A2-II

CENTERING, CS	$x'_{ij} = x_{ij} - \bar{x_j}$
MAXIMUM SCALING, MS	$x'_{ij} = x_{ij}/U_j$
RANGE SCALING, RS	$x'_{ij} = (x_{ij} - L_j)/(U_j - L_j)$
AUTOSCALING, AS	$x'_{ij} = (x_{ij} - \bar{x_j})/s_j$
$\boldsymbol{U_j}$ = maximum value *j*-th column; $\boldsymbol{L_j}$ =minimum value *j*-th column; $\boldsymbol{s_j}$ = standard deviation	

In many cases it happens that no samples are available for all the values for all variables that describe them. This unpleasant aspect of the data forces us to take some decisions, because none of the common mathematical methods can deal with problems where some values are missing in the data matrix. To solve this problem there are several possibilities:

- **Removal of samples**

The simplest method uses the removal of samples for which there is data missing. This is a good solution only if the number of samples is high and therefore the elimination of some sample does not entail the loss of relevant information.

- **Elimination of variables**

An alternative to the previous case is the elimination of one or more variables for which there are a number of missing data, and then retaining only those variables for which all values are available.

- **Substitution with the average**

If missing data are not too many, the missing value can be replaced with the mean value calculated over all the remaining data relevant to the involved variable.

- **Replacement by a random value**

The missing data are replaced by a random number such as -999.

- **Replacing using regression**

The missing values for each variable are predicted using a regression model derived from all samples with no missing values.

- **Replacing using local similarity**

This method for the calculation of missing values is based on the method of classification K-NN (k-nearest neighbors).

- **Replacing using the principal component analysis**

The principal component analysis is performed using all the known values and applying an algorithm that allows to calculate the main components even with incomplete matrices.

A2-4 Principal component analysis

A2-4.1 Introduction

The principal component analysis (PCA) is one of the fundamental techniques for multivariate analysis. It was introduced by Karl Pearson in 1901 and developed in its present form in 1933 by Harold Hotelling. It is the most important *exploration data* technique and it consists in transforming the original variables into new variables, called *principal components* (or *latent variables*), obtained by linear combination of the original variables and orthogonal to each other.

The PCA allows to:

- evaluate correlations between variables and their relevance
- display objects (identification of outliers, classes, etc.)
- summarize the data description (removal of noise or spurious information)
- reduce the data dimensionality
- individuate major properties
- develop a model of data representation in an orthogonal space

A2-4.2 The principal components

The calculation of the principal components is divided into several steps that lead to a matrix **A**, matrix of eigenvectors,

$$\mathbf{Av} = \lambda\mathbf{v} \qquad A2\text{-}1$$

where λ is an eigenvalue and is related to the *explained variance*.

The first step consists in the calculation of the *correlation matrix* **C** (*p x p*), defined as:

$$\mathbf{C} = \frac{\mathbf{X_{AS}^T X_{AS}}}{\text{n-1}} \qquad A2\text{-}2$$

where:

- **X** is the data matrix (*n x p*)
- $\mathbf{X_{AS}}$ is the *autoscalated* data matrix (*n x p*), $x'_{ij}(AS) = \frac{x_{ij} - \overline{x_j}}{s_j}$; $\overline{x_j} = \frac{\sum_{i=1}^{n} x_{ij}}{n}$;

$$s_j = \sqrt{\frac{\sum_{i=1}^{n}(x_{ij}-\overline{x_j})^2}{(n-1)}}$$

The correlation matrix compares variables with each other; the term c_{ij} quantifies the correlation between variables i and j:

- $c_{ij} > 0$ means there is a positive relationship between the variables
- $c_{ij} < 0$ means there is a negative relationship between the variables
- $c_{ij} = 0$ means there is no correlation

The *diagonalization* of the correlation matrix C results in the *eigenvalues* matrix $\mathbf{\Lambda}$ (*p x p*) whose diagonal elements are the *eigenvalues* λ_m (*m= 1,….,p*):

$$\mathbf{\Lambda} = diag\ \mathbf{C} \qquad \textit{A2-3}$$

By the process called Single Value Decomposition (SVD) which is based on the following equation:

$$\mathbf{C} = \mathbf{A}\,\mathbf{\Lambda}\,\mathbf{A^T} \qquad \textit{A2-4}$$

the matrix of the *eigenvectors* **A** is determined:

$$\mathbf{A} = \begin{bmatrix} l_{11} & \cdots & \boxed{l_{1m}} & \cdots & l_{1p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ l_{j1} & \cdots & l_{jm} & \cdots & l_{jp} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ l_{p1} & \cdots & l_{pm} & \cdots & l_{pp} \end{bmatrix}$$

λ_m

The equation A2-4 corresponds to the operation of rotation that defines a new space, called *eigenspace*, whose axes, the eigenvectors, are oriented in the direction of maximum variance, in descending order. Each eigenvalue is proportional to the variance associated with the corresponding principal component and the sum of all the eigenvalues is equal to the total

variance of the data. Considering the matrix **A** and assuming that the smallest eigenvalues are associated with not relevant information, the eigenvalues are examined and the first M largest eigenvalues retained. The matrix obtained with the first M eigenvectors is denoted by **L** (*p x M)* and is called the *loadings matrix,* whose columns represent the PC and whose rows represent original variables.

Loadings are standardized linear coefficients, *i.e*. the sum of squares of the loadings of an eigenvector is equal to 1, or the eigenvectors have unit variance; consequently the relations:

$$-1 \leq l_{jm} \leq +1 \qquad \sum_j l_{jm}^2 = 1 \qquad \textit{A2-5}$$

A value of l_{jm} close to 1, in absolute value, indicates that the *m*-th component is represented mostly by the *j*-th original variable; vice versa, a value of l_{jm} close to zero indicates that the *j*-th variable is not represented in the *m*-th component.

Multiplying together the data matrix **X** and the loadings matrix L , a new matrix **T** called *scores matrix* is determined :

$$\mathbf{T} = \mathbf{XL} \qquad \textit{A2-6}$$

Scores values are the result of a linear combination, in which the variables are the original variables (usually scaled) and the coefficients are the *loadings* of the *m*-th component:

$$t_{im} = \sum_j x_{ij} \boldsymbol{l}_{jm} \qquad \textit{A2-7}$$

Thus:

$$t_{im} = \mathbf{x}_i^T . \boldsymbol{l}_m \qquad \textit{A2-8}$$

where $\mathbf{x}^{\mathbf{T}}$ and $\boldsymbol{l}$ are *p*-dimensional vectors being *p* the number of original variables. Unlike loadings, scores have mean value equal to zero, but can take any numerical values. The scores represent the new coordinates of the objects in the space of principal components.

Applying the inverse formula:

$$\widehat{\boldsymbol{X}} = \mathbf{TL}^{\mathbf{T}} \qquad \textit{A2-9}$$

the *data reproduced* matrix is obtained: if the number of principal components is less than the original variables, the new matrix is an approximation of the original one, as spurious information (*i.e.* the noise) has been eliminated.

A2-4.3 Loading and score plots

In each multivariate analysis is important to display the results obtained by simple and intuitive plots allowing an immediate graphical interpretation.

PCA provides an algebraic solution that determines very effective graphical representations both of objects (*score plot*) and of variables (*loading plot*).

The *loading plot* allows analyzing the role of each variable in the different components, their direct and inverse correlations, and their importance (Figure A2-4).

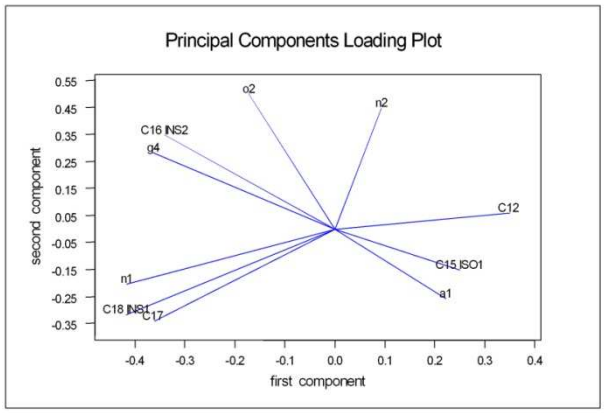

Figure A2-4 Example of a *loading plot*

Consider the simplest case and choose only two principal components, PC1 and PC2, which are the two columns of the matrix of loadings **L** (*p x 2*). This matrix has *m* rows, which correspond to the original variables or to the columns of the data matrix **X**. The elements of **L** vary between -1 and 1. Each variable is characterized by a pair of loadings, l_{j1} and l_{j2}:

- l_{j1} quantifies the relevance component of the variable x_j in PC1
- l_{j2} quantifies the relevance of the variable x_j in component PC2.

The *loading* plot is constructed by representing, in the plane l_{j2} vs. l_{j1}, the points corresponding to the *p* original variables. The interpretation of the loading plot is based on some fundamental observations.

- Points near the origin correspond to variables not relevant to any principal components choices.
- Points on the horizontal axis correspond to variables that are irrelevant for PC2.
- Points on the ordinate correspond to variables that are irrelevant for PC1.
- Points with abscissa of absolute values are very important for high PC1.

- Points ordered with high absolute value are very important for PC2.
- Points close to each other correspond to variables that carry similar information.
- Points symmetrical to the origin correspond to varying inversely related.

The *score plot* shows the behavior of the objects in the different components and their similarity (Figure A2-5).

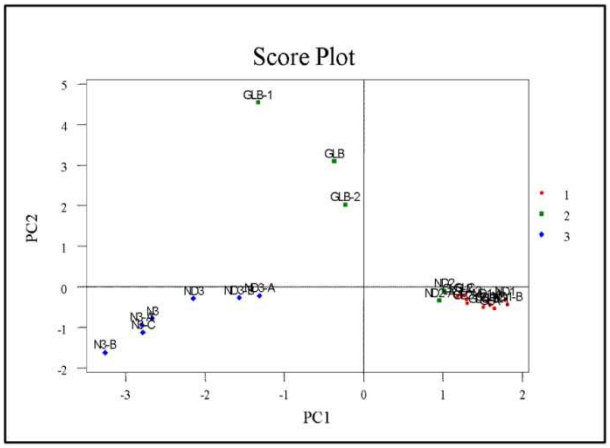

Figure A2-5 Example of a *score plot*

As in the previous example, only two principal components PC1 and PC2, which are the first two columns of the matrix of loadings **L**, have to be considered. The product of **X** and **L** gives the matrix **T** (*n x 2*). This matrix has *n* rows, which correspond to objects or to the rows of the data matrix **X**. Each object is characterized by a pair of scores, t_{j1} and t_{j2} that are the coordinates in the space of the components PC1 and PC2. The most commonly used plot in multivariate analysis is the score vector for PC2 versus the score vector PC1. This is to due to the fact that these are the two directions of largest and second largest variances.

The *score plot* is constructed in the plane t_{j2} vs. t_{j1} reporting the *n* points corresponding to the *n* objects of the problem. There are two rules of thumb concerning score plots:

- Always use the same couple of principal component in all score plots: this will help getting the desired overview of the compound structure data.
- Choose as ordinate the principal component that has the largest "problem relevant" variance. For many applications this will turn out to be PC2, but it is possible that, in other cases, PC2 lies along a direction that for some problem-specific is not interesting.

The *score plot* represents the objects in the space of components principal and is very effective to search for *outliers* and *clusters* graphically. The comparison of score plots and loading plots allows determining which variables are relevant to the *n* objects. There is a "quadrants" correspondence between the two graphs (Figure A2-6):

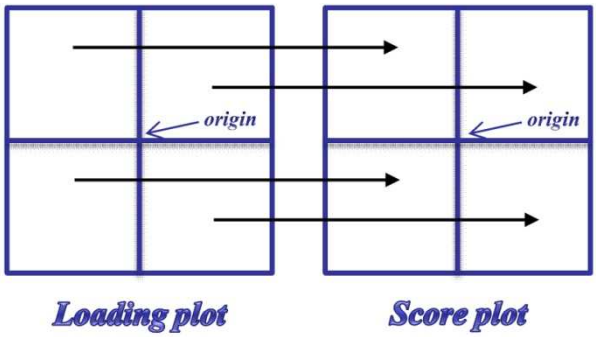

Figure A2-6 Comparison between *loading plot* and *score plot*

A2-4.4 Geometric interpretation of PCA

Geometrically, the PCA appears as a *rotation* from the original space to the space of principal components; this process is performed in such a way that the first new axis, *i.e.* the *first principal component*, PC1, is oriented in the direction of *maximum variance* data and all the next axes are *orthogonal* to each other.

Consider for simplicity the case of a two-dimensional system; the variables space will have 2 axes: *var1* and *var2*. Each object is represented by a set of variable measurement and characterized by its coordinate (x_1, x_2) in the Cartesian coordinate system (Figure A2-7).

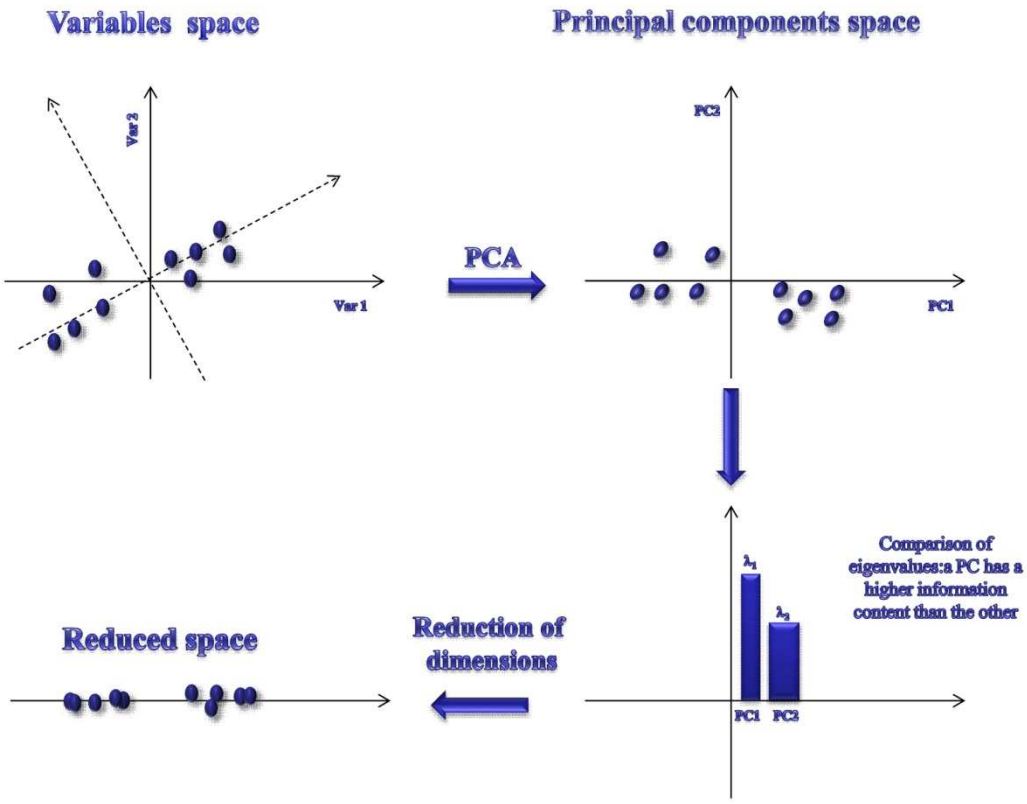

Figure A2-7 Geometric interpretation of PCA

Each object can be plotted as a *point*, and the result is a swarm of points showing a trend of the objects to arrange in a particular direction. A central axis could be drawn through the swarm: this line would describe the data almost as efficiently as all the original variables. This central axis is positioned along the direction of *maximum variance* and is called *first Principal Component* or PC1. The second principal component will lie along a direction orthogonal to first PC and in the direction of the *second largest* variance.

One of the purposes of PCA is to reduce the dimensionality of a multivariable system: a transformation from two dimensions to only one may seem trivial but it allows to understand the true meaning of this method of exploration. If it is assumed to be able to represent the points with the only new coordinate PC1 is right to assess the extent of this approximation. This means verifying how much information is lost. In Figure A2-7 a comparison between the eigenvalues associated to the principal components PC1 and PC2 is reported.

It's evident that $\lambda_1 > \lambda_2$ and since the variance (and therefore the information) is related to this parameter, a dimensions reduction performed by a projection of the objects on the axis PC1 does not cause a significant loss of information.

The example illustrated above can easily be generalized to a *p*-dimensional system, determining the principal components PC1, PC2, PC3 etc. The first principal component contains the highest percentage of variance, the second a smaller percentage and so on, until the last components bring with them a negligible amount of variance.

A2-4.5 *The number of significant components*

One of the fundamental problems that the principal component analysis presents is the determination of the significant number *M* of components (factors), with $M < p$. The search for the number of significant components is known as *rank analysis*. In fact, if data contains an information structure (*i.e.*, non random), separation between the variability due to experimental noise or spurious information and the useful information is achieved through a proper definition of the number of significant principal components. Any selection procedure of a reduced number of significant components assumes that the variability of useful information is greater than the variability associated with the experimental noise or secondary information.

Generally two methods, a numerical and a graphic one, are used.

The numerical method is based on the concept of *percentage explained variance* by *M* principal components, which represents an estimate of the importance of information derived from a certain principal component. Since each eigenvalue represents the variance associated with the corresponding principal component and the sum of all the eigenvalues coincides with the total variance, the *percentage explained variance* for the *m* component compared to the total variance is given by:

$$EV_m(\%) = \frac{\lambda_m}{\sum_{m=1}^{p} \lambda_m} \times 100 \qquad \textit{A2-10}$$

The *cumulative explained variance* relative to the first *M* principal components is given by:

$$Cum.EV_M(\%) = \frac{\sum_{m=1}^{M} \lambda_m}{\sum_{m=1}^{p} \lambda_m} \times 100 \qquad \textit{A2-11}$$

The *unexplained residual variance* relative to the first *M* principal component is given by :

$$RV_M(\%) = \frac{\sum_{m=M+1}^{p} \lambda_m}{\sum_{m=1}^{p} \lambda_m} \times 100 \qquad \textit{A2-12}$$

The right number of PCs would be the one corresponding to a cumulative variance greater than a given threshold value; a common choice of thumb is to consider as a threshold the 80% of the total variance. What is important is to minimize the value of RV to be sure not to lose too much information.

The number of principal components to be considered can be determined by the analysis of the eigenvalues graphics plotted against the number of factors. In this type of graph, known as *scree plot*, the number of components is reported on the abscissa and on the ordinate axis the corresponding eigenvalues (Figure A2-8).

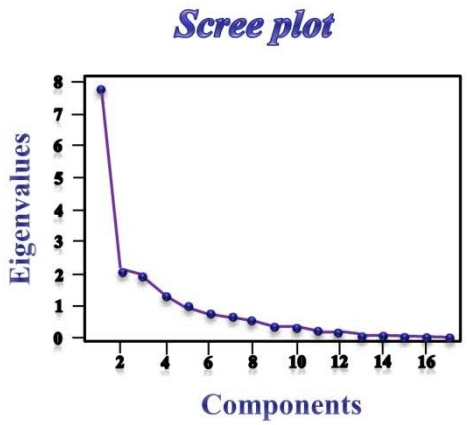

Figure A2-8 Example of *scree plot*

It's evident that the first principal components are associated with a greater percentage of explained variance, while the subsequent PCs are associated with a smaller amount of information (*asymptotic behavior*). A proper criterion of choice is to maintain only the first *n* PC for which a profound change in the slope of plotted line is manifested.

A2-5 SIMCA (Soft Independent Modeling of Class Analogy)

Classification is one of the fundamental aims of multivariate data analysis:

- To identify and quantitatively characterize subgroups within a given set of samples.
- To assess whether a new sample is similar to other samples, or to which group it belongs,
- To identify samples that are dissimilar compared to some standard

The *Soft Independent Modeling of Class Analogy* (SIMCA) represents one of most used chemometric classification approach based on similarities and is able to determine whether one or more new samples belong to an already existing group of similar samples.

The basis of this classical chemometric technique (Wold, 1976) is that objects in one class, or group, show similar rather than identical behavior and new objects are assigned to the class to which they show the largest similarity.

The SIMCA classification distinguishes two cases:

- Classes are known *a priori i.e.* the specific belonging of all the training set objects is known
- There is no *a priori* class membership available

The SIMCA approach presents some powerful advantages compared to methods like *e.g.* Linear Discriminant Analysis (LDA). Firstly SIMCA does not require that the number of objects is significantly larger than the number of variables as is invariably the case with classical statistical techniques: the present bilinear methods are effective regardless of the value of the ratio objects/ variables, be it either (very) many objects with respect to variables - or *vice versa*.

A further reason that makes SIMCA a powerful tool for classification regards the possibility of graphically representing the results with a great insight relatively to the specific data structures behind the modeled patterns.

In the classification of a new set of data, SIMCA calculates extensive statistics that enable to quantify “model envelopes”, or classification spaces, *surrounding* the classes. By defining the *distance* between the envelope and the model, it can be geometrically assessed whether new samples lie *inside* or *outside* a given model. It is thus possible to establish the model distances almost entirely by graphical means, although the full numerical result complement is of course also available in the background

Many plots are available to interpret the object/class-model relationships such as the well-known *Coomans plot* that gives the information about the class membership to any two models simultaneously. The Unscrambler® is the software here applied for this kind of data analysis.

A2-5.1 SIMCA Class-Models

SIMCA can be defined as a multi-application use of PCA-modeling as classification model; it usually consists of several PC-models, one for each class recognized. The first step is always a *class modeling*, *i.e.* to group the different object into homogeneous classes or clusters. Then each class is centered, scaled and modeled *separately*. Finally the new objects can be allocated to the established classes - or they may fall outside all "known" patterns.

In detail, the various steps of method are reported:

1. A data pretreatment is suggested.

2. A projection model of all objects to start to identify the individual classes is performed.

3. The pattern-specific classes and the discrimination between classes are simultaneously defined. All the relevant score plots have to be studied and all the problem specific groups or clusters identified. It has to be determined which objects belong to each subgroup in this training stage. There is always a great deal of interaction between the general problem context and the initial data analysis results in this stage.

4. A separate model for each class is delineated and all classes must be validated in the exact same way, or the membership limits will not be comparable.

5. It's necessary to remove outliers and to study the appropriate score plots to see if there should be more classes present than what is "known" in advance .

6. The optimal number of PCs for each class is individuated.

At this stage, the new data set can be classified; the data must of course be described by the *same set of variables* as for the training class models. If the calibration data (used for creating the class models) were in some way transformed, the new data must also be transformed in exactly the same way before classification. After choosing the class models to be used for the *pattern recognition*, the number of PCs to use in each model has to be defined (this is strongly problem-depend). The appropriate number of PCs depends on the data set, the goal, and the application.

Classifying a new object can result in several different results:

1. The object is *uniquely allocated* to one class, *i.e.* it fits to a single model within the given limits and is very far from the other classes.

2. The object may *fit several classes*, *i.e.* it has a distance that is within the critical limits of several classes simultaneously. This behavior can depend on either the given data are insufficient to distinguish between the different classes or the object actually does belong to several classes. It may be a borderline case or have properties of several classes.

3. The object *fits none* of the classes within the given limits. This is a very important result in spite of its apparent negative character. This *may* mean that the object is of a *new type*, *i.e.* it belongs to a class that was unknown before or - at least - to a class that has not been used in the classification. Alternatively it may simply be an *outlier*.

An important and peculiar aspect which makes SIMCA a powerful tool for the classification is to consider “failed” pattern recognition as an option of classification output. Clearly it is important to be able to identify such potentially important “new objects” with some measure of objectivity. It is therefore important to be able to specify the level of *statistical significance* associated with SIMCA results.

