UNIVERSITA' DEGLI STUDI DI NAPOLI FEDERICO II

Ph.D Thesis in Chemical Sciences
XXXI CICLO

Targeted tandem mass spectrometry strategies to quantify proteins biomarkers of inflammatory diseases

Anna Illiano

Tutor
Prof. Angela Amoresano

Supervisor
Prof. Marco Trifuoggi

Coordinator
Prof. Luigi Paduano

Beato l'uomo che si dedica alla sapienza

e riflette con la sua intelligenza,

che medita nel cuore le sue vie

e con la mente ne penetra i segreti.

[Siracide, 14:20-21]

INDEX

Abstract

Mass spectrometry is a sensitive technique used to detect, identify and quantitate molecules based on their mass-to-charge (m/z) ratio in simple and complex mixtures. Originally developed almost 100 years ago to measure elemental atomic weights and the natural abundance of specific isotopes, MS was first used in the biological sciences to trace heavy isotopes through biological systems.

The following PhD project was involved in the application of advanced methods of mass spectrometry in Multiple Reaction Monitoring ion mode (MRM) into different metabolomics and proteomics research areas.

For the most part, the studies conducted over the past three years have had a single common thread: development and application of MRM mass spectrometry methods for the identification and quantification of targeted metabolites and/or proteins involved in inflammatory processes.

It is clear that, for each project, we started from the study of the different biological matrices to find the most effective extraction strategy for the target analytes and subsequent steps have covered in-depth literature studies to identify the best condition to perform chromatographic separation and the subsequent optimization of instrumental parameters.

The three main areas of application explored in these years concerned the following three points:

1. MRM/MS analysis of metabolites

In collaboration with Prof. Greco, Prof. Auricchio and Prof. Ruoppolo of the Department of Translational Medical Sciences and Department of Molecular Medicine and Medical Biotechnology of the University of Naples Federico II.

The project provided for the development of a method that allow the investigation of the lipidomic profile of genetically predisposed children to celiac disease in order to identify potential molecular biomarkers for disease prediction, first of all, in genetically predisposed patients but also for subjects whose clinical history is unknown.

Serum samples of two cohorts: 23 children who became coeliac and 23 not yet (used as control), do share a similar genetic background, since they come from families with one celiac proband and bear the specific HLA haplotype (DQ2 or DQ8). It does appear that the genetic profile may not explain fully the great differences found between the two cohorts. The

developed MRM method allowed to monitor and quantify 83 different classes of analytes and allowed us to identify some classes of lipids as putative molecular biomarkers by comparing the results obtained from the analysis of the samples of serum collected at 4 months, before introduction of gluten, at 12 months, with the introduction of gluten in the diet of the child and a t> 12 months for children who have been diagnosed with celiac disease.

2. MRM/MS analysis of proteins.

A peculiar feature of a MRM method is the ability to monitor multiple precursor ion-product ion transitions. This greatly increase the selectivity and specificity of the analysis and this represents a huge advantage in the proteomic field because each target analyte, in this case peptide, it can be identified within complex mixtures (such as biological fluids) by monitoring transitions closely related to its own amino acid sequence. Different biological aspects were investigated:

- The project, in collaboration with Prof. Piccoli and Dr. Arciello of the Department of Chemical Sciences of the University of Naples Federico II, involved the development and optimization of an MRM method for the quantification of proteins involved in inflammatory processes: TNF - α, INF-γ, IL-8 and IL-10 in THP-I cell samples. In particular, differentiated cells have been treated with LPS, a well-known endotoxin, to stimulate the onset of an inflammatory process. A time course analysis was performed on differentiated and stimulated cells with LPS for 2h, 4h, 6h, 9h and 24h.
  These analyses allowed to monitor the variation in protein expression during the whole inflammatory process, in both acute and late phase of the inflammation and the obtained data are consistent with published works. Quantitative analysis was conducted using the external standard method. In order to increase the selectivity and the specificity of the method, for each target protein, two or three peptides have been identified thanks to the aid of bioinformatic software which have a unique amino acid sequence and can be used as a stoichiometric representation of the protein in the quantitative analysis.
- Project in collaboration with prof. Francisco Blanco and Dr. Cristina Ruiz-Romero of the INIBIC Biomedical Research Institute of A Coruña, Spain. At this research institute I spent six months for the foreign period of the PhD.

In these months I have worked on both the optimization of a SIS-MRM/MS method for the quantification of proteins involved in rheumatoid arthritis (RA) pathogenesis and the application of this method to a cohort of 80 serum samples of subjects whom RA has been diagnosed.

Quantitative analysis was performed using the internal standard method: stable isotope labelled standard peptides (SIS) were used. The target peptides belong to proteins that were statistically significant ($p < 0.05$) in previous experiments of 8-plex iTRAQ and large-scale proteomics. RA diagnosis is complex and nowadays it is carried out by putting together radiographic data, DAS 28 and serum parameters such as C-reactive protein, rheumatoid factor (RF) and anti-citrullinated antibodies (ACPA) levels. The main problem in diagnostics is incurring false positives, as in the case of RF that shows the same trend also in other diseases like: chronic hepatitis, chronic viral infections, leukaemia, dermatomyositis, mononucleosis, scleroderma, Hashimoto's thyroiditis, systemic lupus erythematosus and Sjögren's syndrome.

First step involved the development of the method, the validation of the selected transitions and the choice of the concentration of labelled peptides to be added to the real samples to obtain the best signal/noise ratio. Reverse calibration curves for each heavy target peptide were realized and analytical parameters Detection limit (LOD), Quantification limit (LOQ) and linearity range were calculated.

The developed method was used for the identification and quantification of the 10 target proteins in a cohort of 80 samples of sera from RA patients subdivided into 4 sub-groups based on rheumatoid factor (RF) and anti-citrullinated antibodies (ACPA) values. The four sub-groups were: RF-/ACPA-, RF-/ACPA+, RF+/ACPA- and RF+/ACPA+.

Data analysis allowed to select some of the proteins monitored for the subsequent steps of method validation by using complementary techniques such as ELISA immunoassay.

3. MRM/MS analysis for discovery.

- An interesting application of mass spectrometry in MRM ion mode was in forensic field for the identification of biomarker proteins of biological fluids. This project was carried out in collaboration with Ten.Col. Peluso of the RIS department of the Carabinieri in Rome. The basic idea was to exploit the potential of mass spectrometry in MRM ion mode to identify

the nature of biological traces found at a crime scene from which the DNA was extracted. In this way it is possible to conduct complementary investigations on both the identity of the suspect and the sequence of events that lead to the crime. Presumptive and confirmatory tests are needed to be absolutely sure of the identity of the biological fluid found at a crime scene. These tests suffer from limitations due, above all, to poor specificity and to the necessity to conduct cascade tests to evaluate the nature of the trace. The developed MRM method allows to overcome these limitations as it allows to discriminate between four biological matrices: blood, saliva, seminal fluid and urine through a single analysis and through a single sample treatment, which involves hydrolysis with trypsin, to carry out the extraction DNA and the subsequent proteomic analysis.

- Project in collaboration with Prof Nicosia from Department of Molecular Medicine and Medical Biotechnologies of the University of Naples Federico II for the identification of proteins in HCT116 cells (human colorectal tumour cells) deriving from alternative splicing processes and which could be closely related to the onset and progression of this type of cancer. Previous studies performed on mRNA have shown that such alternative splicing phenomena lead to proteins that have a mutated sequence to C-ter. This sequence was used for setting up of the MRM / MS method thanks to the high specificity and selectivity that derives from the use of mass spectrometry in this mode, but above all the high sensitivity to which it is possible to arrive (amol/µL). A method was developed to identify the presence of 5 mutated proteins using the mutated C-ter aminoacidic sequence as molecular target and it was possible to verify the presence of these mutated sequences confirming the translation of these proteins.
Finally, for the method validation HT 29 and LS147T cell lines were used. The results seem promising, in all the selected cell lines the mutated proteins were detected. Further developments concern the implementation of the MRM / MS method developed with the use of isotopically labelled peptides to validate and perform quantitative analysis.

Results obtained in the present PhD thesis show the broad applicability of the MRM/MS methodology. This strategy was effective both during the discovery phase and for the quantitative analysis of metabolites and proteins, demonstrating high sensitivity, selectivity and specificity. The chance to simultaneously analyse a panel of numerous analytes allows to optimize analysis times and costs.