A2-5.2 *Statistical Significance Level in SIMCA*

Statistical significance tests are based on the *concern of making mistakes*. In the setup used in SIMCA-classification, the test quantifies the risk of saying that a particular object lies *outside* a specific model envelope. In The Unscrambler® classification results can be studied with varying *significance levels* - usually between 0.1% and 25% - whose value is chosen depending on the particular problem.

The "normal" statistical significance level used is 5%. *i.e.* there will be a 5% risk that a particular object falls outside the class, even if it truly belongs to it; 95% of the object which truly belong will thus fall inside the class. At opposite ends of the significance levels typically used it is noted that:

- Choosing a high significance level (*e.g.* 25%) means that only very "certain" objects will belong to the class, and (many) more "doubtful' outside it. Fewer objects that truly belong will fall inside the class (in this case,75%).
- Choosing a low significance level (*e.g.* 1%) means that cases, which are doubtful, will still be classified as be belonging to the class. More objects will be classified as members, *i.e.* almost all of the true member objects (*i.e.* 99%) will be members of the class.

A2-5.3 *The Coomans Plot*

This plot shows the orthogonal (transverse) distances from all new objects to two selected models (classes) at the same time. The critical, cut-off class membership limits are also indicated. These limits *may* be changed by editing the significance level.

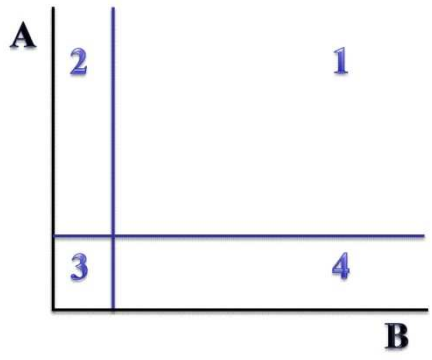

Figure A2-9 Coomans plot scheme

The two straight lines divide the space into four quadrants; the position of the straight lines is linked to the chosen significance level (Figure A2-9). The objects are assigned in this way:

1. In the first of the four quadrants the objects assigned to none of the selected models are placed
2. The zone 2 contains objects assigned to model B
3. In the zone 3 the objects assigned to both methods are placed
4. The objects placed in the last quadrant are assigned to model A

To make interpretation easier, the objects are color coded in The Unscrambler® (Figure A2-10).

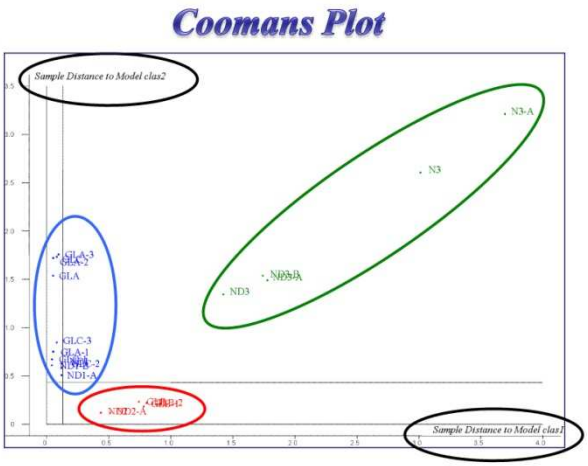

Figure A2-10 Example of *Coomans plot* scheme

SECTION B APPLICATIONS

CHAPTER B1 BEHAVIOUR OF PHOSPHOLIPIDS IN ANALYTICAL REACTIVE PYROLYSIS*

*Manuscript of the following article:

D. Melucci, G. Chiavari, S. Montalbani, S. Prati J Therm Anal Calorim (2011) 104:415–421

B1-1 Keywords

Gas chromatography-mass spectrometry

Pyrolysis

Derivatisation

Phospholipids

Tempera

Pigments

B1-2 Summary

Checking for the presence of egg in a painting layer allows to decide whether or not it is a tempera.

Several already assessed analytical techniques may be used to perform the chemical analysis for the detection of egg in paintings.

As an advantageous, alternative methodology for the determination of egg, a new application of analytical pyrolysis, hyphenated with gas chromatography-mass spectrometry (GC-MS) system, in presence of hexamethyldisilazane (HMDS) and tetramethylammonium hydroxide (TMAH), is here reported. The innovation lays mainly in the choice of new markers for the presence of egg. It is here demonstrated that in art diagnostic Tris-TMS-ester and methyl-ester of phosphoric acid, generated by the pyrolysis of standards phospholipids and synthetic painting layers containing egg as binding medium, may be used as new markers for identification of egg in tempera layers. The adoption of these new markers in analytical pyrolysis allows to obtain higher analytical performance with respect to classical markers (fatty acids), especially in terms of yield and, as a consequence, in terms of limit of detection.

B1-3 Introduction

Egg is a very complex matrix consisting of two separated phases: egg-white and egg-yolk.

Egg-white is an aqueous colloidal solution of proteins, mainly albumins, and low quantities of fats, while egg-yolk can be considered as an emulsion, that is a colloidal dispersion of phosphorated proteins and lipids. The lipidic fraction is the most consistent in egg-yolk, and is made up of triglycerides (65%w/w), phospholipids (29% w/w) and cholesterol (5.2% w/w) [12-13]. The phospholipids are triglycerides in which a phosphoric group esterifies one hydroxyl group of glycerol. This gives rise to a mono-glyceride of the tribasic phosphoric

acid further combined with other compounds. When phosphoric acid is esterified with coline, the resulting groups of compounds are called lecithin, mainly contained in egg-yolk, whose general formula is reported in Figure B1-1.

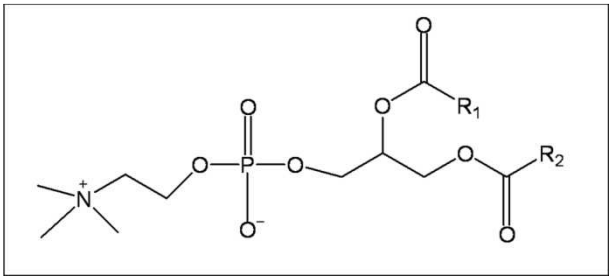

Figure B1-1 Example of a phosphatidylcholine, a type of phospholipid in lecithin (Rx refers to a hydrocarbon group)

The concentration of elementary phosphorus is equivalent to 0.9%w/w of dried whole egg [12].

The presence of egg in a painting layer may be evidenced in analytical pyrolysis, as reported in previous papers [14-17], using as markers the high concentration of palmitic acid and stearic acid, the presence in low concentration of indole and methyl indole, and cholesteryl compounds in widely variable concentrations.

Low concentration of azelaic acid allows distinguishing egg from siccative oil painting layers, where the intensity of the pyrographic peak of this acid is usually higher: this marker it is not sometimes sufficiently reliable because the concentration of the dicarboxylic products may be low, when the oil layer is fresh.

Other papers in literature deal with pyrolysis of egg painting layers [18-20].

The detection of egg in painting layers is commonly performed by means of the characterisation of the aminoacid content and sometimes of the total fatty acids composition with GC-MS analysis [21-25].

Spectroscopic techniques, like for instance Micro Raman analysis [26], are also used for the characterisation of the ligands.

Van den Brick et al. used matrix assisted laser desorption ionization - Fourier transform mass spectrometry to detect the alteration products from diacylphosphatidylcholines and triacylglycerols in egg tempera paintings for monitoring indoor museum conditions [27].

The determination of phospholipids is important in several fields: marine geochemistry [28], pharmaceutical research [29-30], biology [31-32], and studies concerning fat materials [33-38].

The phospholipids content has been taken into consideration in some papers to study egg painting layers, considering for instance the concentration of phosphorus [39], but the elementary phosphorus can arise even from some pigments.

The conventional method for the determination of phospholipids is known as the acid digest/arseno-molybdate Barlett method and it consists in phosphorus analysis.

Several methods are available for the determination of phospholipids using thin-layer chromatography [28, 32, 34, 38], Fourier-transform infrared spectroscopy (FTIR) [33], high-performance liquid chromatography [30] and gas-liquid chromatography for the characterisation of the fatty acids composition [34], colorimetric methods for the analysis of organic phosphorus complexed with a chromogenic reagent [37].

In this paper, pyrolysis-methylation [5] and pyrolysis–silylation [40,7,41] are described as useful techniques to identify the presence of phospholipids by the methyl-ester and tris-TMS-ester phosphoric acid that these analytes produce when undergoing pyrolysis. The identification of this product of pyrolysis has been carried out and it has been stated that it can be a useful marker for the detection of egg in painting layers.

B1-4 Materials and Methods

The derivatising reagents used, tetramethylammonium hydroxide (TMAH) 25%$_{w/v}$ water solution and Hexamethyldisilazane (HMDS), 98% $_{w/v}$, were from Aldrich.

Standard phospholipids, 1, 2-Dipalmitoyl-sn-glycero-3-phosphocholine and 1, 2-Dimyristoyl-rac-glycero-3-phosphocholine, were obtained from Aldrich.

Samples of painting standard layers, which are more than ten years old, were prepared by the Opificio delle Pietre Dure (Florence, Italy) following ancient recipes.

Samples (0.5 mg) were scraped from the support, inserted into a quartz capillary tube and added with 5 µl of derivatising reagent prior to pyrolysis.

Pyrolysis was carried out at 600°C for 10 s at the maximum heating rate using a CDS 1000 pyroprobe heated filament pyrolyser (Chemical Data System, Oxford, USA), directly connected to the injection port of a Varian 3400 gas-chromatograph coupled to a Saturn II ion-trap mass spectrometer (Varian Analytical Instruments, Walnut Creek, USA). A Supelco

SPB5 capillary column (30m, 0.25mm I.D., 0.25 µm film thicknesses) was used with a temperature programme from 50°C (held for 2 min) to 310°C (held for 5 min) at 5°C min^{-1} with helium as carrier gas. Temperatures of split/splitless injector (split mode) and Py-GC interface were kept at 250°C. The Py-GC interface was kept at 250°C and the injection port at 250°C. Injection mode was split (1:50 split ratio) and gas carrier was helium at flow rate of 1.0 ml min^{-1}. Mass spectra were recorded at 1 scan s^{-1} under electron impact at 70 eV, scan range 45 to 650 *m/z*. Structural assignment of the pyrolytical fragments was based on match with the NIST mass spectra library or literature data; the principal products identified are listed in Table B1- I and Table B1-II.

Table B1-I Products arising from pyrolysis-methylation of the samples

Peak No.	Fragment	M_w	Characteristic ions m/z (relative abundance)[a]
1	Phosphoric acid, trimethyl ester	140	65 (7), 79 (43), 95 (30), 110 (100), **140** (15)
2	Octanoic acid, methyl ester	158	74 (100), 87 (43), 115 (23), 158 (2), **159** (12)
3	Nonanoic acid, methyl ester	172	74 (100), 87 (47), 143 (16), **172** (5), 173 (14)
4	Octanedioic acid, dimethyl ester	202	129 (44), 138 (39), 171 (77), **202** (8), 203 (100)
5	Dodecanoic acid, methyl ester	214	74 (100), 87 (60), 143 (16), 171 (13), **214** (13)
6	Nonanedioic acid, dimethyl ester	216	55 (77), 152 (100), 185 (86), **216** (10), 217 (67)
7	Tetradecanoic acid, methyl ester	242	74 (100), 87 (66), 143 (28), 199 (18), **242** (18)
8	9-Hexadecenoic acid, methyl ester	268	55 (100), 83 (45), 96 (52), 236 (28), **268** (7)
9	Hexadecanoic acid, methyl ester	270	74 (100), 87 (67), 143 (25), 227 (15), **270** (27)
10	9-Octadecenoic acid, methyl ester	296	55 (100), 69 (63), 83 (55), 264 (17), **296** (5)
11	Octadecanoic acid, methyl ester	298	74 (100), 87 (71), 143 (30), 155 (18), **298** (40)

[a] The value in bold indicates the molecular ion

Table B1-II Products arising from pyrolysis-silylation of the samples

Peak N°	Fragment	Mw	Characteristic ions *m/z* (relative abundance) [a]
1	Phosphoric acid, bisTMS monomethyl ester	256	73 (18), 133 (16), 211 (14), 241 (100), **256** (15)
2	Phosphoric acid, triTMS ester	314	73 (60), 283 (8), 299 (100), 300 (24), **314** (16)
3	Octanedioic acid, bisTMS ester	318	75 (100), 169 (36), 187 (50), 303 (63), **318** (3)
4	Nonanedioic acid, bisTMS ester	332	73 (47), 201 (27), 317 (100), **332** (2), 333 (23)
5	Tetradecanoic acid, TMS ester	300	73 (68), 117 (100), 129 (39), 285 (59), **300** (8)
6	9-Hexadecenoic acid, TMS ester	326	73 (100), 117 (72), 129 (59), 311 (41), **326** (8)
7	Hexadecanoic acid, TMS ester	328	73 (49), 117 (100), 129 (38), 313 (46), **328** (22)
8	9-Octadecenoic acid, TMS ester	354	73 (100), 117 (64), 129 (57), 339 (30), **354** (6)
9	Octadecanoic acid, TMS ester	356	73 (65), 117 (100), 129 (40), 341 (57), **356** (22)

[a] The value in bold indicates the molecular ion

B1-5 Results and discussion

B1-5.1 Pyrolysis-methylation

B1-5.1.1 Analysis of standard phospholipids

The figures B1-2 and B1-3 show the pyrolytical patterns related to pyrolysis-methylation of standard phospholipids, 1, 2-Dipalmitoyl-sn-glycero-3-phosphocholine and 1, 2-Dimyristoyl-rac-glycero-3-phosphocholine, and the chromatographic and mass spectrometric data of the main peaks are reported in Table B1-1.

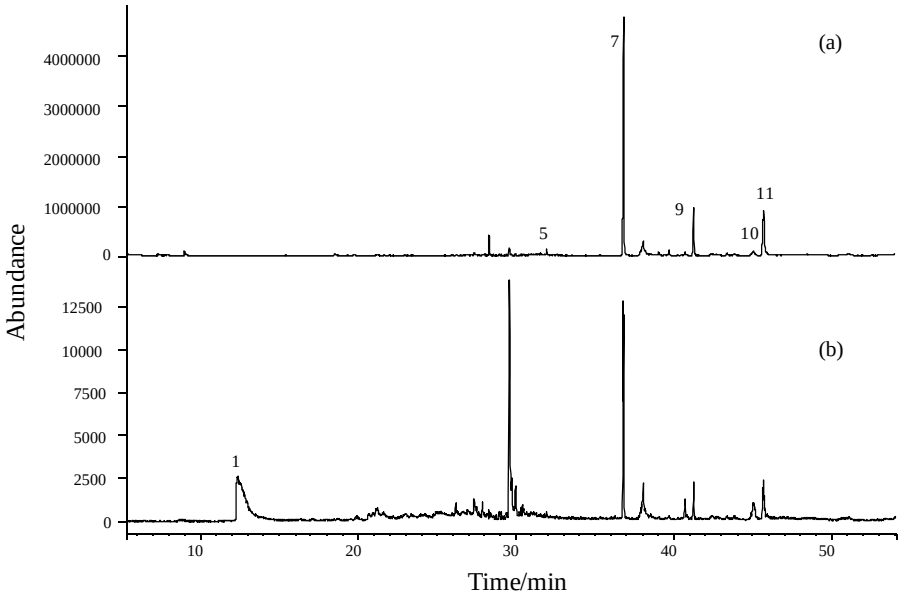

Figure B1-2 (a) Reconstructed ion chromatogram arising from pyrolysis-methylation of 1, 2-Dimyristoyl-rac-glycero-3-phosphocholine (b) single ion monitoring for *m* / *z* (110+ 140) the ions deriving from phosphoric acid, trimethyl ester

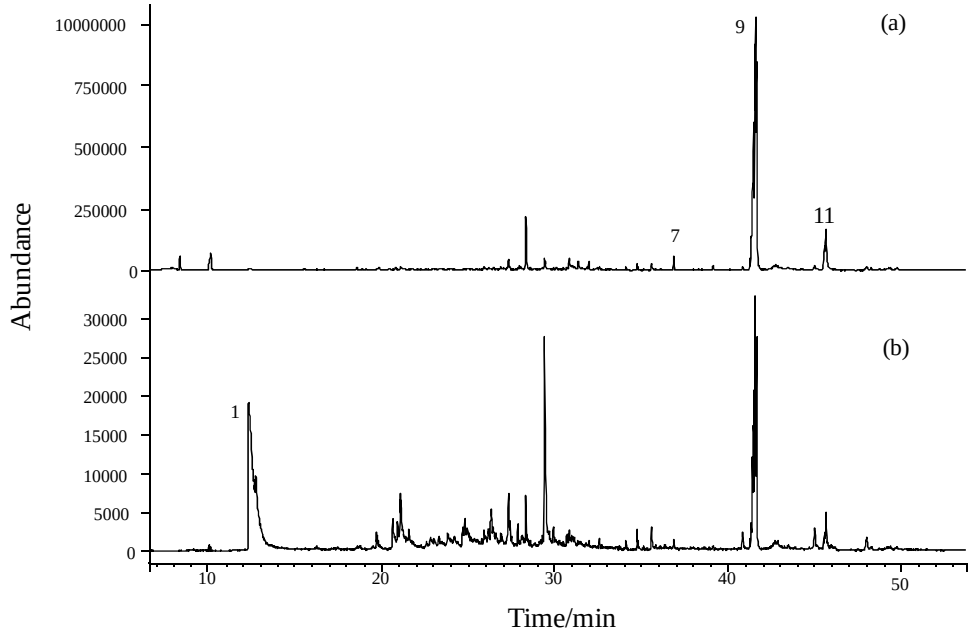

Figure B1-3 a) Reconstructed ion chromatogram arising from pyrolysis-methylation of 1, 2-Dipalmitoyl-sn-glycero-3-phosphocholine (b) single ion monitoring for *m* / *z* (110+ 140) the ions deriving from phosphoric acid, trimethyl ester

Analytical pyrolysis in presence of tetramethylammonium hydroxide is a powerful technique to characterize different kinds of organic compounds, but it has been pointed out that its strong alkalinity (pKb =0.9) [42] causes some secondary reactions such as degradation or isomerisation. It is likely that this behaviour leads to a low yield in trimethyl phosphate, so that it's advisable to use the Single Ion Monitoring (SIM) relative to the characteristic ions of the trimethyl phosphate to detect its presence in a sample.

In spite of this situation, the recognition of the phosphate is still possible thanks to a good specifity of the mass spectrum. The mass spectrum of trimethylphosphate (Figure B1-4) shows the peak at 110 *m/z* as the main peak derived from the molecular ion by the loss of a neutral formaldehyde molecule.

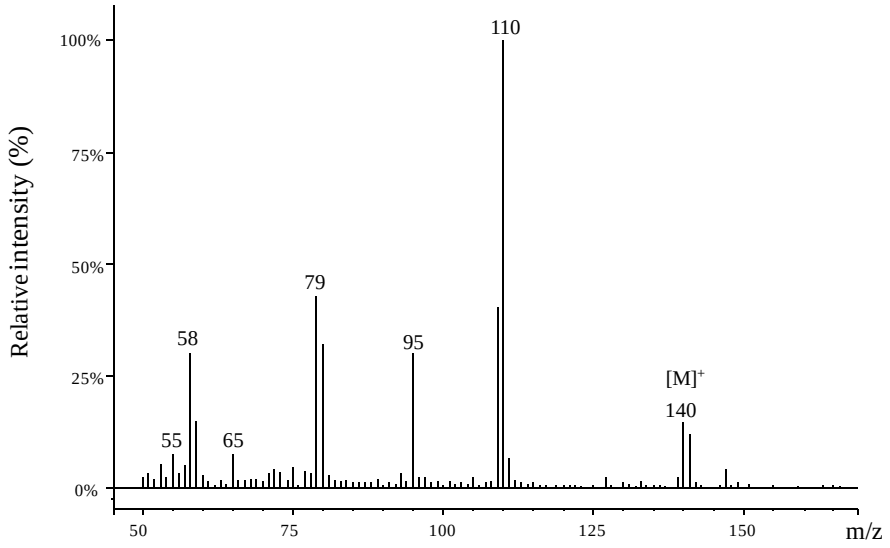

Figure B1-4 Mass spectrum of phosphoric acid, trimethyl ester

B1-5.1.2 Analysis of standard painting layers

The result, obtained with the pyrolysis-methylation of a standard painting layer prepared from whole egg (Figure B1-5, Table B1-I), has confirmed the trend highlighted for the standard phospholipids: the presence of the peak corresponding to the phosphoric acid, trimethyl ester, is evident only in the Single Ion Monitoring .

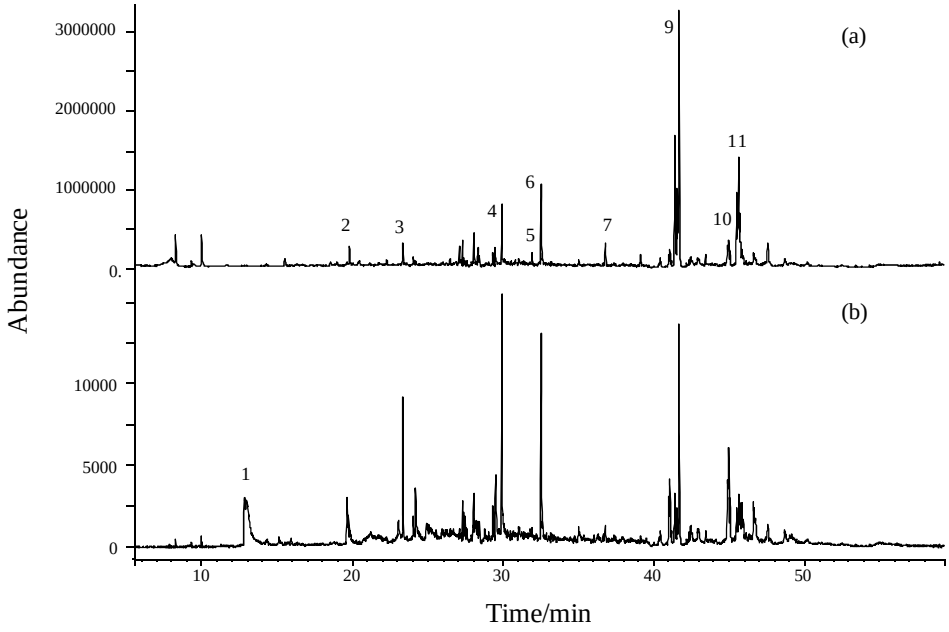

Figure B1-5 (a) Reconstructed ion chromatogram arising from pyrolysis-methylation of standard painting layer prepared from whole egg (b) single ion monitoring for *m* / *z* (110+ 140) the ions deriving from phosphoric acid, trimethyl ester

In Figure B1-6 the Single Ion Monitoring for *m/z* (110+140) arising from pyrolysis-methylation of standard painting layers prepared from whole egg with different pigments such as malachite [$Cu_2(CO_3)(OH)_2$], smalt blue [SiO_2,K_2O,Al_2O_3,CoO], azurite [$Cu_3(CO_3)_2(OH)_2$] and cinnabar [HgS] are reported.

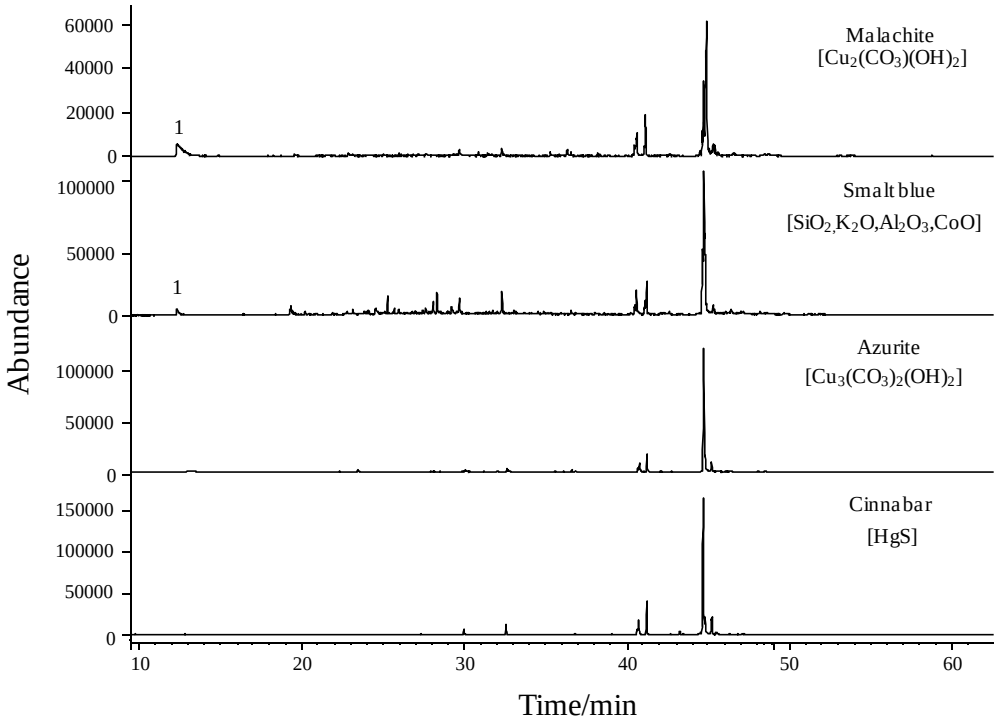

Figure B1-6 Single ion monitoring for m / z (110+ 140) arising from pyrolysis-methylation of standard painting layers prepared from whole egg with different pigments : malachite, smalt blue,azurite and cinnabar

The presence of a pigment changes the yield of trimethyl phosphate, in particular regarding smalt blue and cinnabar. This behaviour can be explained considering the fact that some inorganic salts, like some painting pigments, can affect the formation of some pyrolysis products or lead to secondary pyrolysis reaction [43-44]: the ability of the powder pigment to influence the pyrolitic behaviour could not be excluded, and can justify the difficulty to detect the presence of phosphate when this technique is used for recognition of egg in tempera layers.

B1-5.2 *Pyrolysis-silylation*

B1-5.2.1 *Analysis of standard phospholipids*

In Figures B1-7 and B1-8 (Table B1-II) the pyrolytical patterns related to the pyrolysis-silylation of the standard phospholipids, 1,2 -Dipalmitoyl-sn-glycero-3-phosphocholine and 1,2-Dimyristoyl-rac-glycero-3-phosphocholine are reported.

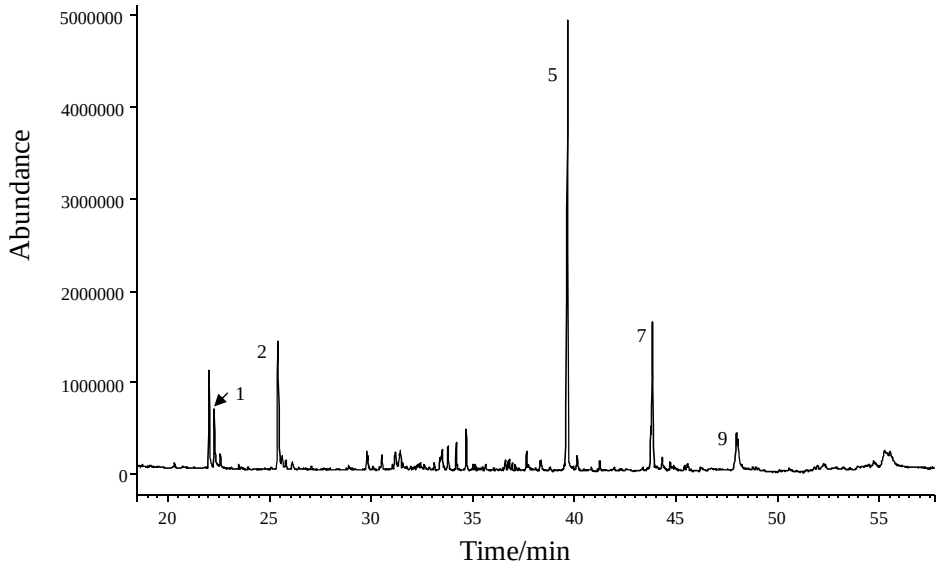

Figure B1-7 reconstructed ion chromatogram arising from pyrolysis-silylation of 1, 2-Dimyristoyl-rac-glycero-3-phosphocholine

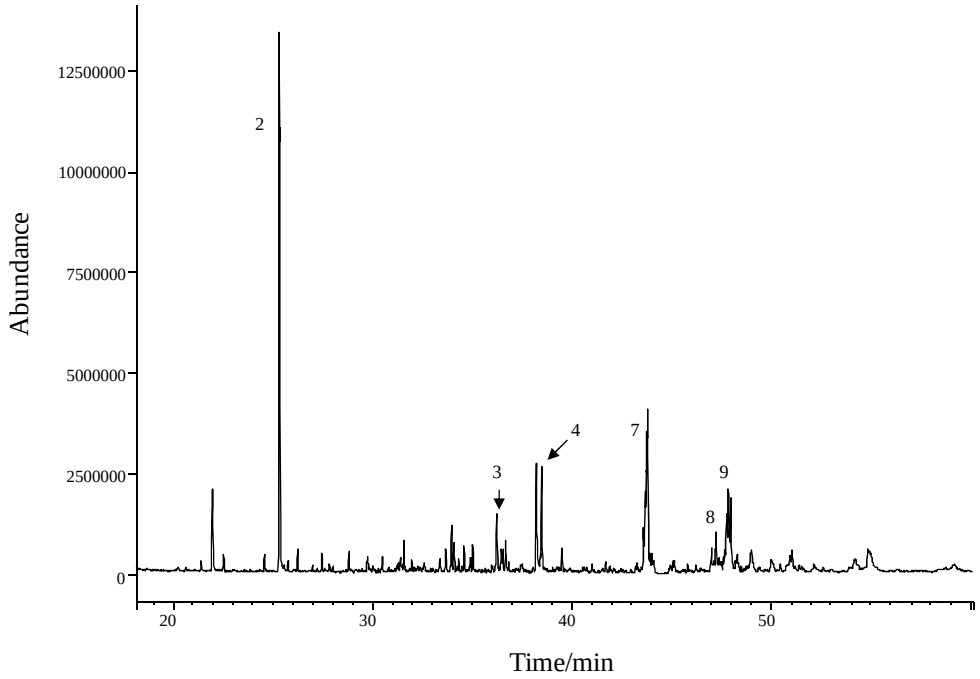

Figure B1-8 Reconstructed ion chromatogram arising from pyrolysis-silylation of 1, 2-Dipalmitoyl-sn-glycero-3-phosphocholine

The chromatographic profile obtained in both cases shows that the use of a derivatising reagent less aggressive than TMAH leads to an increased yield of tri-TMS-phosphate, which is clearly visible in the total ion chromatogram: the peak 2 (Figures B1-9) has a molecular peak corresponding at 314 *m/z*, which is the molecular weight of tri-TMS-phosphate with base peak 299 *m/z* obtained from the loss of a methyl group.

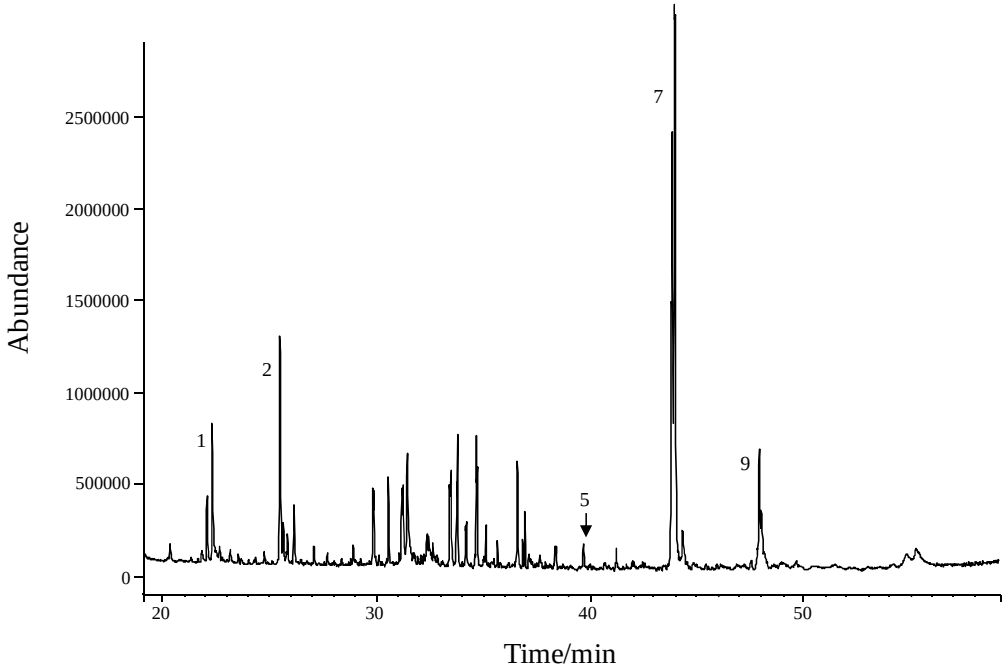

Figure B1-9 Mass spectrum of phosphoric acid, triTMS ester

B1-5.2.2 Analysis of standard painting layers

Figure B1-10 reports the reconstructed ion chromatogram arising from pyrolysis-silylation of a standard painting layer prepared from whole egg.

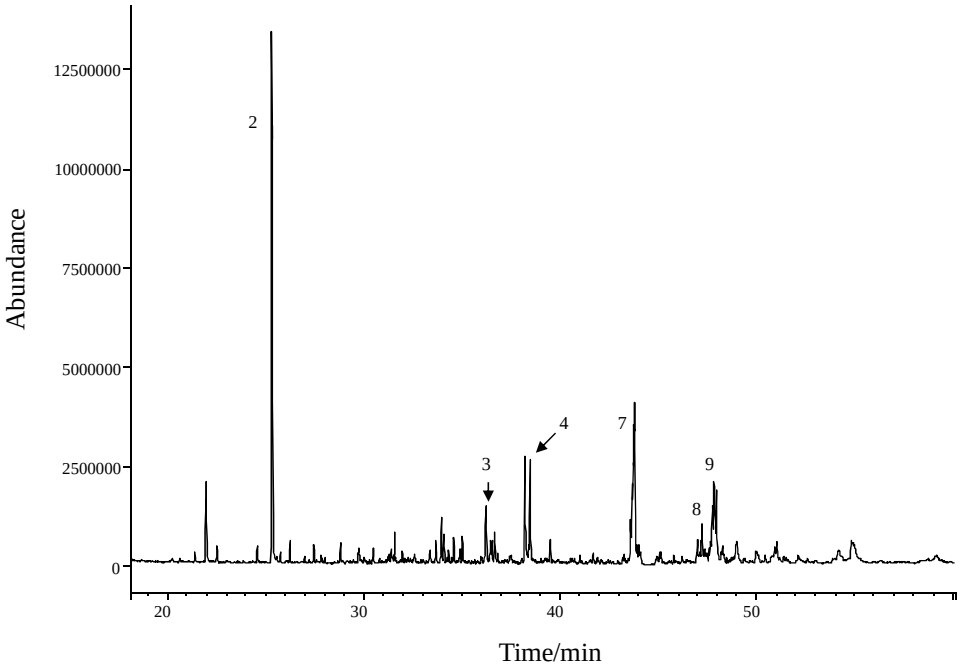

Figure B1-10 Reconstructed ion chromatogram arising from pyrolysis-silylation of standard painting layers prepared from whole egg

The result confirms the hypothesis expressed in the case of standard phospholipids: analytical pyrolysis performed in the presence of HMDS leads to the formation of a greater amount of phosphate. The corresponding peak (Figure B1-10, peak 2) is the most intense and gives the mass spectrum reported in Figure B1-12.

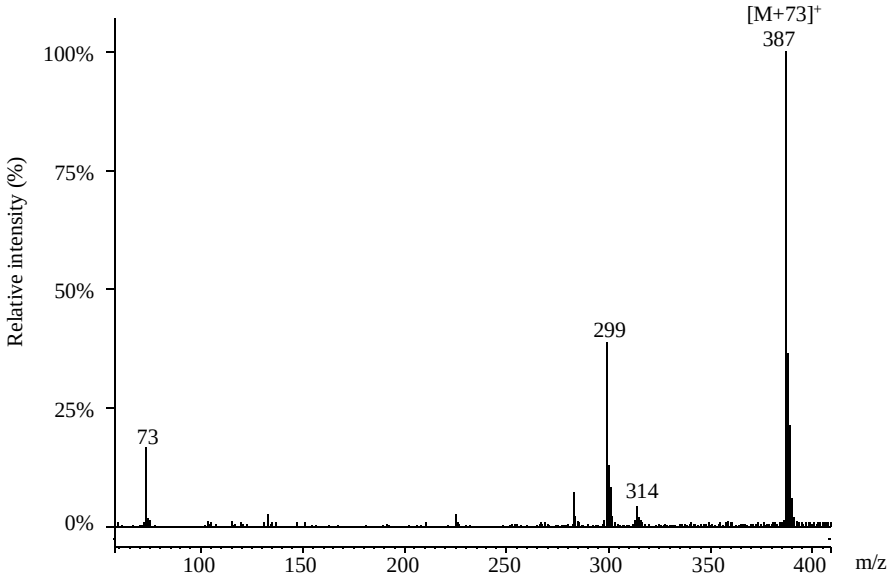

Figure B1-11 Mass spectrum of compound corresponding to the peak 2

The consistent peak at 387 *m/z* may be attributed to an association ion $M(TMS)^+$: we have found that the TMS^+ gives such type of association with the molecular compound, working with ion trap detector.

In Figure B1-12 the total ion-chromatograms arising from pyrolysis-silylation of standard painting layers prepared from whole egg with different pigment – malachite [$Cu_2(CO_3)(OH)_2$], smalt blue [SiO_2,K_2O,Al_2O_3,CoO], azurite [$Cu_3(CO_3)_2(OH)_2$] and cinnabar [HgS] - are reported.

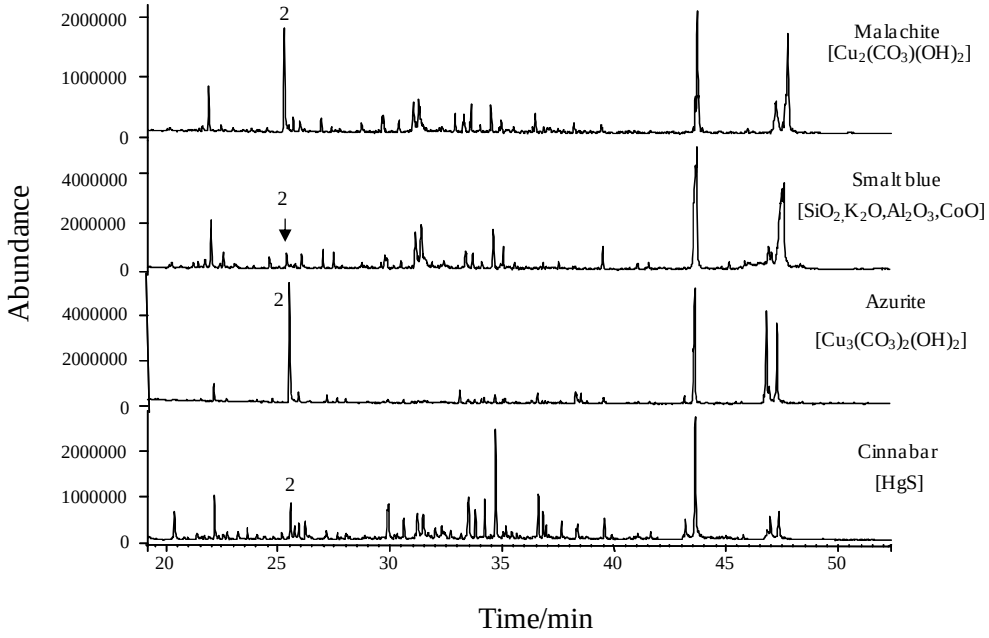

Figure B1-12 Reconstructed ion chromatogram arising from pyrolysis-silylation of standard painting layers prepared from whole egg with different pigments: malachite, smalt blue, azurite and cinnabar

As in the case of pyrolysis-methylation, the effect of the pigment on the formation of tris-TMS-phosphate is evident; in particular, it is clear that the presence of smalt blue as that of cinnabar is associated with the lowest intensity of the peaks.

B1-6 Conclusions

The study here presented shows a new application of analytical pyrolysis in the characterisation of complex organic compounds such as the egg yolk largely used in the old painting technique as ligand.

We have compared two of the reagents usually used in pyrolysis for the speciation of high polar compounds: tetramethylammonium-hydroxide (TMAH) as methylating reagent and hexamethyldisilazane (HMDS) as silylating reagent. Both the reagents give a positive recognition of phosphoric acid esters present in lipidic structures with the formation of trimethyl phosphate and tris-TMS-phosphate, offering a further potential marker for the analysis of painting ligands.

However, the silylation appears to be more useful, thanks to the higher yields obtained.

Considering the positive results obtained with OPD tempera layers, the next objective is the application to the real painting cases but, until now, we have obtained scarce results probably because of the difficulties to find reliable ancient tempera sample and the high degradability that distinguishes this kind of compounds.

CHAPTER B2 MULTIVARIATE METHODS FOR EXPERIMENTAL-DATA ANALYSIS APPLIED TO IDENTIFICATION OF BACTERIA BY MEANS OF PY-GC-MS

B2-1 Summary

There is the need to identify bacteria by means of reliable but rapid analytical methodologies: the already assessed current analytical methodologies require laborious preparative steps; moreover, they are often very expensive. In order to solve these problems, a more rapid and less expensive analytical technique based on Py-GC-MS has been here developed.

Analytical Pyrolysis (Py) allows for the characterization of a material by controlled thermal chemical degradation in inert atmosphere; degradation products are characteristic of the original sample and can be characterized using a GC-MS system. When Py-GC-MS is applied to bacteria, trivial visual examination of *pyrograms* easily allows discrimination between different *genera* of bacteria. But in order to obtain discrimination in terms of *species*, chemometric data analysis is necessary.

Among chemometric methods for *data exploration*, *Principal Components Analysis* (*PCA*) is extensively used in Analytical Chemistry. As for *classification procedures*, Soft Independent Modeling of Class Analogy (SIMCA) is particularly suitable for analytical problems, since it works even when the number of samples is very low with respect to the number of variables characterizing samples.

The purpose of the present work was to develop a rapid Py-GC-MS methodology for the analysis and identification of bacteria in terms of *genus* and *species*. Methylated fatty acids (FAMEs) were chosen as biomarkers of bacteria in the pyrolysates. *In situ* Thermal Hydrolysis and Methylation (THM) was applied. Pyrographic peak areas, and relevant normalized or relative values, were chosen as variables for chemometric data processing by PCA and SIMCA.

Satisfactory results were obtained in classifying bacteria in terms of *genus* and *species*.

B2-2 Keywords:

Gas chromatography-mass spectrometry

Pyrolysis

Bacteria

Principal Components Analysis (PCA)

Soft Independent Modeling of Class Analogy (SIMCA)

B2-3 Introduction

Bacteria are ubiquitous microorganisms whose presence can be detected in water, soil, air and in other more complex living organisms. Their characterization is essential to solve common medical, biological, food, environmental, cultural heritage and forensic problems.

Traditional microbiological identification of bacteria is based on appearance under a microscope (shape, size, presence of particular structures), response to staining (*e.g.*, the classic Gram stain), or indirect characteristics (growth under aerobic or anaerobic conditions, generation of specific enzymes or biochemical products, etc.) [1]. Molecular analysis is another method based on identifying the genetic characteristics of the microorganism [45].

The study of morphological characteristics and microbial culture allows to obtain information in a relatively short time. However, the obtained results are generic and do not imply differentiation at the level of bacterial species. Molecular analysis leads to undeniable results and, therefore, to uniquely identify a micro-organism down to the level of species, but this analysis is very long and laborious and, moreover, the relevant costs are rather high [46].

Chromatographic techniques and coupled to mass spectrometry offer a significant lowering of costs and time of analysis, together with efficient results as regards the determination of *markers* characteristic such as methylated fatty acids (FAMEs) [47,48].

On the other hand, for the sake of completeness, it should be noted that the gas chromatographic analysis of microorganisms includes a step of pretreatment of the sample still relatively laborious. Chemical components characteristic of specific microorganisms or microbial groups are usually enclosed in part of a polymeric cellular matrix. Such nonvolatile and intractable biological samples determine some difficulties in direct characterization by GC-MS. When analytical pyrolysis is applied to a bacteria sample, depolymerization and volatilization of components are accomplished simultaneously. The volatile thermal products may be detected directly by MS (Py-MS) [49-52] or separated on-line by capillary GC and detected by MS (Py-GC-MS) [53-56]. In contrast to derivatization-based methods and like other direct approaches, bacteria characterization by analytical pyrolysis requires minimal sample pretreatment and short total analysis times.

The purpose of the present work was to develop a rapid Py-GC-MS methodology for the analysis and identification of bacteria in terms of *genus* and *species*. Two lines of research have been developed: the first regards the analysis of different bacterial species belonging to the genera *Bacillus* and *Pseudomonas* grown under the same culture conditions while the

second concerns the study of the strain of *Rhodococcus Aetherovorans BCP1* exposed to different sources of carbon during growth [57].

Methylated fatty acids (FAMEs), deriving from phospholipids forming the cell wall of the microorganism, were chosen as biomarkers of bacteria in the pyrolysate [58]. *In situ* Thermal Hydrolysis and Methylation (THM) was applied [5]. Pyrographic peak areas, and relevant normalized or relative values, were chosen as variables for chemometric data processing by PCA (*Principal Component Analysis*) and SIMCA (*Soft Independent Models of Class Analogy*) were performed on the experimental dataset [59,9,10].

B2-4 Materials and methods

B2-4.1 Bacteria samples

The bacterial samples (Table B2-II) were provided by Prof. Stefano Fedi of the Department of Experimental Evolutionary Biology (BES).

Table B2-II Bacteria samples

Genus	**Species**
Rhodococcus	*Aetherovorans BCP1*
Bacillus	*Subtilus*
	Licheniformis
	Firmus
Pseudomonas	*Fluorescens*
	Putida
	Stutzeri C12
	Stutzeri C/IS A
	Pseudoalcaligenes
	Bif 7

B2-4.1.1 Preparation of Bacillus samples

Each species of *Bacillus* was subjected to the same preparative procedure that involves:

- smear samples on Petri plates
- growing for 24-48 h in a growth chamber at 37°C
- removal of bacterial colonies using sterilized loop
- inoculum flasks containing LB (*Luria-Bertani Broth*), the medium used for the enrichment (Table B2-II)
- thermal agitation (T = 30°C; t = all night; v = 150 rpm)
- centrifugation (t = 10 min; v = 5000 rpm)

- elimination of supernatant, resuspension in physiological solution, centrifugation
- drying in an oven at 105 ° C overnight
- storage at -20 ° C until analysis

Table B2-II *Luria-Bertani Broth, LB*

Luria-Bertani Broth, LB (pH 7)	
Tryptone	10 g/l
Yeast Extract	5 g /l
NaCl	10 g /l
H_2O	to volume

B2-4.1.2 Preparation of Pseudomonas samples

Each species of *Pseudomonas* was subjected to the same preparative procedure that involves:

- smear samples on Petri plates
- growing for 24-48 h in TSA medium, in a growth chamber at 37°C. TSA (*Tryptic Soy Agar*), is the medium used for the culture plate (Table B2-III)
- removal of bacterial colonies using sterilized loop
- inoculum flasks containing TSB (*Tryptic Soy Broth)* the medium used for the enrichment (Table B2-IV)
- thermal agitation (T = 30°C; t = all night; v = 180 rpm)
- centrifugation (T=4°C; t = 10 min; v = 8000 rpm)
- elimination of supernatant, resuspension in physiological solution, centrifugation
- drying in an oven at 105 ° C overnight
- storage at -20 ° C until analysis

Table B2-III *Tryptic Soy Agar, TSA*

Tryptic Soy Agar, TSA (pH 7.2)	
Casein Peptone	17 g/l
Soya Peptone	3.0 g/l
Glucose	2.5 g/l
K_2HPO_3	2.5 g/l
NaCl	5.0 g/l
Agar	15 g/l
H_2O	to volume

Table B2-IV *Tryptic Soy Broth, TSB*

Tryptic Soy Broth, TSB (pH 7.2)	
Pancreatic digest of casein	17 g/l
Enzymatic digest of soybean meal	3 g/l
NaCl	5 g/l
Dipotassium Phosphate	2.5 g/l
Dextrose	2.5 g/l
H_2O	to volume

B2-4.1.3 Preparation of Rhodococcus samples

The *BCP1* strain was isolated from chloroform contaminated groundwater. The procedure used for the culture of this strain was performed in several stages:

- smear samples on Petri plates
- growing for 24-48 h in LB medium , in a growth chamber at 37°C
- removal of bacterial colonies using sterilized loop
- resuspension in physiological solution
- inoculum flasks containing MM (*minimal medium*), Table B2-V
- addition of a different carbon source (C12, C16, C20, 1-pentanol, toluene, succinate, and hexane) in each flask
- thermal agitation (T = 30°C; t = all night; v = 180 rpm)
- centrifugation (T = 4°C; t = 10 min; v = 8000 rpm)
- elimination of supernatant, resuspension in physiological solution, centrifugation (3 times)
- drying in an oven at 105 ° C overnight
- storage at -20 ° C until analysis

Table B2-V *Minimal medium, MM*

Minimal medium, MM (pH 7.2)	
MM (pH 7.2) (100X) – in BD water:	
K_2HPO_4	155 g/l
$NaH_2PO_4 \cdot H_2O$	73.9 g/l
$(NH4)_2SO_4$	10.52 g/l
$NaNO_3$	76.5 g/l
$MgSO_4$ stock solution (1000X) – in BD water:	
$MgSO_4 \cdot 7H_2O$	60.2 g/l
$CaCl_2$ stock solution (1000X) – in BD water:	
$CaCl_2$	14.7 g/l
Microelements stock solution (1000X) – in BD water:	
$FeSO_4 \cdot 7H_2O$	6.283 g/l
$MnCl_2 \cdot 4H_2O$	0.3 g/l
$ZnSO_4 \cdot 7H_2O$	0.147 g/l
H_3BO_3	0.061 g/l
$Na_2MoO_4 \cdot H_2O$	0.109 g/l
$NiCl_2 \cdot 2H_2O$	0.024 g/l
$CuCl_2 \cdot 2H_2O$	0.017 g/l
$CoCl_2 \cdot 6H_2O$	0.024 g/l
Yeast extract stock solution (10000X) – in BD water	5gr/l

B2-4.2 Experimental conditions

The derivatising reagents used, tetramethylammonium hydroxide (TMAH) 25%$_{w/v}$ water solution, was from Aldrich.

Samples (0.5 mg) were inserted into a quartz capillary tube and added with 5 µl of derivatising reagent prior to pyrolysis.

Pyrolysis was carried out at 600°C for 10 s at the maximum heating rate using a CDS 1000 pyroprobe heated filament pyrolyser (Chemical Data System, Oxford, USA), directly connected to the injection port of a Varian 3400 gas-chromatograph coupled to a Saturn II ion-trap mass spectrometer (Varian Analytical Instruments, Walnut Creek, USA). A Supelco SPB5 capillary column (30m, 0.25mm I.D., 0.25 µm film thicknesses) was used with a temperature programme from 50°C (held for 5 min) to 130°C at 5°C min^{-1}, from 130°C to 170° C at 2,5°C min^{-1} , from 170° C to 300°C at 5°C min^{-1} with helium as carrier gas. Temperatures of split/splitless injector (split mode) and Py-GC interface were kept at 250°C. The Py-GC interface was kept at 250°C and the injection port at 250°C. Injection mode was split (1:50 split ratio) and gas carrier was helium at flow rate of 1.0 ml min^{-1}. Mass spectra were recorded at 1 scan s^{-1} under electron impact at 70 eV, scan range 45 to 650 *m/z*. Structural assignment of the pyrolytical fragments was based on match with the NIST mass spectra library or literature data; the principal products identified are listed in Table B2- I

Table B2-I Products arising from pyrolysis-methylation of the samples

Peak No	Fragments	M_w	Characteristic ions m/z (relative abundance)[a]
1	Dodecanoic acid, methyl ester	214	74 (100), 87 (62), 143 (22), 171 (16), **214** (12)
2	Tridecanoic acid, methyl ester	228	74 (100), 87 (66), 143 (27), 185 (16), **228** (15)
3	Tridecanoic acid, 12-methyl-, methyl ester	242	74 (100), 87 (70), 143 (23), 199 (51), **242** (11)
4	Tetradecanoic acid, methyl ester	242	74 (100), 87 (67), 143 (33), 199 (20), **242** (22)
5	Tetradecanoic acid, 13-methyl-, methyl ester	256	74 (100), 87 (71), 143 (24), 213 (24), **256** (17)
6	Tetradecanoic acid, 12-methyl-, methyl ester	256	74 (100), 143 (27), 199 (57), 213 (13), **256** (13)
7	Pentadecanoic acid, methyl ester	256	74 (100), 87 (64), 157 (14), 213 (16), **256** (28)
8	Pentadecanoic acid, 14-methyl-, methyl ester	270	74 (100), 87 (70), 143 (25), 227 (21), **270** (24)
9	9-hexadecenoic acid, methyl ester	268	55 (100), 152 (26), 194 (20), 236 (26), **268** (8)
10	Isomer peak 9	268	55 (100), 152 (15), 194 (10), 236 (16), **268** (5)
11	Hexadecanoic acid, methyl ester	270	74 (100), 87 (65), 143 (34), 227 (18), **270** (52)
12	11-hexadecenoic acid, 15-methyl-, methyl ester	282	55 (100), 96 (67), 213 (25), **282** (4), 283 (6)
13	Hexadecanoic acid, 15-methyl-, methyl ester	284	74 (100), 87 (77), 143 (25), 241 (21), **284** (6)
14	Hexadecanoic acid, 14-methyl-, methyl ester	284	74 (100), 143 (28), 185 (20), 199 (209), **284** (15)
15	Isomer peak 12	282	55 (100), 96 (38), 208 (13), 250 (20), **282** (7)
16	Isomer peak 15	282	55 (100), 166 (11), 250 (21), **282** (4), 283 (11)
17	Heptadecanoic acid, methyl ester	284	74 (100), 87 (68), 143 (30), 241 (16), **284** (37)
18	Heptadecanoic acid, 16-methyl-, methyl ester	298	74 (100), 143 (56), 199 (30), 255 (21), **298** (31)
19	9-octadecenoic acid, methyl ester	296	55 (100), 180 (12), 222 (12), 264 (24), **296** (8)
20	Octadecanoic acid, methyl ester	298	74 (100), 143 (36), 199 (20), 255 (22), **298** (62)

[a] The value in bold indicates the molecular ion

The mass spectrum relevant to a methylated fatty acid presents a peak base with *m/z*=74 that derives from a particular fragmentation reaction, the *McLafferty rearrangement*. Since FAMEs were chosen as *markers* for the bacteria characterization, it is appropriate to consider the chromatographic profiles obtained in Single Ion Monitoring (SIM) at *m/z*=74; in this way the background is minimized allowing a more accurate determination of peak areas.

B2-5 Results and discussion

B2-5.1 *Identification of pyrolysis compounds*

In Figures B2-1, B2-2 and B2-3, pyrograms (SIM; *m/z*=74) arising from pyrolysis-methylation of bacteria samples (*Rhodococcus, Bacillus e Pseudomonas),* are reported (Table B2-I).

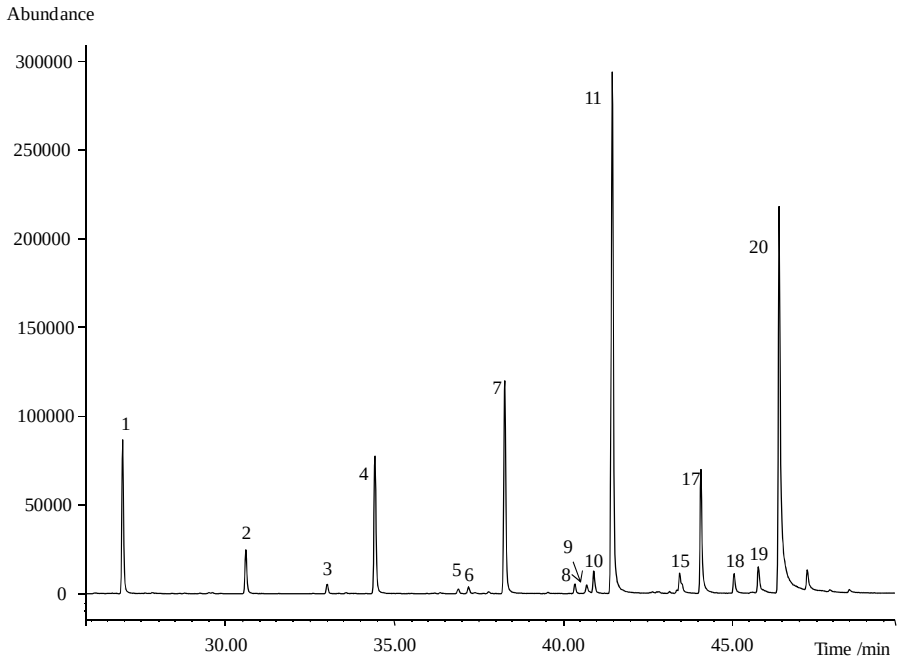

Figure B2-1 Single ion monitoring (*m/z* = 74) arising from pyrolysis-methylation of a sample of Rhodococcus Aetherovorans BCP1 (Source of carbon: 1-pentanol)

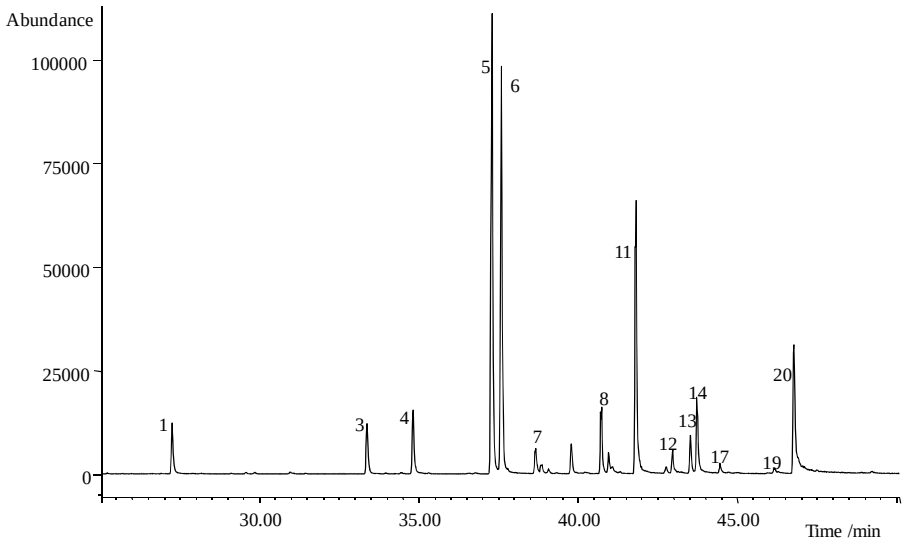

Figure B2-2 Single ion monitoring (*m/z* = 74) arising from pyrolysis-methylation of a sample of Bacillus Firmus

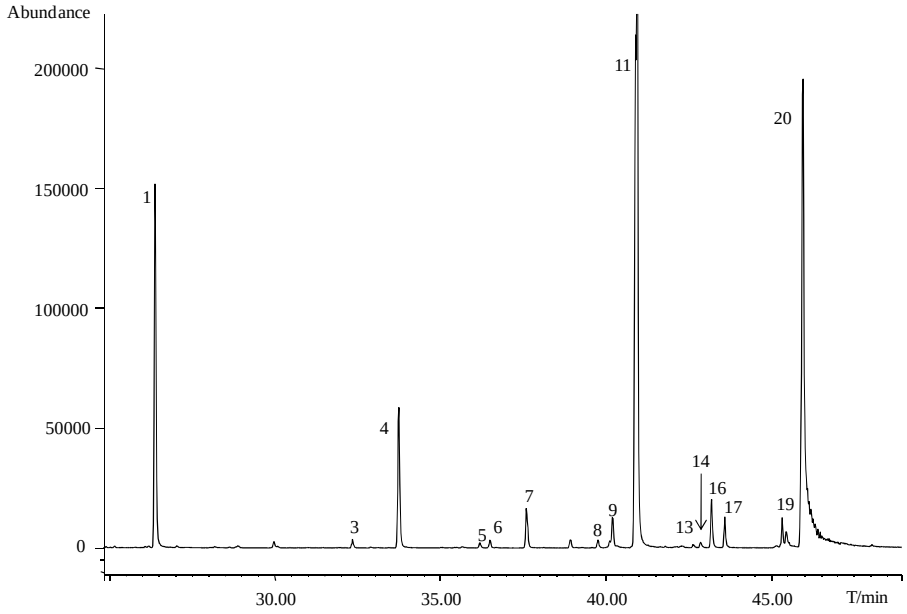

Figure B2-3 Single ion monitoring (*m/z* = 74) arising from pyrolysis-methylation of a sample of Bacillus Firmus

Pyrolysis of bacterial sample leads to the formation of a great variety of FAMEs, branched and linear, saturated and unsaturated fatty acids. Their identification required a thorough study of the relevant mass spectra, especially as regards the distinction of various isomers, linear and branched. By way of example, the mass spectra relevant to tetradecanoic acid, 13-methyl-, methyl ester (peak 5), tetradecanoic acid, 12-methyl-, methyl ester (peak 6) and pentadecanoic acid, methyl ester (peak 7) are reported in Figures B2-4, B2-5 and B2-6.

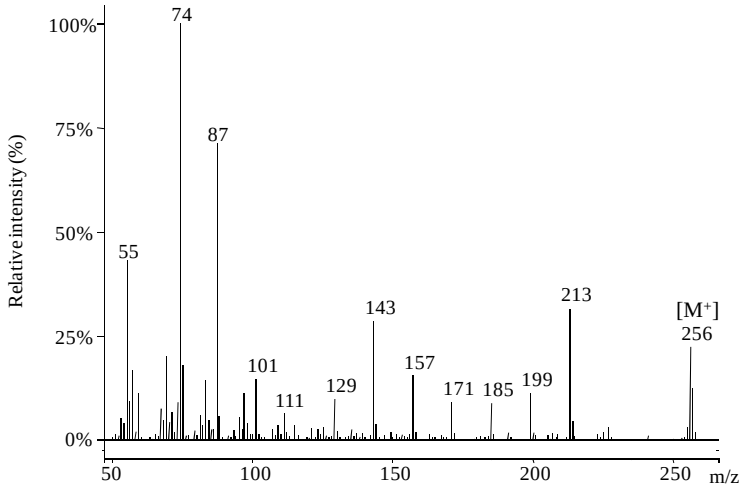

Figure B2-4 Mass spectrum of tetradecanoic acid, 13-methyl-, methyl ester (peak 5)

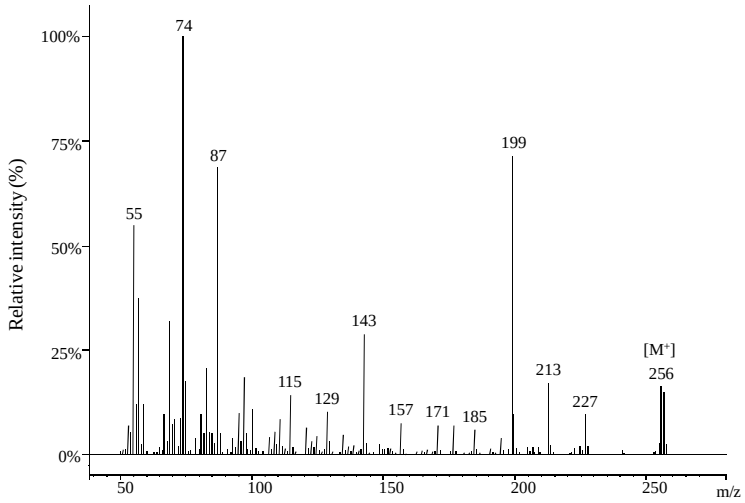

Figure B2-5 Mass spectrum of tetradecanoic acid, 12-methyl-, methyl ester (peak 6)

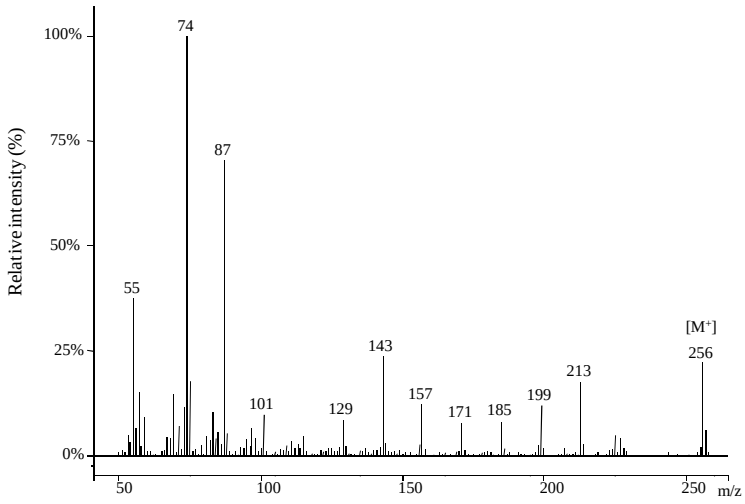

Figure B2-6 Mass spectrum of pentadecanoic acid, methyl ester (peak 7)

B2-5.2 *Reproducibility*

Reproducibility in an analytical method is essential to validate the results obtained. In Figure B2-7 a comparison between the pyrograms (SIM: *m/z*=74) for three repeated analysis performed on the sample of *R. Aetherovorans* BCP1 is reported.

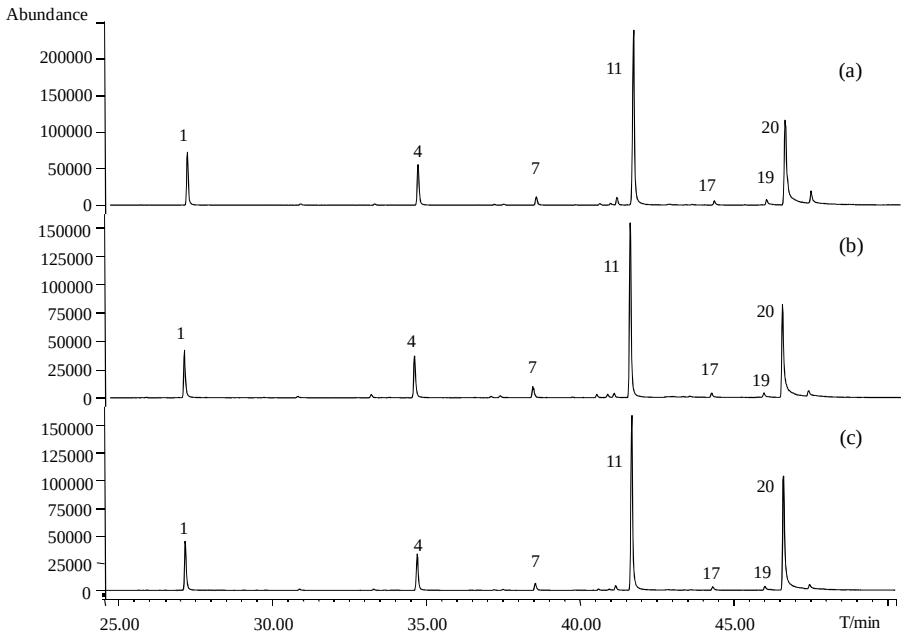

Figure B2-7 Comparison between pyrograms (SIM: *m/z*=74) for three repeated analysis performed on the sample of R. *Aetherovorans* BCP1 grown in the presence of hexane

The chromatographic profiles show a high reproducibility of the analyses performed by the integrated system Py-GC-MS.

B2-5.3 *Distinction between bacterial genera*

In Figure B2-8 a comparison between pyrograms (SIM: *m/z*=74) for the bacteria samples (*Bacillus*, *Rhodococcus*, *Pseudomonas)* is reported.

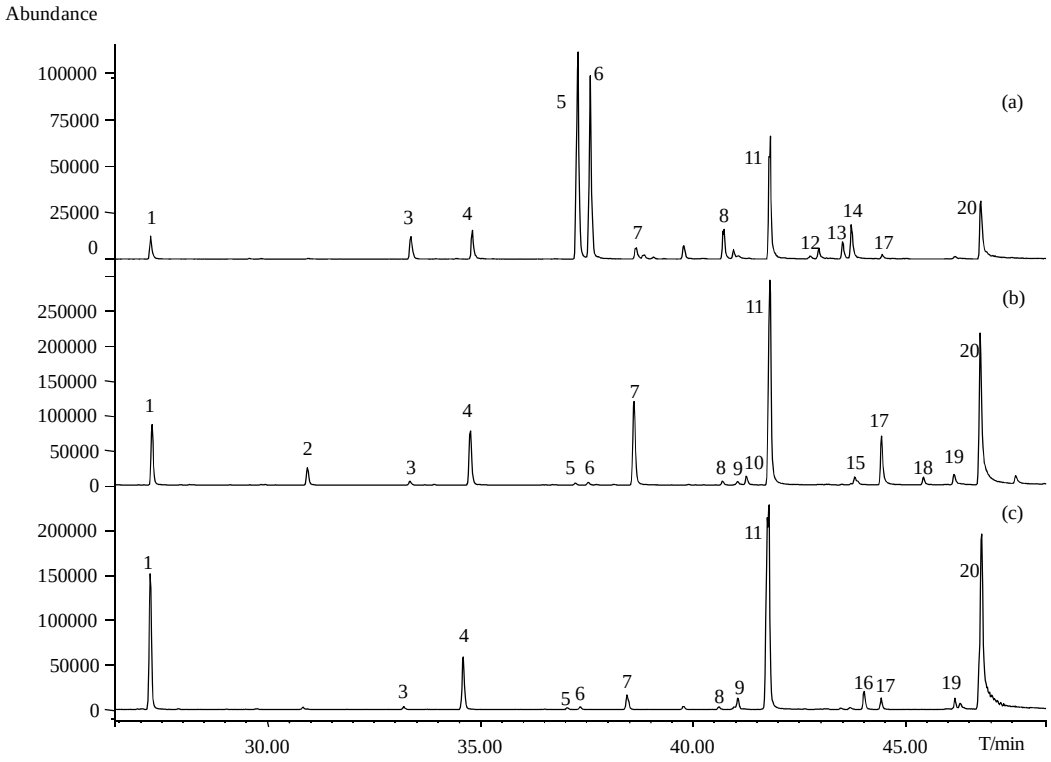

Figure B2-8 Comparison between pyrograms (SIM: *m/z*=74) for the samples of a) *Bacillus*; b) *Rhodococcus*; c) *Pseudomonas*

The three bacterial samples from three different genera show very different chromatographic profiles.

The genus Bacillus is characterized by the presence of branched isomers of fatty acids.

The bacteria of the genus Pseudomonas have mostly linear and unsaturated fatty acids; the hallmark of genus is the high relative intensity of the peak corresponding to dodecanoic acid.

Finally, the bacterial samples belonging to the genus Rhodococcus have a fatty acid profile characterized by different unsaturated compounds and linear isomers with very high relative intensity.

B2-5.4 Distinction between bacterial species of Bacillus

Three bacterial species (*Subtilus, Licheniformis, Firmus*) belonging to the genus Bacillus were analyzed by the system Py-GC-MS (Figure B2-9; Table B2-I).

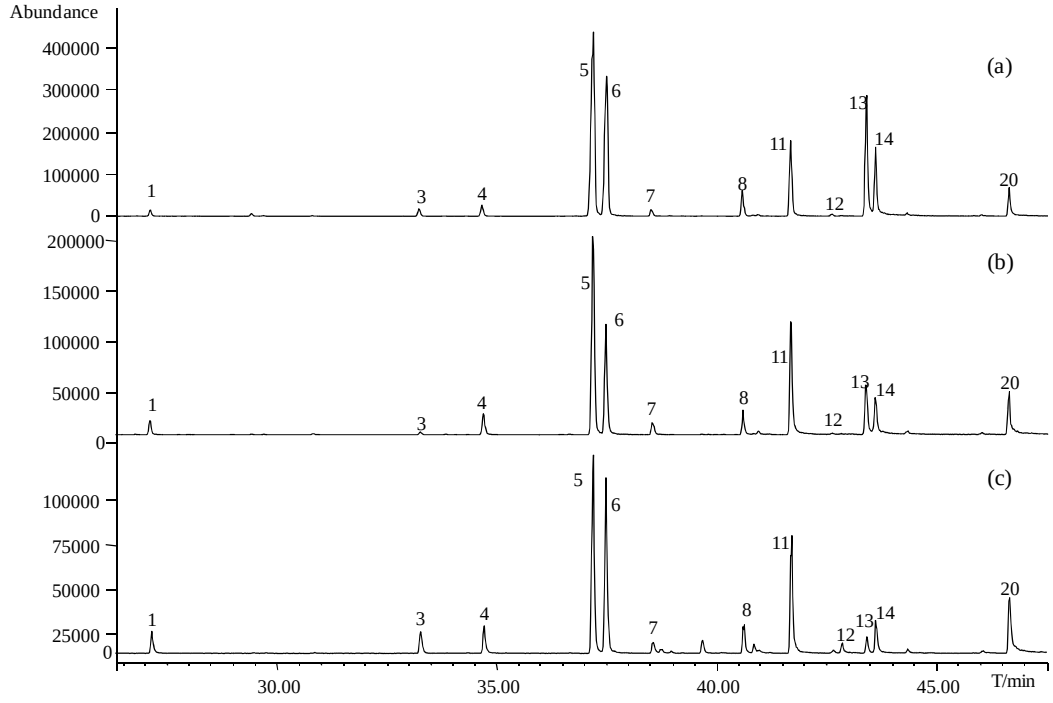

Figure B2-9 Comparison between pyrograms (SIM: *m/z*=74) for three species of *Bacillus*
a) *B. Subtilus*; b) *B. Licheniformis*; c) *B. Firmus*

It is evident that the fatty acids pattern is the same for each bacterial species of Bacillus.
The differences between the pyrograms are related to the relative intensity of the peaks, but are not so obvious as to allow a visual distinction between three different samples.
In this case it will be necessary to use a different approach to be able to differentiate between samples at the species level.

B2-5.4.1 Chemometric analysis

Multivariate data analysis proved to be the best tool to highlight differences between the bacterial species here analyzed. The starting dataset comprises 23 objects, divided into 3 classes which were known *a priori*: Table B2-II shows the typical setting of the input data

matrix. The variables (*var*) that are exploited for the chemometric analysis are ratios between the most significant chromatographic peak areas, and by values of normalized areas.

Table B2-II

VARIABLES

OBJECTS	*Sample*	*Var1*	*Var2*	…	*VarA*	*VarB*	…	*Class*
	ND1							**1**
	…							**1**
	GLA							**1**
	…							**1**
	GLC							**1**
	…							**1**
	ND2							**2**
	…							**2**
	GLB							**2**
	…							**2**
	ND3							**3**
	…							**3**
	N3							**3**
	…							**3**

B2-5.4.1.1 Principal Components Analysis

The principal component analysis was performed using the software SCAN ®.

After carrying out an analysis of the PC considering all the variables, only some of the original variables were selected (Table B2-VII).

Table B2-VII

VARIABLES			
a	**Peak 20 / Peak 13**	**e**	**Peak 7 / Peak 5**
b	**Peak 13 / Peak 8**	**f**	**Peak 7 / Peak 6**
d	**Peak 6 / Peak 5**	**n**	**Peak 19 / Peak 17**
(*) *Normalized to the sum of the peaks considered*			

The *loading plot* obtained by carrying out the PCA with these variables is shown in Figure B2-10. The variables are arranged in three groups: the first group consists of {d, e}, the second by {a, f, n}, {b} be the third. All the selected variables have a significant relevance in the principal components plotted.

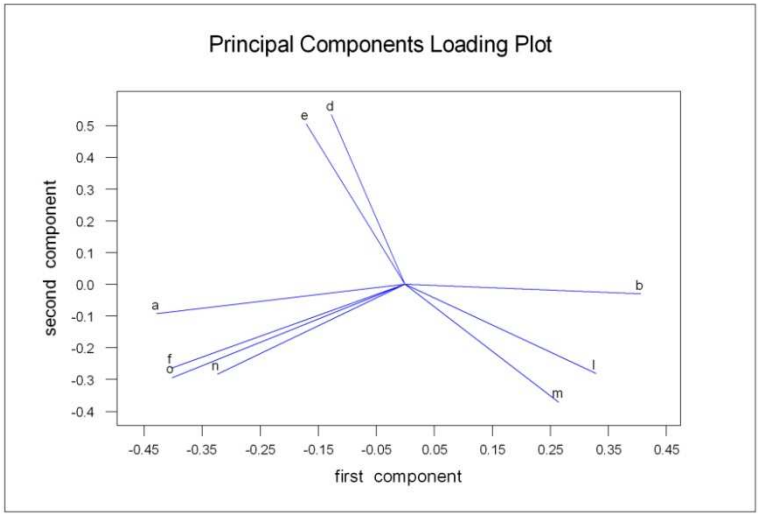

Figure B2-10 *Loading plot* relevant to Table B2-VII

In figure B2-11 the corresponding *score plot* is reported.

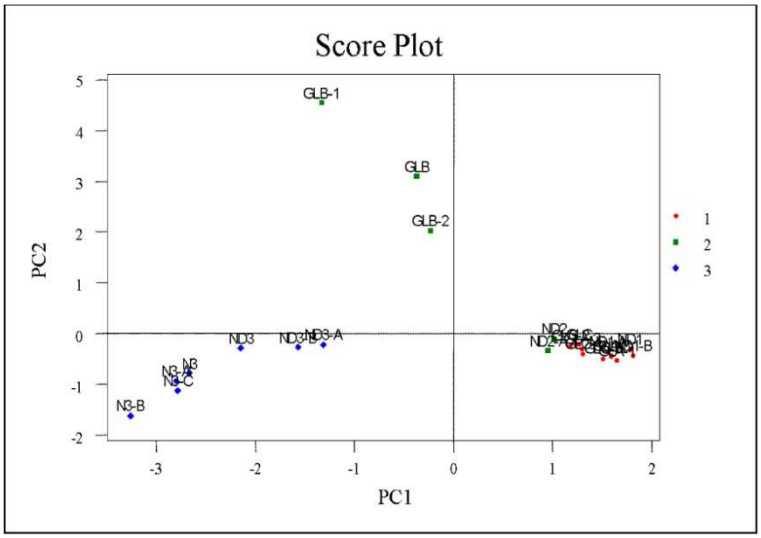

Figure B2-11 *Score plot* relevant to Table B2-VII

It's evident that objects belonging to classes **1** and **3** (*Subtilus* and *Firmus*) form very compact groups and well separated from each other: the elements of group 1 are located in the fourth quadrant, while the objects of the group **3** fall in the lower left. The elements belonging to the class **2** (*Licheniformis*), however, show great variability and are not so close as to constitute a well-defined group: two objects of this class are closer to group **1**. To improve the results other variables were chosen (Table B2-VIII)

Table B2-VIII

VARIABLES			
a	**Peak 20 / Peak 13**	**d**	**Peak 6 / Peak 5**
b	**Peak 13 / Peak 8**	**f**	**Peak 7 / Peak 6**
c	**Peak 8 / Peak 5**	**13**	**Area Peak 13***
(*) *Normalized to the sum of the peaks considered*			

In the Figures B2-12 and B2-13 the *loading plot* and the *score plot* are reported.

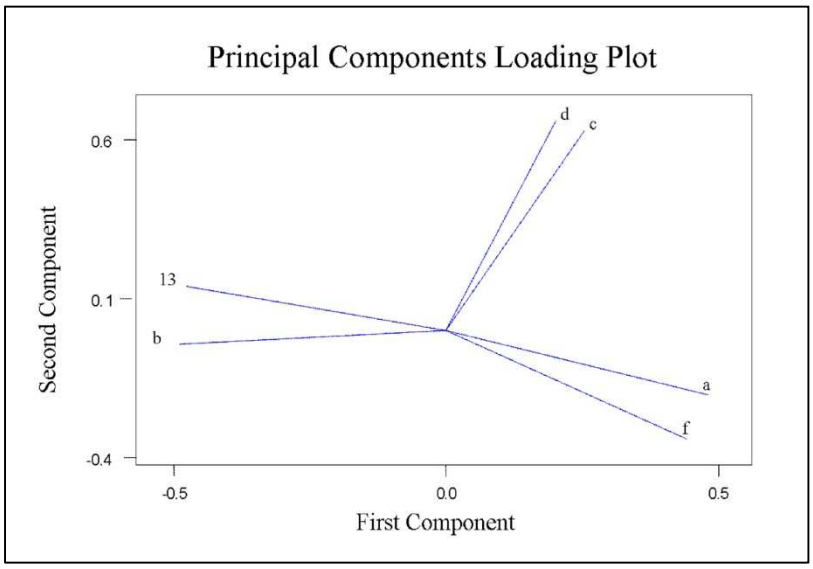

Figure B2-12 *Loading plot* relevant to Table B2-VIII

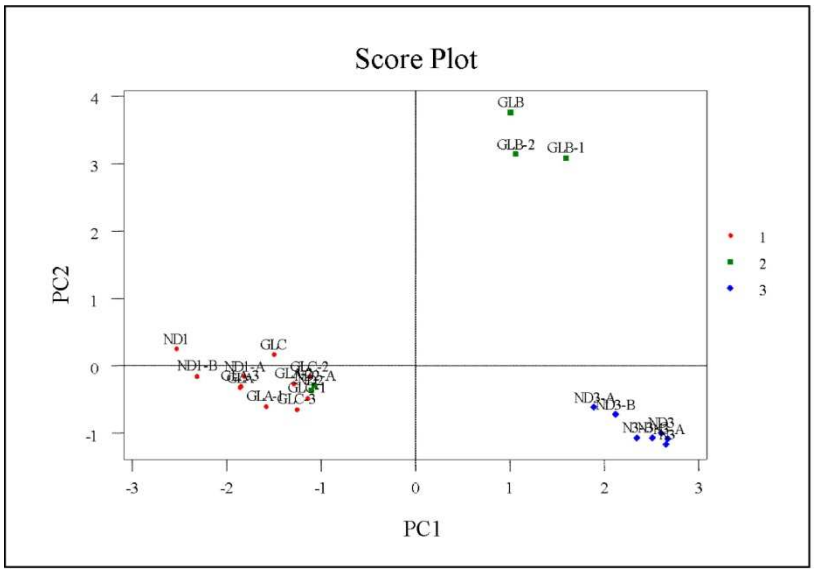

Figure B2-13 *Score plot* relevant to Table B2-VIII

The *score plot* shows a good distinction, already significant in the previous analysis, among the objects of class **1** and **3** and the large variability of the objects of class **2**.

B2-5.4.1.2 SIMCA classification

The SIMCA classification was performed by the commercial software The Unscrambler® . Variables are those reported in Table B2-VIII.

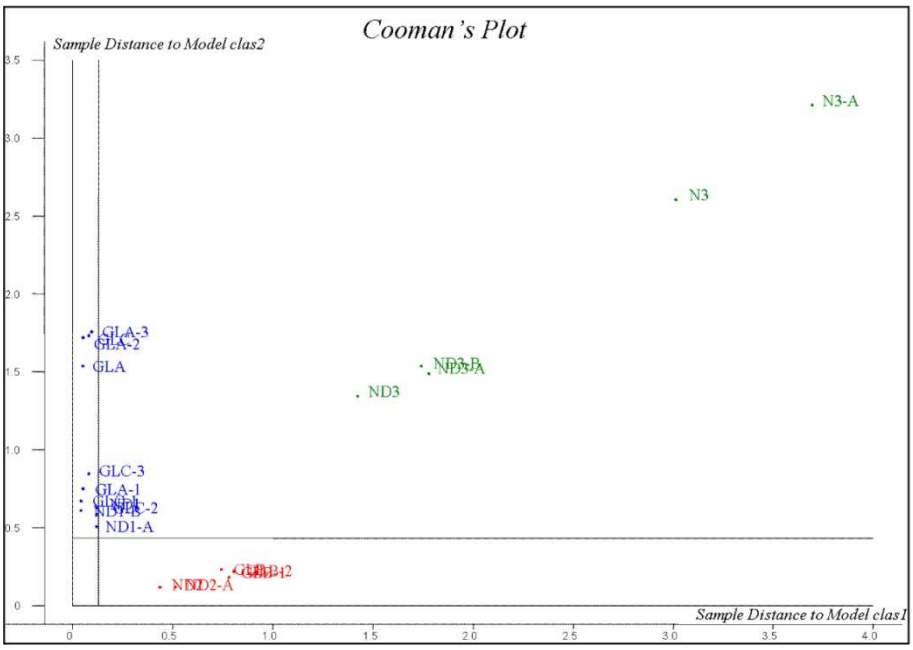

Figure B2-14 ***Coomans plot.*** **Model *'class2' vs.* Model *'class1'***
Blue: *Bacillus Subtilus*, Red: *Bacillus Licheniformis*, Green: *Bacillus Firmus*

The *Coomans plot* in Figure B2-14 shows the classification results obtained using the models built on the elements of classes **1** and **2**; objects belonging to class **1** (blue in the figure) are placed in the second quadrant of the graph and are then recognized by **model 1**, the elements of class **2** (red in figure) are assigned to the **model 2** and are therefore subject to the fourth quadrant. On the contrary, the objects of class **3** (green in the figure), that are not assigned to any of the two models, fall in the first quadrant of graph.

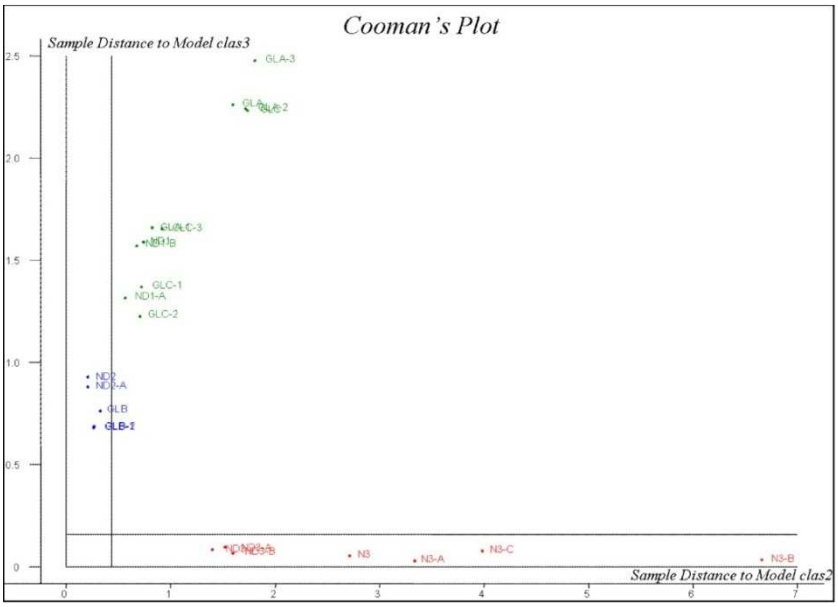

Figure B2-15 ***Coomans plot.*** **Model *'class3' vs.* Model *'class2'***
Blue: *Bacillus Subtilus*, Red: *Bacillus Licheniformis*, Green: *Bacillus Firmus*

The *Coomans plot* in Figure B2-15 shows the classification results obtained by considering the models built on the objects belonging to classes **2** and **3**; in this case, the elements of the classes **2** (blue) and **3** (red) are recognized by the **models 2** and **3**, respectively, while the objects of class **1** (green) are positioned in the first quadrant, and are not assigned to any group.

The above discussed *Coomans plots* confirm that the examined bacteria are correctly classified. This is a very important result, because it indicates that proper original variables were chosen, and the multivariate classification of bacteria based on these variables is validated.

B2-5.5 Distinction between bacterial species of Pseudomonas

Six different species of genus *Pseudomonas* were analyzed (Table B2-IX).

Table B2-IX

Genus	Species	Sample
Pseudomonas	*Fluorescens*	*BPSF*
Pseudomonas	*Putida*	*BPSPT*
Pseudomonas	*Pseudoalcaligenes*	*BPSPS*
Pseudomonas	*Stutzeri*	*BPSS*
Pseudomonas	*Stutzeri CIS/10A*	*BPSSA*
Pseudomonas	spp *Bif 7*	*BPS7*

The chromatographic profiles shown in Figure B2-16 prove the inability to distinguish the two considered species on the basis of a comparison between the FAMEs patterns.

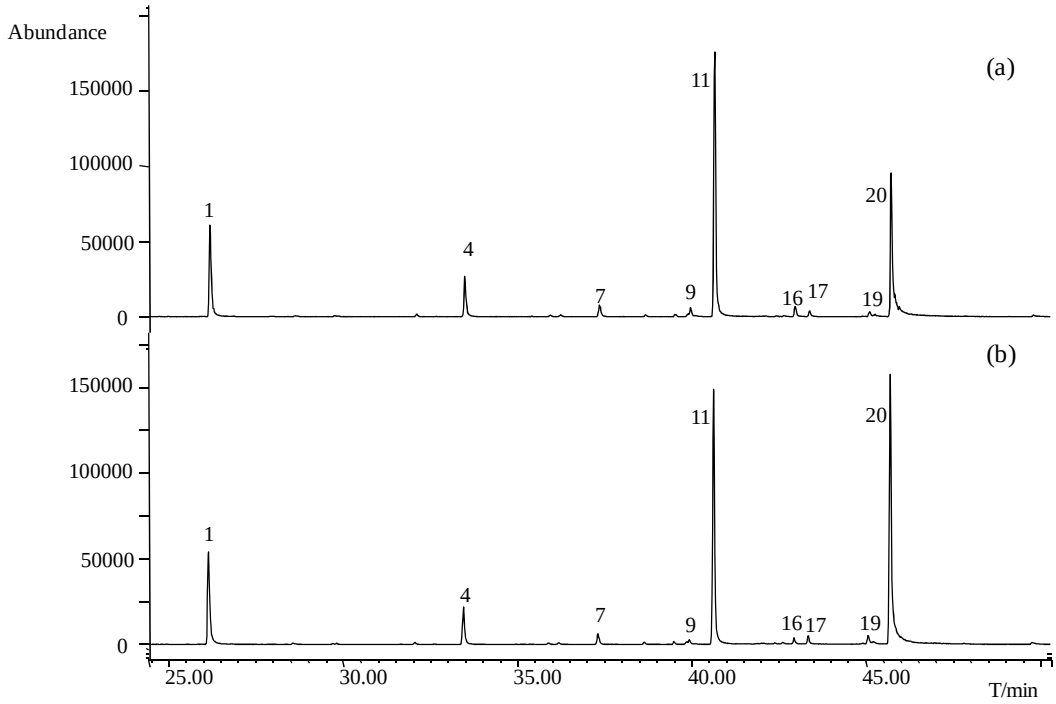

Figure B2-16 Comparison between pyrograms (SIM: *m/z*=74) for two species of Pseudomonas: a) *P. Putida*; b) *P. Fluorescens*

As in the previously examined case, it is necessary to apply *multivariate data analysis* and check if it is possible to obtain an accurate distinction even in the case of a greater number of species.

B2-5.5.1 *Chemometric analysis*

A dataset consisting of 60 objects, divided into 6 classes known *a priori* has been used the chemometric analysis of data obtained by Py-GC-MS analysis, relevant to various species of Pseudomonas (Table B2-X)

As variables, the ratios between peak areas considered most significant and normalized areas values were used.

Table B2-X

	VARIABLES							
OBJECTS	***Sample***	***Var1***	***Var2***	**…**	***VarA***	***VarB***	**…**	***Class***
	BPSF							**1**
	…							**1**
	BPSPT							**2**
	…							**2**
	BPSPS							**3**
	…							**3**
	BPSS							**4**
	…							**4**
	BPSSA							**5**
	…							**5**
	BPS7							**6**
	…							**6**

B2-5.5.1.1 *Principal Components Analysis*

The principal component analysis was performed using the software SCAN ®.

After carrying out an analysis of the PC considering all the variables, only some of the original variables were selected (Table B2-VII).

Table B2-XI

VARIABLES					
c1	**Peak 16 / Peak 13**	**g2**	**Peak 7 / Peak 3**	**C15 iso1**	**Area Peak 5***
d2	**Peak 17 / Peak 8**	**i**	**Peak 14 / Peak 13**	**C17 iso 1**	**Area Peak 13***
e2	**Peak 8 / Peak 7**	**l2**	**Peak 13 / Peak 1**	**C17 ins**	**Area Peak 16***
f	**Peak 6 / Peak 5**	**C12**	**Area Peak 1***		
(*) Normalized to the sum of the peaks considered					

In Figures B2-17 and B2-18 the corresponding *loading plot* and *score plot* are shown.

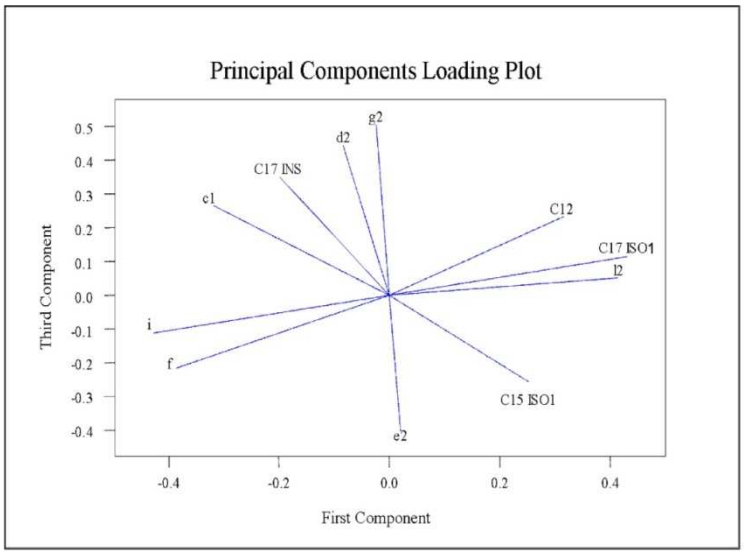

Figure B2-17 ***Loading plot***

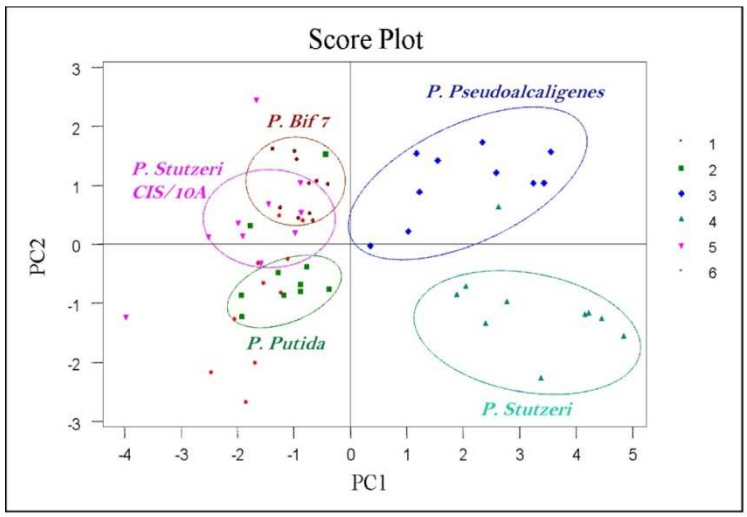

Figure B2-18 ***Score plot***

In the *score plot* some clusters of bacteria are highlighted: the objects belonging to classes **3** (*Pseudoalcaligenes*) and **4** (*Stutzer*i) are well separated from each other and from the other samples, falling in two different quadrants.

As the groupings relative to the other species of Pseudomonas, it's clear that the objects of the classes **2** (*putida*), **5** (*stutzeri CIS/10A*) and **6** (*Bif 7*) form quite compact groups but not well separated from each other.

Objects of class **1** (*fluorescens*), finally, show a great variability.

B2-5.5.1.2 *SIMCA classification*

The SIMCA classification was performed by the commercial software The Unscrambler® .

The Figure B2-19 shows the *Coomans plot* relating to PCA models constructed on the **class 3** and **class 4**.

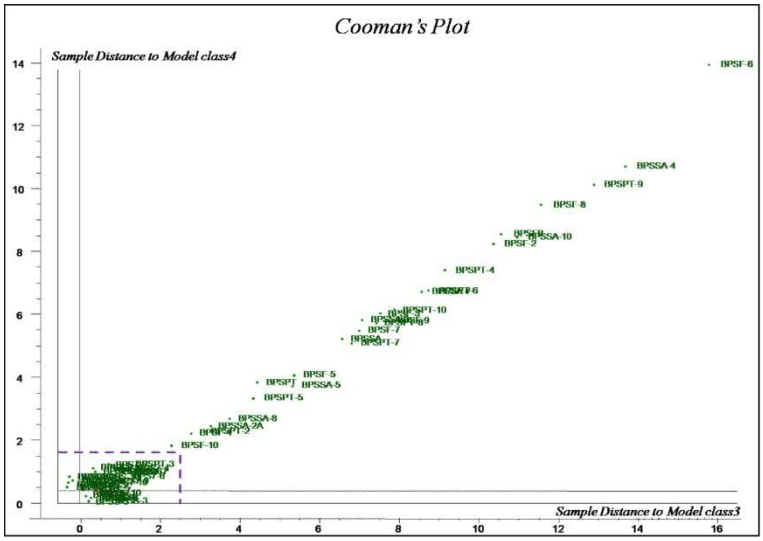

Figure B2-19 *Coomans plot.* Model *'class4' vs.* Model *'class3'*

The more evident result is the fact that most of the objects are not recognized by any model, and are then positioned in the first quadrant. The result of classification of objects of class **3** and **4** can be better evidenced only considering the shaded area (Figure B2-20).

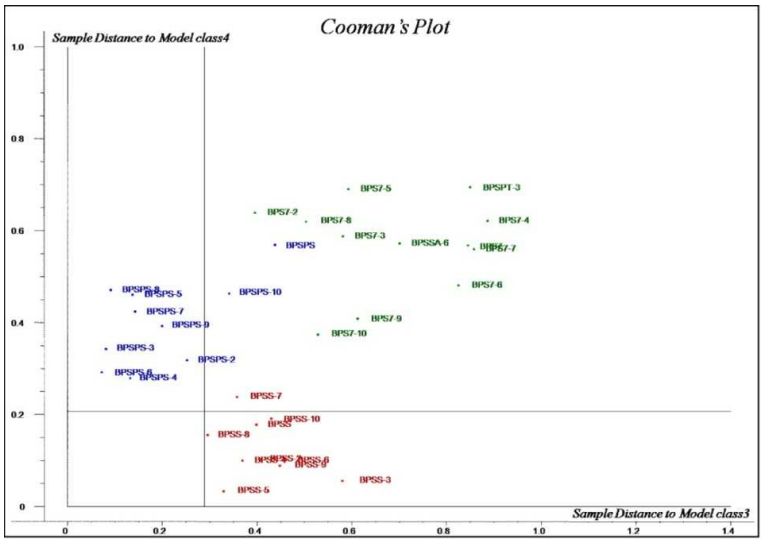

Figure B2-20 Coomans *plot.* Model *'class4' vs.* Model *'class3'*

It's evident that the objects of class **3** (blue) fall in the second quadrant of the graph and are then recognized by **model 3**, the objects of the class **4** (red) are assigned to **model 4** and fall within the fourth quadrant. The other objects (green), as already mentioned above, are not recognized by any of the two models and then fall in the first quadrant.
The obtained results confirm that the chemometric procedure validated in the case of *Bacillus*, works for *Pseudomonas*, too.

B2-5.6 Effect of culture conditions: analysis of Rhodococcus Aetherovorans BCP1

Once validated the chemometric procedure for multivariate classification, an application to the study of the effect of culture conditions on bacteria characteristics was tried.
Seven samples of Rhodococcus *Aetherovorans BCP1* grown in presence of various carbon sources were analyzed to check for possible differences in the fatty acids profile derived from the cell wall (Table B2-XII).

Table B2-XII

NAME	GENUS	SPECIES	STRAIN	CARBON SOURCE
BR1C12	***Rhodococcus***	***Aetherovorans***	***BCP1***	**C12**
BR1C16	***Rhodococcus***	***Aetherovorans***	***BCP1***	**C16**
BR1C20	***Rhodococcus***	***Aetherovorans***	***BCP1***	**C20**
BR1PEN	***Rhodococcus***	***Aetherovorans***	***BCP1***	**1-Pentanol**
BR1TOL	***Rhodococcus***	***Aetherovorans***	***BCP1***	**Toluene**
BR1SUC	***Rhodococcus***	***Aetherovorans***	***BCP1***	**Succinate**
BR1ESA	***Rhodococcus***	***Aetherovorans***	***BCP1***	**Hexane**

In Figure B2-21, a comparison between pyrograms (SIM: *m/z*=74) for the samples of *R.Aetherovorans* BCP1 grown with different source of carbon (hexane, toluene and C20) is reported (Table B2-I). The chromatographic profiles do not show substantial differences to allow any differentiation of the samples analyzed. It is therefore appropriate to apply the validated *multivariate data analysis* to check for differences between the bacterial samples analyzed.

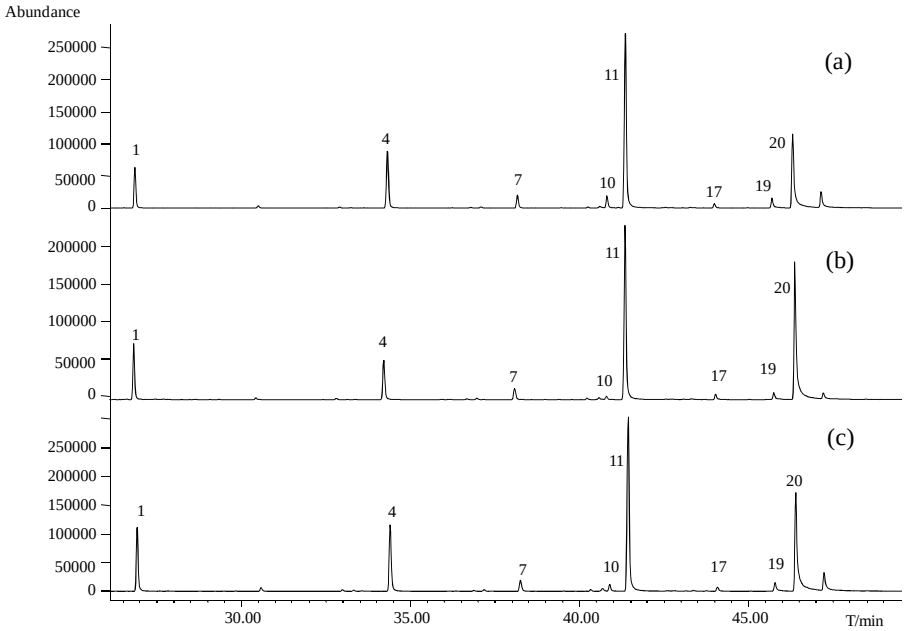

Figure B2-21 Comparison between pyrograms (SIM: *m/z*=74) for the samples of *R.Aetherovorans BCP1* grown with different source of carbon: a) hexane, b) = toluene c) C20

B2-5.6.1 Chemometric analysis

The dataset includes 51 objects that are divided into 7 groups (Table B2-XIII).

The chosen variables consist of specific ratios between the areas of chromatographic peaks corresponding to the most characteristic FAME and area values relating to the same peaks (each area value is normalized to the sum total of the areas).

Table B2-XIII

VARIABLES

OBJECTS	*Sample*	*Var1*	*Var2*	…	*VarA*	*VarB*	…	*Class*
	BR1C12							**1**
	…							**1**
	BR1C16							**2**
	…							**2**
	BR1C20							**3**
	…							**3**
	BR1ESA							**4**
	…							**4**
	BR1PEN							**5**
	…							**5**
	BR1SUC							**6**
	…							**6**
	BR1TOL							**7**
	…							**7**

B2-5.6.1.1 *Principal Components Analysis*

The principal component analysis was performed using the software SCAN ®.

After carrying out an analysis of the PC considering all the variables, only some of the original variables were selected (Table B2-XIV).

Table B2-XIV

VARIABLES			
a2	**Peak 20 / Peak 15**	**n2**	**Peak 19 / Peak 17**
f	**Peak 7 / Peak 6**	**o1**	**Peak 9 / Peak 17**
g4	**Peak 10 / Peak 19**	**p**	**Peak 1/ Peak 2**

The *loading plot* obtained with these variables is shown in Figure B2-22.

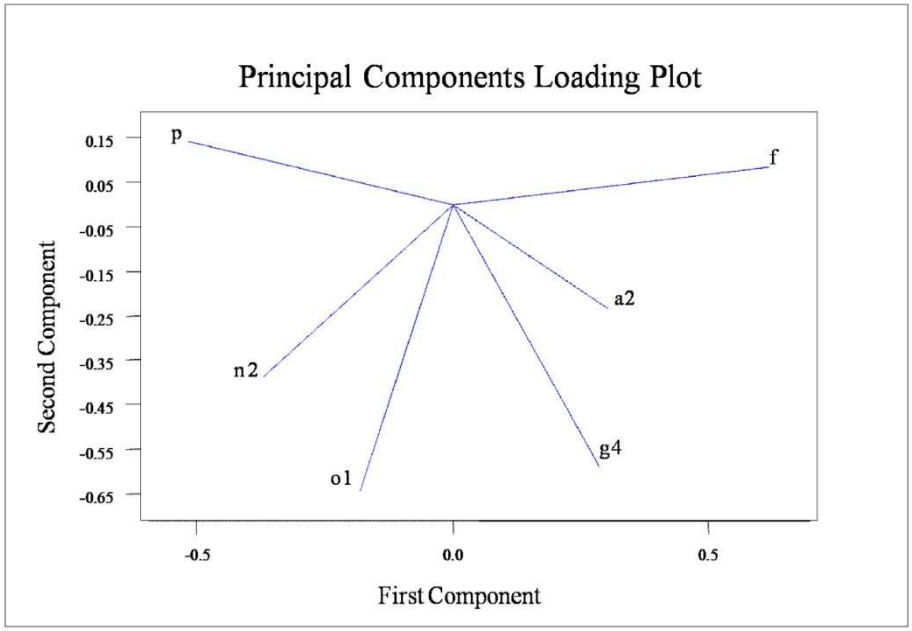

Figure B2-22 Loading plot

The selected variables show information and relevance different in the *principal components* considered: variables *f* and *p* are most relevant for the first principal component, while the variables *g4*, *o1* and *n2* for the second principal component.

The corresponding *score plot* is shown in Figure B2-23.

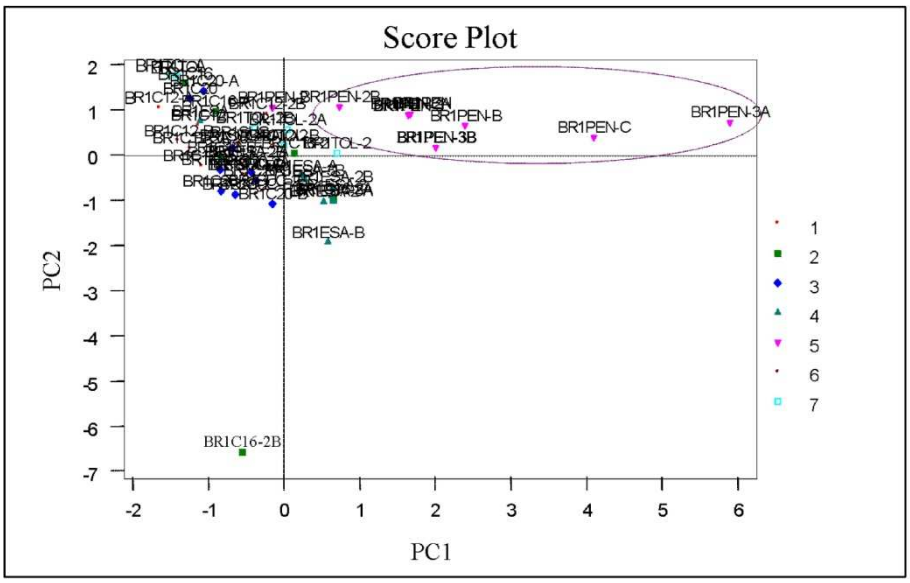

Figure B2-23 Score plot

In the *score plot*, a very compact group of points near to the axes origin is present. Moreover, it's evident a series of objects (indicated by the pink color), distant from the origin and characterized by a great variability within the belonging group.

Further information that can be easily obtained from this graph is the presence of an *outlier*, an object considerably far from all others.

The next step was to perform the PCA using, as variables, both the ratios of the peak areas and the values of the normalized areas.

Table B2-XV

VARIABLES			
g3	**Peak 9 / Peak 19**	**C17**	**Area Peak 17***
n2	**Peak 19 / Peak 17**	**C18**	**Area Peak 20***
o1	**Peak 9 / Peak 17**		
(*) Normalized to the sum of the peaks considered			

The *loading plot* and the *score plot* obtained with these new variables are shown in Figures B2-24 and B2-25.

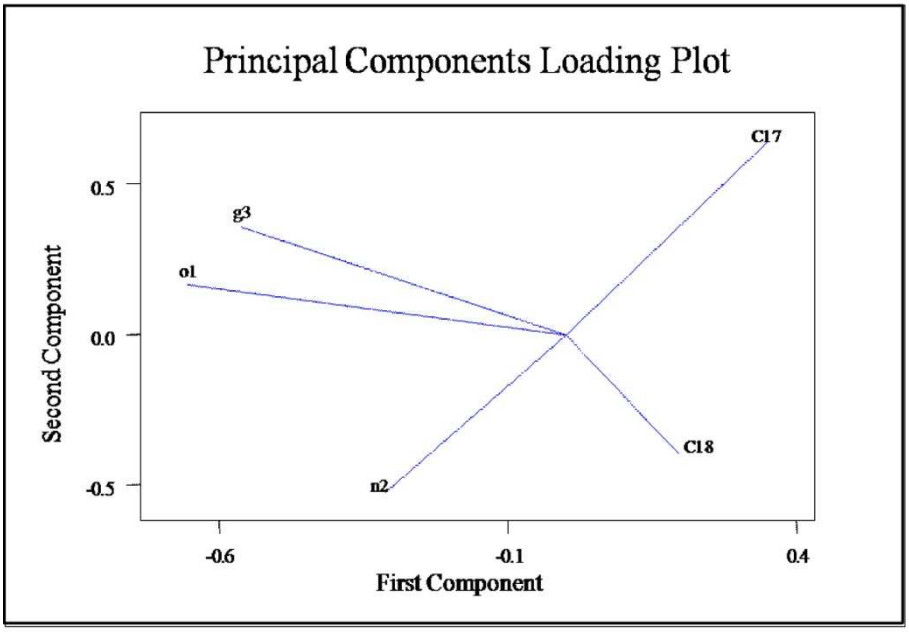

Figure B2-24 Loading plot

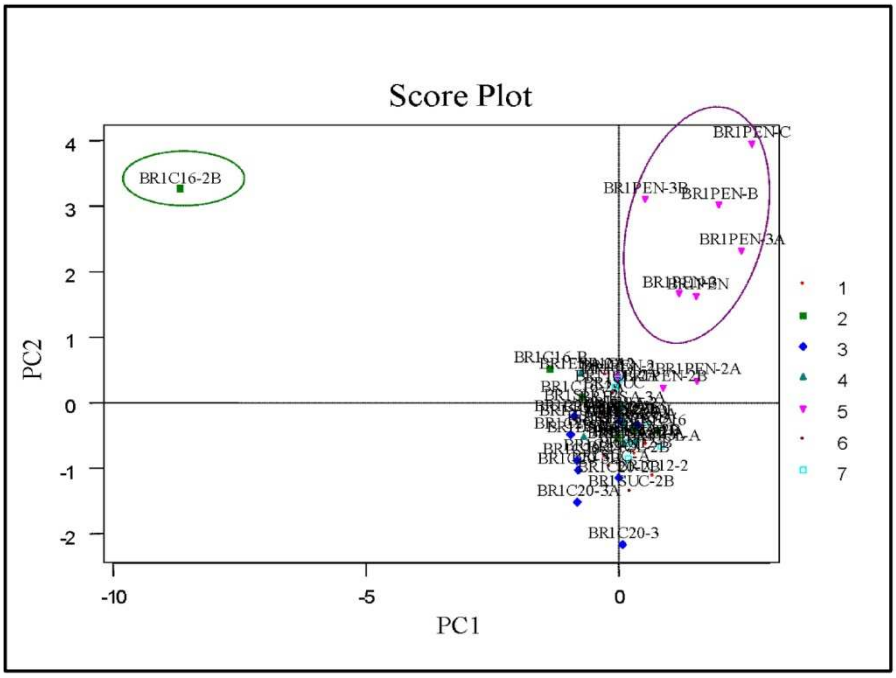

Figure B2-25 Score plot

The results obtained from this second PCA are consistent with the previous and reveal greater the topics discussed first before: the character of *outlier* of the samples belonging to class **2** (marked green) is enhanced and the separation between the samples of class **5** (pink colored) and all other objects is increased.

B2-5.6.1.2 SIMCA classification

The SIMCA classification was performed by the commercial software The Unscrambler® .

The Figure B2-26 shows the *Coomans plot* relating to PCA models constructed on the class, known *a priori*, Hexane and class ***Toluene.*** For the PCA modeling the variables described in Table B2-XV were used.

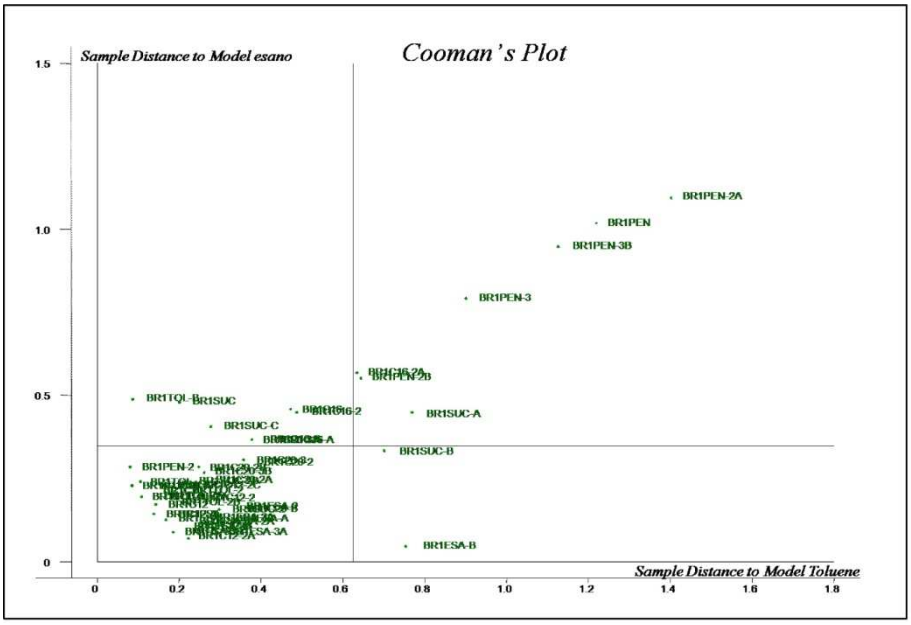

Figure B2-26 Coomans plot. Model *'hexane' vs Model 'toluene'*

It's evident that almost all objects corresponding to the bacterial samples grown in the presence of 1-pentanol are not assigned to any of the two classes chosen and fall in the first quadrant; on the contrary, all other objects are placed in the third quadrant and are then described by both models selected.

The results validate the observation that all the bacterial samples do not show a structural difference dependent on the medium used for growth with the exception of those cultivated in the presence of 1-pentanol. A further test can be obtained by selecting as one of the two models PCA the class of bacteria grown in the presence of 1-pentanol, the class, *i.e.*, showing a behavior distant from the others (Figure B2-27).

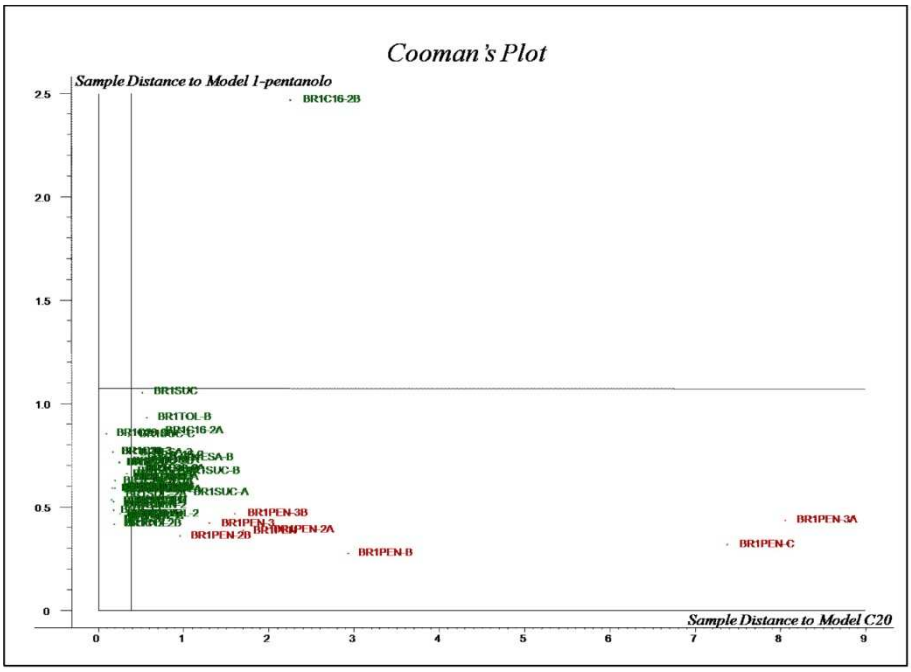

Figure B2-27 Coomans plot. Model *'1-pentanol' vs Model 'C20'*

The objects related to the bacteria grown in the presence of 1-pentanol are correctly recognized by the corresponding model while all the other samples fall in the third quadrant of the graph and, therefore, are still recognized once by both models.

B2-6 Conclusions

In this work two lines of investigation were followed, both on the analysis of bacterial samples by using the analytical integrated system Py-GC-MS and on multivariate data analysis.

In the first part the issue of bacterial identification was tackled; choosing appropriate markers it was possible to implement the differentiation of bacteria at the genus level, performing a simple comparison between the chromatographic profiles obtained. The application of chemometric tools such as *Principal Component Analysis* (PCA) and the *Soft Independent Modeling of Class* (SIMCA) allowed a further differentiation at the level of bacterial species. The results obtained showed that the analytical pyrolysis, coupled to gas chromatography and mass spectrometry, is a suitable technique for the analysis of bacterial samples as regards the identification and differentiation, significantly reducing the pre-treatment of samples requested by the recognition techniques commonly used.

In the second part the study was directed on samples belonging to the same bacterial strain, but subjected to different growth conditions. The data obtained by the Py-GC-MS analysis and their exploration performed by chemometric analysis led to the conclusion that the fatty acid composition of bacterial cell wall is the same in almost all samples analyzed, with the exception of the case of the samples grown in the presence of 1-pentanol, which show a behavior significantly different with respect to all other bacteria analyzed.

CHAPTER B3

CHARACTERIZATION OF A MIXED-BIOFILM BY MEANS PY-GC-MS

B3-1 Summary

Biofilms are communities of microorganisms adhering to a surface and embedded in an extracellular polymeric matrix. Biofilms are frequently associated with disease and contamination, but they are also important for engineering applications, such as biocatalysis and bioremediation.

In the natural environment most of the biofilms are constituted by both eukaryotic and prokaryotic microbial species. In this context, a mixed biofilm formed by bacteria and fungi may provide an optimal habitat for addressing contaminated areas and at the same time heavy metal and organic pollutants. To exploit the potential of natural microbial communities consisting of bacteria and fungi, it is essential to understand and control their formation.

In this work a characterization of a mixed-biofilm consisting of a bacterium (*Pseudomans Mendocina*) and a fungus (*Pleorotus ostreatus*) by means Py-GC-MS is reported. In situ Thermal Hydrolysis and Methylation (THM) was applied. The results were exhaustive for the characterization and differentiation of biofilms. The methylated fatty acids (FAMEs) were used as *markers* for bacterium identification. A characteristic pyrographic pattern deriving from the pyrolysis of polysaccharides present in great amount in wall cell was used as a *marker* for fungus.

B3-2 Keywords

Mixed-Biofilm

Bacteria

Fungi

Pyrolysis

Gas-Chromatography-Mass Spectrometry

B3-3 Introduction

Biofilms are differentiated groups of microorganisms (e.g. bacteria and fungi) arranged as aggregated structures which constitute a unique mode of growth that allows survival in hostile environments [60]. Their formation is an ordered process, complex and not fully understood, that is dependent on the response of the cell to environmental cues, which in turn regulates specific genes. Stages of biofilm formation include motility to the surface, attachment, formation of clusters, development of differentiated structures, and dispersal. It consists of water, microbial cells and a wide range of self-generated extracellular polymeric substances

(EPS) referred to as the matrix [61]. Biofilms are important contributors to the state of human disease, considering most of bacterial chronic inflammatory and infectious diseases involve biofilms [62-64]. They are frequently associated with biofouling and biocorrosion [65]. However, due to their ability to hold out the chemical and physical stresses, they are exploited for many beneficial applications of biotechnology.

Biofilms have been important for wastewater treatment [66-67], for decreasing biocorrosion [68], as well as for other applications, such as biocatalysis [69], which includes the production of biofuels, specialty/bulk chemicals, biologics, and food additives for the important process of bioremediation. In bioremediation applications, microorganisms are used to remove, detoxify, or immobilize pollutants, without harmful chemicals [70-73].

To take advantage of the potential of biofilms, their formation must be studied and controlled. Most studies on biofilm examine monospecies cultures, whereas nearly all biofilm communities in nature comprise a variety of microorganisms. The species that constitute a mixed-biofilm and the interactions between these microorganisms critically influence the development and shape of the community [74].

Interactions among species within a biofilm can be antagonistic, such as competition over nutrients and growth inhibition, or synergistic, and may occur both among different bacterial species and between bacteria and fungi. The latter can result in the development of several beneficial phenotypes such as biofilm formation by co-aggregation, metabolic cooperation where one species utilizes a metabolite produced by a neighboring species, and increased resistance to antibiotics or host immune responses compared to the mono-species biofilms. These beneficial interactions in mixed-biofilms have important environmental, industrial and clinical implications.

The purpose of this work was to verify the possibility of characterizing a mixed-biofilm, consisting of a bacterium (*Pseudomans Mendocina*) and a fungus (*Pleorotus ostreatus*), by means of the integrated system pyrolysis-gas chromatography/mass spectrometry [75,76,56,58,77]. Bacterium and fungus have been grown as microbial biofilms using the MBEC Assay™. Pyrolysis-methylation was applied to three different biofilms: the first deriving from bacterium, the second from fungus and the third consisting of bacterium and fungus grown together. The chromatographic profiles were found to be exhaustive for the characterization and differentiation of biofilms. The characterization of the bacterium is carried out according to the profile of the methylated fatty acids (FAMEs) [5]; as regards

fungus, the relevant presence of polysaccharides in the wall cell produces, under these experimental conditions, a characteristic pyrographic pattern. The biofilms samples were provided by Dr. Stefano Fedi (Department of Experimental Evolutionary Biology (BES), University of Bologna (Italy)), who studied the application of biofilm in bioremediation for a long time.

B3-4 Materials and methods

B3-4.1 Preparation of microorganisms for the growth in biofilms

For the bacterial biofilm the following procedure was performed:

- first smear of bacterial strain on agar plates with medium LB (*Luria-Bertani Broth*) (Table B3-I)
- second smear of bacterial strain on agar plates with medium LB (*Luria-Bertani Broth*) (Table B3-I)
- a part of biomass was recovered by a cotton swab and resuspended in LB liquid until reaching an optical density at 600 nm equal to 0.78.
- from this solution a dilution (1:30) was performed in LB liquid medium to obtain the suspension of inoculum.
- the suspension was used to inoculate a plate "MBEC assay"
- 150 ml of bacterial suspension were inoculated to each well
- the plate was closed with the lid and set to agitate at 150 rpm for 48 hours at 30 ° C.
- after 48 hours, a number of pins was removed with sterile forceps for subsequent analysis

Table B3-I

Luria-Bertani Broth, LB (pH 7)	
Tryptone	10 g/l
Yeast Extract	5 g/l
NaCl	10 g/l
H_2O	to volume

For the fungal biofilm was performed the following procedure:

- A plate of fungus was inoculated and grown on agar MEG with the following composition: glucose 10 g l^{-1} malt extract 5g l^{-1}
- the fungus was collected and inoculated for about a week into 100 ml of liquid MEG and grown for about 4 days at 150 rpm and 30 ° C.
- At the end of growth, the fungal biomass was repeatedly homogenized in a potter
- At the end of homogenization, the solution was used to inoculate a plate MBEC assay in liquid medium LB (*Luria-Bertani Broth*) (Table B3-I)
- The biofilm fungus was grown in agitation at 30 rpm and 30 ° C for 4 days
- Subsequently a number of pins have been removed with sterile forceps according to the procedure followed for the growth of bacterial biofilm,
- the pins were washed, dried and then stored at -20 ° C

To prepare the mixed-biofilm, the rest of the upper part of the plate containing the individual fungal biofilm was transferred into a new 96-well plate inoculated with a bacterial solution prepared as described above. The plate was subjected to 150 rpm and 30 ° C for 48 hours, and then additional pins were taken for the analysis of mixed-biofilms following a scheme similar to that previously described.

B3-4.2 *Experimental conditions*

The derivatising reagents used, tetramethylammonium hydroxide (TMAH) 25%$_{w/v}$ water solution, and the monosaccharides standard (glucose, mannose, galactose and glucuronic acid) were from Aldrich.

Samples (0.5 mg) were inserted into a quartz capillary tube and added with 5 µl of derivatising reagent prior to pyrolysis.

Pyrolysis was carried out at 600°C for 10 s at the maximum heating rate using a CDS 1000 pyroprobe heated filament pyrolyser (Chemical Data System, Oxford, USA), directly connected to the injection port of a Varian 3400 gas-chromatograph coupled to a Saturn II ion-trap mass spectrometer (Varian Analytical Instruments, Walnut Creek, USA). A Supelco SPB5 capillary column (30m, 0.25mm I.D., 0.25 µm film thicknesses) was used with a temperature programme from 50°C (held for 5 min) to 130°C at 5°C min^{-1},from 130°C to 170° C at 2,5°C min^{-1} , from 170° C to 300°C at 5°C min^{-1} with helium as carrier gas. Temperatures of split/splitless injector (split mode) and Py-GC interface were kept at 250°C.

The Py-GC interface was kept at 250°C and the injection port at 250°C. Injection mode was split (1:50 split ratio) and gas carrier was helium at flow rate of 1.0 ml min^{-1}. Mass spectra were recorded at 1 scan s^{-1} under electron impact at 70 eV, scan range 45 to 650 *m/z*. Structural assignment of the pyrolytical fragments was based on match with the NIST mass spectra library or literature data; the principal products identified are listed in Table B3- I.

Table B3-1 Products arising from pyrolysis-methylation of the samples

Peak No	Fragments	M_w	Characteristic ions *m/z* (relative abundance)[a]
1	Dodecanoic acid, methyl ester	214	74 (100), 87 (62), 143 (22), 171 (16), **214** (12)
2	Tetradecanoic acid, methyl ester	242	74 (100), 87 (67), 143 (33), 199 (20), **242** (22)
3	Pentadecanoic acid, methyl ester	256	74 (100), 87 (64), 157 (14), 213 (16), **256** (28)
4	9-hexadecenoic acid, methyl ester	268	55 (100), 152 (26), 194 (20), 236 (26), **268** (8)
5	Isomer peak 9	268	55 (100), 152 (15), 194 (10), 236 (16), **268** (5)
6	Hexadecanoic acid , methyl ester	270	74 (100), 87 (65), 143 (34), 227 (18), **270** (52)
7	11-hexadecenoic acid, 15-methyl-, methyl ester	282	55 (100), 96 (67), 213 (25), **282** (4), 283 (6)
8	Heptadecanoic acid, methyl ester	284	74 (100), 87 (68), 143 (30), 241 (16), **284** (37)
9	9-octadecenoic acid, methyl ester	296	55 (100), 180 (12), 222 (12), 264 (24), **296** (8)
10	Octadecanoic acid, methyl ester	298	74 (100), 143 (36), 199 (20), 255 (22), **298** (62)

[a] The value in bold indicates the molecular ion

B3-5 Results and discussion

B3-5.1 Analysis of bacterial and fungal biofilms

Since FAMEs were chosen as *markers* for the bacterial biofilm characterization, it was appropriate to consider the chromatographic profiles obtained in Single Ion Monitoring (SIM) at *m/z*=74. In Figure B3-1 a comparison between the pyrograms (SIM; *m/z*=74) arising from pyrolysis-methylation of bacterial and fungal biofilms is shown.

In these pyrolytic conditions, the bacterial biofilm mainly determines saturated and unsaturated linear fatty acids; another distinctive feature is the high relative intensity of the peak corresponding to dodecanoic acid. The profile of the methylated fatty acids (FAMEs) determined from the analysis of the sample of fungal biofilm is clearly different: the intensity of the peak corresponding dodecanoic acid is much lower and are almost totally absent unsaturated fatty acids .

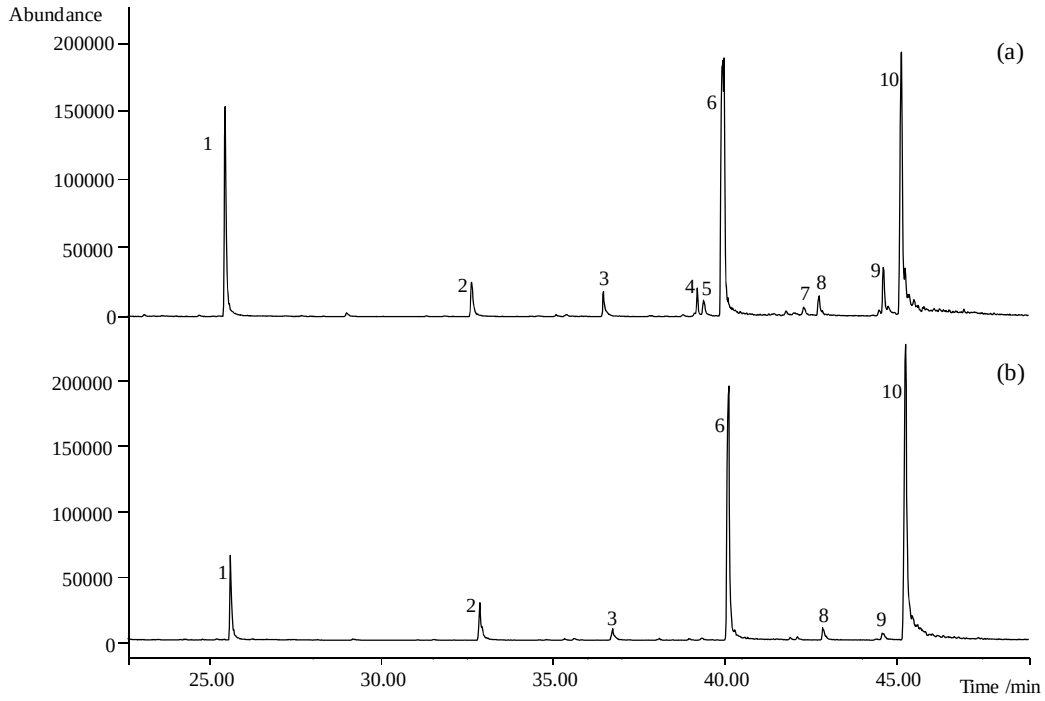

Figure B3-1 Comparison between pyrograms (SIM: *m/z*=74) for the samples of a) bacterial and b) fungal biofilms

As it is well known, the wall cell of fungi is mainly composed of polysaccharides, glucose-based; hence, a search of their characteristic *markers* was performed. There is an extensive

literature relating to the behavior of saccharidic material in these experimental conditions [78-79].

The pyrolysis-methylation of these compounds produces a pair of permethylated saccharinic acids isomers of the main product, all characterized by a mass spectra with a base peak at *m/z* 129. This result is closely related to the behavior of the carbohydrates in a basic environment and at high temperature; these factors favor their "alkaline degradation".

In Figure B3-2 (b) a Single Ion Monitoring at *m/z* 129 arising from pyrolysis-methylation of fungal biofilm is reported.

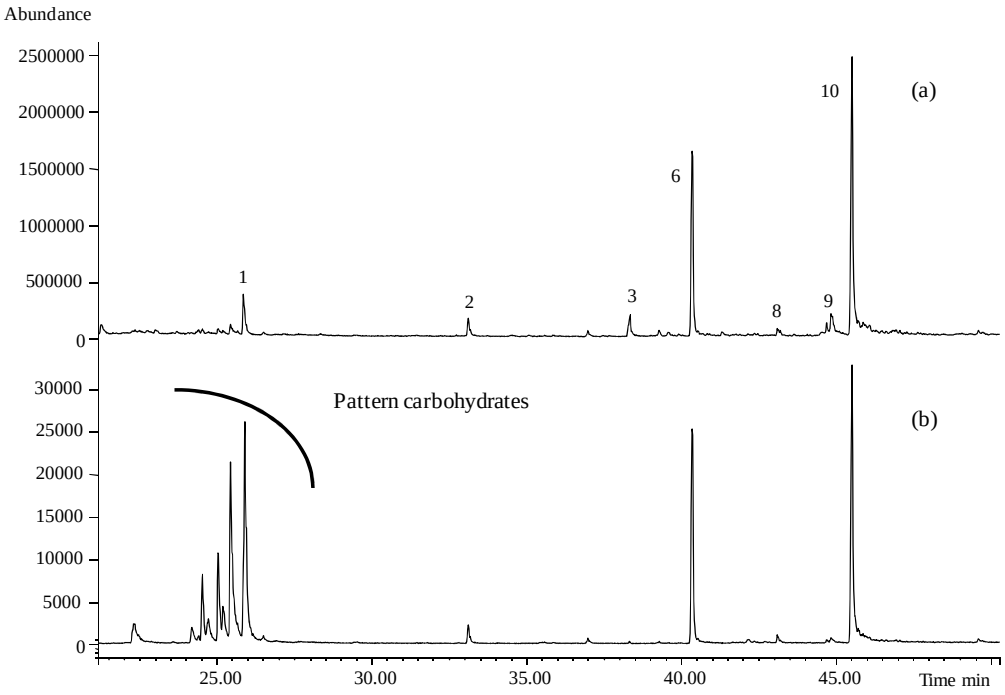

Figure B3-2 (a) Reconstructed ion chromatogram and (b) Single Ion Monitoring at *m/z* 129 arising from pyrolysis-methylation of fungal biofilm

The characteristic pattern, formed in the presence of saccharidic derivatives, has been highlighted. To determine the nature of these compound some standard monosaccharides (glucose, mannose, galactose and glucuronic acid) have been analyzed (Figure B3-3; table B3-III).

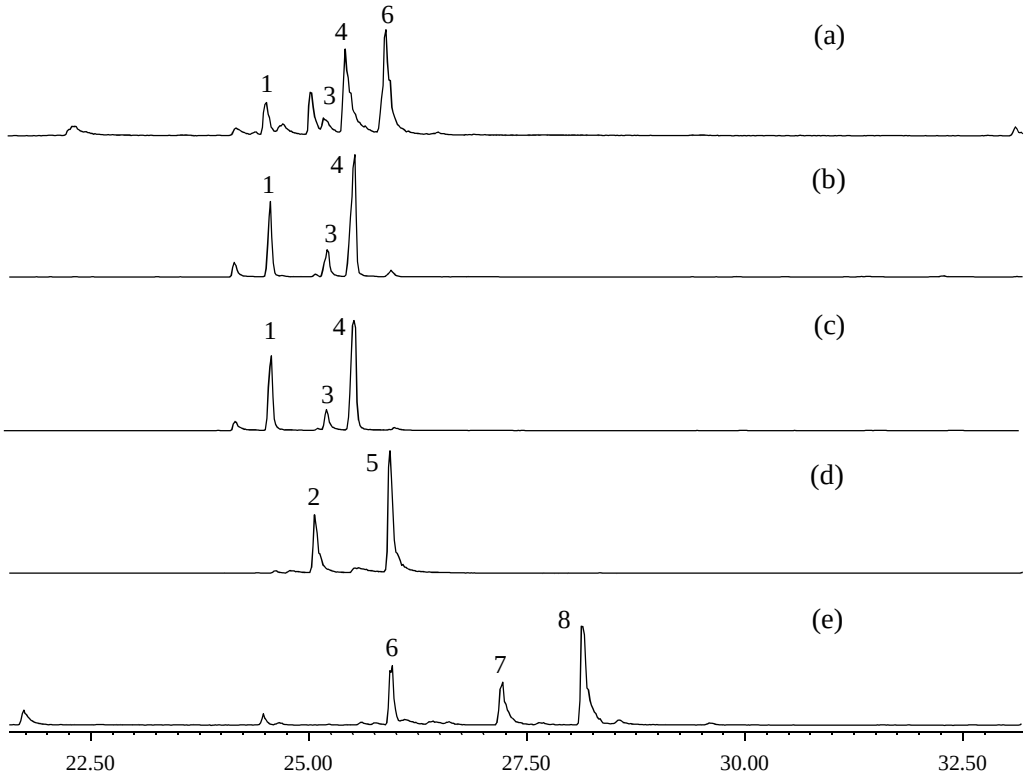

Figure B3-3 Comparison between pyrograms (SIM: *m/z*=129) for the samples of a) fungal biofilm and standard monosaccharides: b) glucose, c) mannose, d) galactose , e) glucuronic acid

Table B3-III

Peak No	Fragments	M_w	Characteristic ions *m/z* (relative abundance)[a]
1	Tetra-*O*-methyl-3-deoxy-D-*ribo*-hexonic acid methyl ester	250	75 (34), 101 (32), 129 (100), 159 (15), 187 (5)
2	Tetra-*O*-methyl-3-deoxy-D-*xilo*-hexonic acid methyl ester	250	75 (37), 101 (36), 129 (100), 159 (12), 191 (9)
3	Partially methylated compound from 1 or 3	220	59 (32), 75 (7), 101 (79), 129 (100), **220** (9)
4	Tetra-*O*-methyl-3-deoxy-D-*arabino*-hexonic acid methyl ester	250	75 (35), 101 (33), 129 (100), 159 (14), **250** (2)
5	Tetra-*O*-methyl-3-deoxy-D-*lyxo*-hexonic acid methyl ester	250	75 (40), 101 (37), 129 (100), 159 (14), **250** (2)
6	Dodecanoic acid, methyl ester	214	74 (100), 87 (62), 143 (22), 171 (16), **214** (12)
7	Tri-*O*-methyl-3-deoxy-glucaric acid methyl ester	264	75 (58), 101 (36), 129 (100), 141 (40), 173 (39)
9	Isomer of 7	264	75 (51), 101 (31), 129 (100), 141 (34), 173 (31)

[a] The value in bold indicates the molecular ion

Comparing retention times and mass spectra, a correspondence between fungal biofilm and the glucose/mannose chromatographic profiles is evident. In these experimental conditions, the two monosaccharides lead to the formation of the same isomers being the starting products epimers in C2.

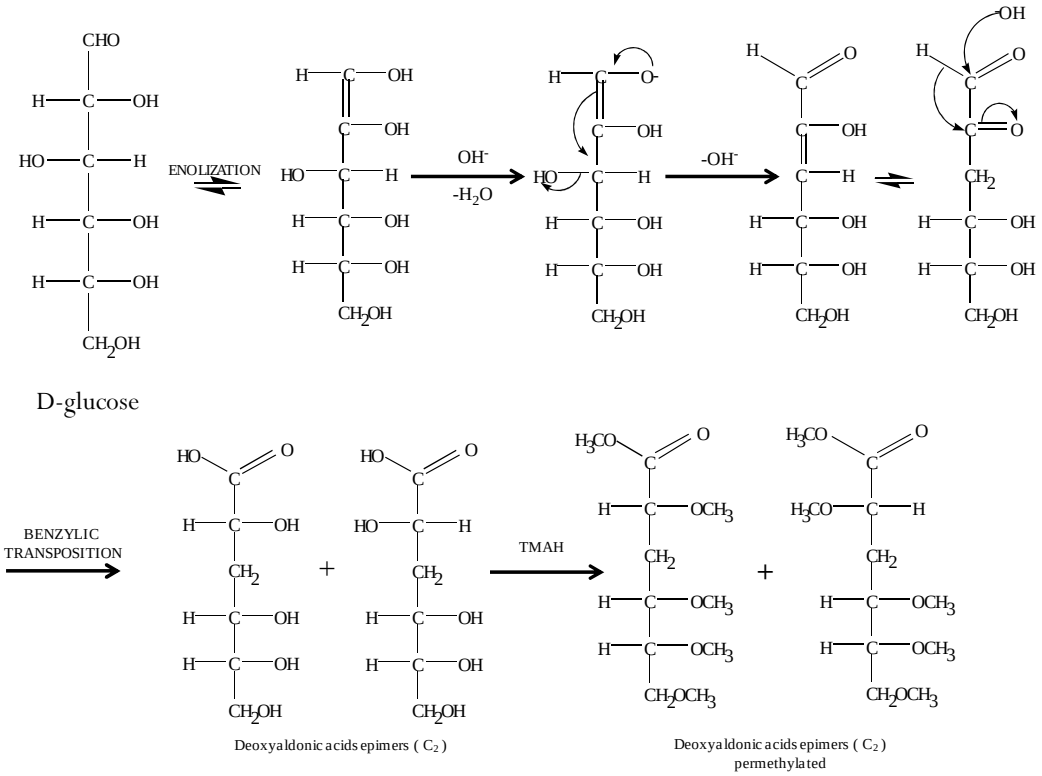

Figure B3-4 Proposed mechanism of the formation of permethylated saccharinic acids from pyrolysis- methylation of glucose (from *Ref*. [79])

The reaction includes a rapid isomerization, for enolization, and a subsequent elimination of a molecule of water which results in the loss of the stereogenic center at C3. In a strong basic environment, the intermediate α-dicarbonylic compound undergoes a rearrangement similar to the benzylic transposition, leading to deoxyaldonic acids, epimers in C2. The TMAH determines a methylation of all the hydroxyl groups (carboxylic and alcohol) forming two permethylated saccharinic acids epimers.

B3-5.2 Analysis of mixed-biofilm

In Figure B3-5 the reconstructed ion chromatogram (a), Single Ion Monitoring at *m/z* 129 (b) and (c) Single Ion Monitoring at *m/z* 74 arising from pyrolysis-methylation of mixed-biofilm are reported.

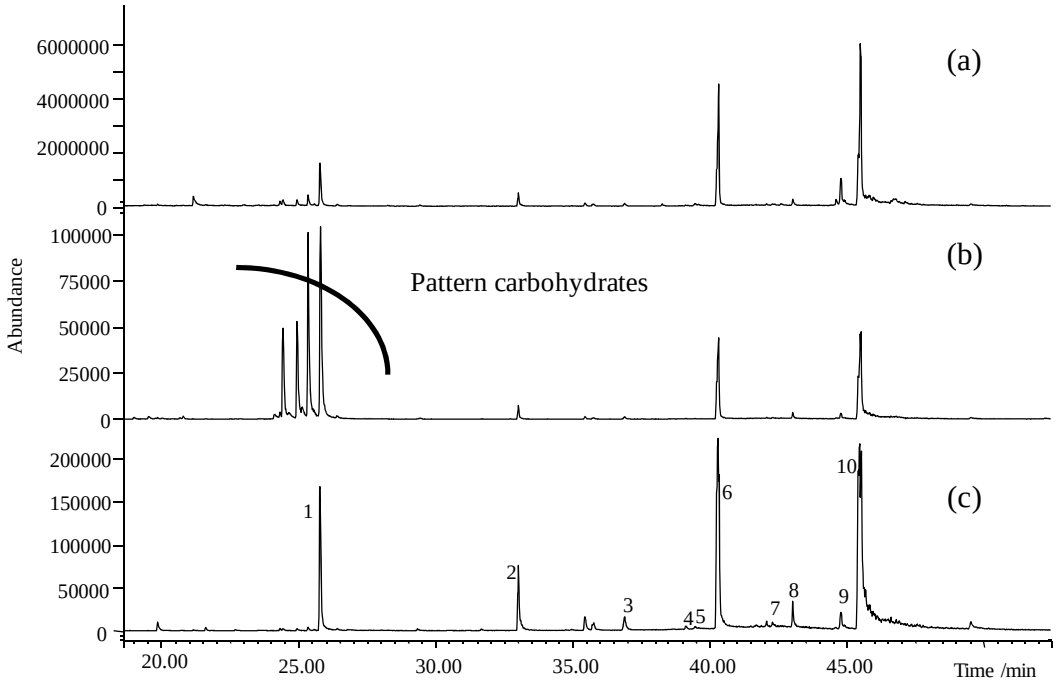

Figure B3-5 (a) Reconstructed ion chromatogram (b) Single Ion Monitoring at *m/z* 129 (c) Single Ion Monitoring at *m/z* 74 and arising from pyrolysis-methylation of mixed-biofilm

The analysis of the chromatographic profiles demonstrates the possibility of determining, at once, the presence of bacterium and fungus in the mixed-biofilm as the characteristic *markers* of the bacterial and fungal biofilms are readily identifiable.

B3-6 Conclusions

The application of analytical pyrolysis coupled to gas chromatography and mass spectrometry to samples of bacterial (*Pseudomonas Mendocina*), fungal (*Pleorotus ostreatus*) and mixed-biofilms led to exhaustive results concerning their characterization and differentiation. Bacterial and fungal biofilm can be simply distinguished comparing the methylated fatty acids (FAMEs) profiles which are different in the two cases. Pyrolysis-methylation of fungal biofilms produces some characteristic compounds derived from the cell wall polysaccharides, a pair of deoxyaldonic permethylated acids epimers. Their identification was performed by comparing the retention times and the corresponding mass spectra with those obtained from standard monosaccharides. As mixed-biofilm the simultaneous presence of the bacterium and the fungus has been identified; the corresponding markers are in fact very well distinguishable.

The Py-GC-MS is then a suitable technique for a rapid analysis of microbial biofilms showing a concrete possibility to distinguish the three different samples on the basis of the respective chromatographic profiles (retention times, mass spectra).

REFERENCES

[1] T. P. Wampler; *Applied Pyrolysis Handbook,* II edition (2006), CRC press

[2] T. P. Wampler; *Introduction to pyrolysis – capillary gas chromatography*; J Chrom A, 842 (1999) 207-220

[3] K. Sobeih, M. Baron, J. Gonzalez-Rodriguez; *Recent trends and developments in pyrolysis – gas chromatography*; J Chrom A, 1186 (2008) 51-66

[4] Serban C.Moldoveanu; *Analytical pyrolysis of natural organic polymers*, (1998),Brown & Williamson Tobacco Corp., Macon, GA, USA

[5] G. Chiavari, D. Fabbri, S. Prati; *In situ pyrolysis and silylation for analysis of lipid materials used in paint layers*; Chromatographia, 53 (2001) 311-314

[6] G. Chiavari, D. Fabbri, S. Prati; *Gas chromatographic – mass spectrometric analysis of products arising from pyrolysis of amino acids in the presence of hexamethyldisilazane*; Journal of Chromatography A, 922 (2001) 235-241

[7] J. M. Challinor; *The development and applications of thermally assisted hydrolysis and methylation reactions*; J Anal Appl Pyrol, 61 (2001) 3- 34

[8] D.Melucci, G.Chiavari, S.Montalbani, S. Prati; *Behaviour of phospholipids in analytical reactive pyrolysis*; J Therm Anal Calorim 104 (2011) 415–421

[9] Roberto Todeschini; *Introduzione alla chemiometria*, (1998), Edises

[10] Kim H. Esbensen; *Multivariate Data Analysis In practice,* (2006), V Edition CAMO Software AS.

[11] R.A.Johnson, D.W.Wichern; *Applied Multivariate Statistical Analysis*; (2007), VI Edition Pearson Education, Inc.

[12] M. Matteini, A. Moles; *La Chimica nel Restauro. I materiali dell'arte pittorica*. Nardini Editori, Firenze;1989

[13] J.S. Mills, R.W.White; The *Organic Chemistry of Museum Objects*; Butterworth-Heinemann; 1996

[14] G. Chiavari, N.Gandini, P. Russo, D. Fabbri ;*Characterisation of standard tempera painting layers containing proteinaceous binders by pyrolysis (/methylation)-gas chromatography/mass spectrometry*; Chromatographia, 47 (1998) 420-426

[15] G.Chiavari, P.Bocchini, G.C. Galletti; *Rapid identification of binding media in paintings using simultaneous pyrolysis methylation gas chromatography;* Sci Technol Cultural Heritage, 1 (1992) 153-158

[16] G.Chiavari, D. Fabbri, S. Prati; *Analysis of fatty materials used in painting layers by in situ pyrolysis and silylation*; Chromatographia, 53 (2001); 311-314

[17] F.Cappitelli , T.Learner, O.Chiantore; *An initial assessment of thermally assisted hydrolysis and methylation-gas chromatography/mass spectrometry for the identification of oils from dried paint films*; J Anal Appl Pyrol, 63 (2) (2002) 339-348

[18] S.Prati, G. Chiavari, D. Cam; DSC *application in the conservation field*; J Thermal Anal, 66 (2001) 315-327

[19] M. Odlyha, A. Burmester; Preliminary *investigations of the binding media of paintings by differential thermal analysis*; J Thermal Anal, 33 (1988) 1041-1052

[20] M.Odlyha, S. Felder-Casagrande; *Development of standard paint films based on artists' materials*; J Thermal Anal, 49 (1997) 1585-1591

[21] Bonaduce, M. Cito, M.P. Colombini; *The development of a gas chromatographic–mass spectrometric analytical procedure for the determination of lipids, proteins and resins in the same paint micro-sample avoiding interferences from inorganic media* ; J Chromatogr A, 1216 (2009) 5931-5939

[22] A.Derieux, S.Rochut, M.C.Papillon, C.Pepe; *Identification of proteins in art works by using GC-MS ions trap;* C.R. Acad. Sci. Paris Chimie/Chem, 4(4) (2001) 295-300

[23] A.Andreotti ,I. Bonaduce, M.P. Colombini, G. Gautier, F,Modugno ,E. Ribechini; *Combined GC-MS Analytical Procedure for the Characterization of Glycerolipid, Waxy, Resinous, and Proteinaceous Materials in a Unique Paint Microsample*; Anal Chem, 78 (2006) 4490-4500

[24] M.P.Colombini, F. Modugno; *Characterization of proteinaceous binders in artistic painting by chromatographic technique*; Sep Sci, 27 (2004) 147-160

[25] R.Mateo-Castro , J.V.Gimeno-Adelantado ,F. Bosch-Reig ,A. Domenech-Carbo , M.J.Casas-Catalan, L.Osete-Cortin , J.De la Cruz-Canizares ,M.T. Domenech-Carbo; *Identification by GC-FID and GC-MS of amino acids, fatty and bile acids in binding media used in works of art*; Fresen J Anal Chem 369(9-8) (2001) 642-646

[26] P.Vandenabeele, B.Wehling, L.Moens, H.Edwards, M. De Reu, G.Van Hooydonk; *Analysis with micro-Raman spectroscopy of natural organic binding media and varnishes used in art*; Anal Chim Acta, 407(1-2) (2000) 261-274

[27] O.F.Van den Brink,J.Boon, P.B.O'Connor , M.C. Duursma ,R.M.A Heeren ; *Matrix-assisted laser desorption/ionization Fourier transform mass spectrometric analysis of oxygenated triglycerides and phosphatidylcholines in egg tempera paint dosimeters used for environmental monitoring of museum display conditions*; J Mass Spectr, 36(5) (2001) 479-492

[28] L.Pinturier-Geiss ,J. Laureillard , C.Riaux-Gobin, J. Fillaux , A. Saliot; *Lipids and pigments in deep-sea surface sediments and interfacial particles from the Western Crozet Basin*; Mar Chem, 75 (4) (2001) 249-266

[29] S. Essig, K.A. Kovar; *Quantitative determination of phospholipids in a pharmaceutical drug by scanning and video densitometry* JAOAC Int. 84(4) (2001)1283-1286

[30] A.Carrier, J.Parent, S. Dupuis; *Quantization and characterization of phospholipids in pharmaceutical formulations by liquid chromatography–mass spectrometry*; J Chromatogr A, 876(1-2) (2000) 97-109

[31] Q.Bai, A.Gattinger, L. Zelles; *Characterisation of microbial consortia in paddy rice soil by phospholipid analysis*; Microbial Ecol, 39 (2000) 273-281

[32] D. Giron; *Applications of thermal analysis and coupled techniques in pharmaceutical industry*; J Thermal Anal, 68 (2002) 335-357.

[33] J.M Nzai, A. Proctor*; Soy lecithin phospholipid determination by Fourier Transform Infrared Spectroscopy and the acid digest/arseno-molybdate method: A comparative study*: J Am Oil Chem Soc, 76 (1999) 61-66

[34] U. Stoll; *Techniques of phospholipid analyses applied to sera and egg-yolk*; Fett Wiss Technol, 96 (1994) 188-194

[35] J. Dekock; *The european analytical subgroup of ilps - a joint effort to clarify lecithin and phospholipid analysis*; Fett Wiss Technol 95 (1993) 325-355

[36] R. Przybylskyl, N. Eskin; *Phospholipid analysis of canola oil*; J Am Oil Chem Soc, 55 (1998) 512

[37] K.E.Hundrieser, R.M. Clark, R.G. Jensen; *Total phospholipid analysis in human-milk without acid digestion*; Am J Clin Nutr 41 (1985) 988-993

[38] M.Guichardant, M. Lagarde; *Phospholipid analysis and fatty acid content in platelets by the combination of high-performance liquid chromatography and glass capillary gas—liquid chromatography*; J Chromatogr B, 257 (1983) 400-406

[39] Eibener; *L'examen microchemique des agglutinants*; Mousseion, 20 (1932) 5-22

[40] G.Chiavari, D. Fabbri, S. Prati; *Effect of pigments on the analysis of fatty acids in siccative oils by pyrolysis methylation and silylation*; J Anal Appl Pyrol, 74(2005) 39-44

[41] G.Chiavari, D. Fabbri; *Analytical pyrolysis of carbohydrates in the presence of hexamethyldisilazane;* Anal Chim Acta, 449 (2001) 271-280

[42] W.Butte, J. Eilers, M. Kirsch; *Trialkylsulfonium and trialkylselenonium hydroxides for the pyrolytic alkylation of acidic compounds*; Anal Lett, 15 (1982) 841-850

[43] M. Kleen, G. Gellerstedt; *Influence of inorganic species on the formation of polysaccharide and lignin degradation products in the analytical pyrolysis of pulps*; J. Anal Appl Pyrol, 35 (1995) 15-41

[44] M. Hida, T. Mitsui, Y.Minami, Y. Fujimura; *Classification of hashish by pyrolysis-gas chromatography*; J Anal Appl Pyrol, 32 (1995) 197-204

[45] A.T. Bull, A.C. Ward, M. Goodfellow; *Search and Discovery Strategies for Biotechnology: the Paradigm Shift*; MMBR, 64 (2000) 573–606

[46] J. G. Green, J. M. Potter; *Pyrolysis GC-MS analysis of a bacterial population in a saline activated sludge system—Identification of the origin of the pyrolysis product 2-methylpyrimidine*; J Anal Appl Pyrol 91 (2011) 40–47

[47] S. L. Morgan, A.Fox, J. Gilbart; *Profiling, structural characterization, and trace detection of chemical markers for microorganisms by gas chromatography-mass spectrometry;* J Microbiol Meth, 9 (1989) 57–69

[48] Jeffrey S. Buyer; *Rapid sample processing and fast gas chromatography for identification of bacteria by fatty acid analysis;* J Microbiol Meth 51 (2002) 209– 215

[49] A. J. Madonna, K. J. Voorshees, T. L. Hadfield, E. J. Hilyard; *Investigation of cell culture media infected with viruses by Pyrolysis / Mass Spectrometry: implication for bioaerosol detection*; J Am Soc. Mass Spectrom, 10 (1999) 502-511

[50] F. Basile, M. B. Beverly, K. J. Voorshes, T. L. Hadfield; *Pathogenic bacteria: their detection and differentiation by rapid lipid profiling with pyrolysis mass spectrometry*; Trends Anal Chem, 17 (1998) 95-109

[51] K.J.Voorhees, S.J. DeLuca, A.Noguerola; *Identification of chemical biomarker compounds in bacteria and other biomaterials by pyrolysis-tandem mass spectrometry*; J Anal Appl Pyrol, 24 (1992) 1-21

[52] K.J. Voorhees, F. Basile, M.B. Beverly, C. Abbas-Hawks, A. Hendricker, R.B. Cody,T.L. Hadfield, *The use of biomarker compounds for the identification of bacteria by pyrolysis-mass spectrometry;* J Anal Appl Pyrol, 40–41 (1997) 111–134

[53] Peter G. Simmonds; *Whole Microorganisms Studied by Pyrolysis-Gas Chromatography-Mass Spectrometry: Significance for Extraterrestrial Life Detection Experiments*; Appl. Environ. Microbiol, (1970) 567-572

[54] A. Fox; *Chemical markers for bacteria in extraterrestrial samples;* Anat Rec , 268(2002) 180-185

[55] B. E. Watt, S. L. Morgan, and A. Fox, *2-butenoic acid, a chemical marker for poly-b-hydroxybutyrate identified by pyrolysis gas chromatography-mass spectrometry in analysis of whole cells;* J Anal Appl Pyrol, 20 (1991) 237-250

[56] P. B. Smith, A. P. Snyder; *Characterization of bacteria by quartz tube pyrolysis – gas chromatography / ion trap mass spectrometry*; J Anal Appl Pyrol, 24 (1992) 23-38

[57] D. Frascari, D. Pinelli, M. Nocentini, S. Fedi, Y. Pii, D. Zannoni; *Chloroform degradation by butane – grown cells of Rhodococcus Aetherovorans BCP1*; Appl Microbiol Biot, 73 (2006) 421-428

[58] J. P. Dworzanski, L. Berwald, H. L. C. Meuzeelar; *Pyrolytic methylation – gas chromatography of whole bacterial cells for rapid profiling of cellular fatty acids*; Appl. Environ. Microbiol, 56 (1990) 1717-1724

[59] N. D. Urbina; *A multivariate model for identification of bacteria using pyrolysis – GC-MS*; Uppsala University, Department of mathematics, (2001)

[60] T.K.Wood, S.H. Hong, Q Ma*; Engineering biofilm formation and dispersal*; Trends Biotechnol, 29 (2011) 87-94

[61] T. Coenye, H.J.Nelis; *In vitro and in vivo model systems to study microbial biofilm formation;* J Microbiol Methods 83 (2010) 89–105

[62] K. Sauer, A.H. Rickard, D.G.Davies; *Biofilms and biocomplexity;* Microbe, 2 (2007) 347–353

[63] L. Yang, Y. Liu, H.Wu, Z. Song, N. Høiby, S. Molin, M. Givskov; *Combating biofilm*; FEMS Immunol Med Mic, (2011) 1-12

[64] H. Kobayashi; *Airway biofilm disease;* Int J Antimicrob Ag, 17(2001), 351–356

[65] I.B. Beech, J.A. Sunner, K.Hiraoka; *Microbe-surface interactions in biofouling and biocorrosion processes;* Int Microbiol, 8(2005) 157-168

[66] C. Nicolella, M.C.M. van Loosdrecht, J.J. Heijnen; *Wastewater treatment with particulate biofilm reactors;* Journal of Biotechnology, 80 (2000) 1–33

[67] S. Wuertz, P. L. Bishop, P. A. Wilderer; *Biofilms in wastewater treatment: an interdisciplinary approach;* WA Publishing, 2003

[68] R. Zuo; *Biofilms: strategies for metal corrosion inhibition employing microorganisms*; Appl Microbiol Biotechnol, 76 (2007) 1245-1253

[69] B.Rosche, X. Zhong Li, B. Hauer, A.Schmid, K.Buehler; *Microbial biofilms: a concept for industrial catalysis?*; Trends Biotechnol, 27, (2009), 636-643

[70] G. M. Teitzel , M. R. Parsek; *Heavy Metal Resistance of Biofilm and Planktonic Pseudomonas aeruginosa;* Appl EnvironMicrobiol, *69*(2003) 2313-2320

[71] V.Tremaroli, S.Fedi, R.J. Turner RJ, H. Ceri, D. Zannoni; *Pseudomonas pseudoalcaligenes KF707 upon biofilm formation on a polystyrene surface acquire a strong antibiotic resistance with minor changes in their tolerance to metal cations and metalloid oxyanions*; Arch Microbiol, 190(2008)29-39

[72] V.Tremaroli,C. Vacchi Suzzi , S. Fedi, H. Ceri ,D. Zannoni, R.J. Turner ;*Tolerance of Pseudomonas pseudoalcaligenes KF707 to metals, polychlorobiphenyls and chlorobenzoates: Effects on chemotaxis-, biofilm- and planktonic-grown cells;* FEMS Microbiol Ecol , 74 (2010) 291-301

[73] D.C. Yee, J. A. Maynard, T. K. Wood; *Rhizoremediation of Trichloroethylene by a Recombinant, Root-Colonizing Pseudomonas fluorescensStrain Expressing Toluene ortho-Monooxygenase Constitutively;* Appl Environ Microbiol, 64 (1998) *112-118*

[74] S.Elias, E.Banin; *Multi-Species Biofilms: Living with Friendly Neighbors*; Microbiol Rev. (Accepted 3 Jan 2012)

[75] M.J.Wargo, D. A.Hogan; *Fungal—bacterial interactions: a mixed bag of mingling microbes;* Curr Opin Microbiol, 9 (2006) 359–364

[76] V. Lazarova, J. Manem; *Biofilm characterization and activity analysis water and wastewater treatment;* Wat Res, 29 (1995) 2227-2245

[77] C. Schwarzinger*; Identification of fungi with analytical pyrolysis and thermally assisted hydrolysis and methylation;* J Anal Appl Pyrol, 74(2005) 26–32

[78] C. Schwarzinger; *Identification of methylated saccharinolactones and partially methylated saccharinic acids in the thermally assisted hydrolysis and methylation of carbohydrates;* J Anal Appl Pyrol, *71*(2004) 501-514

[79] D. Fabbri, R. Helleur; *Characterization of the tetramethylammonium hydroxide thermochemolysis products of carbohydrates*, J Anal Appl Pyrol, 49 (1999) 277-293

ACKNOWLEDGMENTS

I wish to express my deepest thanks to my advisor as well as counselor and friend Prof. Dora Melucci, who shared with me many important moments in these three years. I feel I am very grateful for all the opportunities she gave me to learn new things that will be useful for my future.

I would like address a very special thanks to Prof. Chiavari who always proved to have confidence in my abilities and encouraged me to improve and do even the most.

I wish to thank Prof. Locatelli for his availability, expertise and friendliness.

I want to thank prof. Stefano Fedi for his valuable collaboration and availability.

I would also like to thank Dr. Simona Raffo with whom I did a part of this work and who has cheered my days with her freshness and fun.

I want to dedicate a great thanks to Francesca, Paola, Romina and Sonia which shared with me very special moments and helped me a lot when I really needed.

I want to thank my beloved parents who have always been cheering for me and I hope they know how this it was important for me. As for my sister, I want to thank her for all she has done and is doing every day.

A special thanks to Remo and Magda for their constant and sincere support, made of small and great actions.

Finally I would like to express my heartfelt thanks to Simone who encourages and supports me in every moment. But most of all is still able to make me laugh and be amazed as when he decided to accept our beloved kitty Leila.