Finally, the "multiple" ability of this method allows all methods to be implemented by inserting the characteristic transitions of heavy isotopically labelled standards that have the same chemical and physical properties of the target molecules but different m/z ratio. The use of an internal standard is fundamental to evaluate the matrix effect, the efficiency of an extraction methodology and the identification of an analyte in complex samples.

Subsequent steps for the presented projects are mainly focused on the implementation of the developed MRM methods with isotopically labelled standards, the validation of the obtained results by complementary techniques (ELISA) and the development of kits that can be used in clinical practice for diagnostics or follow-up of patients suffering from various diseases or in forensic investigations.

Chapter 1

1.1 Introduction

1.1.1 Omics and system biology

'Omic' technologies adopt a holistic view of the molecules that make up a cell, tissue or organism. They are aimed primarily at the universal detection of genes (genomics), mRNA (transcriptomics), proteins (proteomics) and metabolites (metabolomics) in different biological compartments such as whole blood, plasma and urine. The integration of these techniques is called systems biology[1].

Systems biology is based on the understanding that the whole is greater than the sum of the parts. It has been responsible for some of the most important developments in the science of human health and environmental sustainability[2]. It is a holistic approach to deciphering the complexity of biological systems integrating many scientific disciplines – biology, computer science, engineering, bioinformatics, physics and others – to predict how these systems change over time and under varying conditions, and to develop solutions to the world's most pressing health and environmental issues.

Components data yield information regarding the specific molecular content of the cell or system. Interactions data specify the connectivity that exists among the molecular species, thereby defining the network 'scaffold' within the cell or system. Finally, functional-states data reveal the overall behaviour, or phenotype, of the cell or system [3].

This ability to design predictive, multiscale models enables our scientists to discover new biomarkers for disease, stratify patients based on unique genetic profiles, and target drugs and other treatments. Systems biology, ultimately, creates the potential for entirely new kinds of exploration, and drives constant innovation in biology-based technology and computation.[4]

1.1.2 Components data

Genomics. Genomics is defined as the study of the whole genome sequence and the information contained therein. It is clearly the most mature of the omics-sciences. Since 1995, nearly 300 genome-sequencing projects, with representative species from each of the three kingdoms of life, have been completed [5]and hundreds more are underway.

The genomic sequence is used to study the function of the numerous genes (functional genomics), to compare the genes in one organism with those of another (comparative genomics), or to generate the 3-D structure of one or more proteins from each protein family, thus offering clues to their function (structural genomics).

Genome annotation provide information about human and other major model organisms on complement of proteins and functional RNAs, transcription-factor-binding sites in genomic sequences and on complete set of translated open reading frames (ORFs) and the exon-intron structures from which they are assembled.

Transcriptomics. The transcriptome is the total mRNA in a cell or organism and the template for protein synthesis in a process called translation. The transcriptome reflects the genes that are actively expressed at any given moment (Horgan RP 2011). Microarrays and serial analysis of gene expression (SAGE) represent the most well-used approaches and have been applied to many model systems, as well as to the study of genes that are predominantly expressed in stem cells, to classifying the molecular subtypes of human cancers, and to monitoring the host-cell transcriptional response to pathogens (Joyce A. R. 2006).

Proteomics. The proteome is defined as the set of all expressed proteins in a cell, tissue or organism. Proteome analysis presents specialized analytical problems in two major areas: i) dynamic expression range and ii) diversity of protein expression (multiple protein forms). In order to clarify proteome–phenotype relationships, several studies are dedicated to the development of new technology to explain the link between protein-expression profiles and distinct cellular processes or conditions[6]. In particular, one strategy that is being developed involves overcoming the problem of detecting only the most highly represented proteins in biological samples by focusing on unique characteristic peptides for each protein or protein isoform .[3]

Metabolomics. The metabolome refers to the complete set of low molecular weight compounds in a biological sample. These compounds are the substrates and byproducts of enzymatic reactions and have a direct effect on the phenotype of the cell. Metabolites are the final products of the genome, transcriptome and proteome integration. The metabolome consists of molecules that have high chemical and physical heterogeneity. Metabolomics is fast becoming a popular tool for studying the cellular state of many systems, including plants, the human red blood cell and microbes, as well as in

metabolic-engineering applications, in pharmacology and toxicology and in human nutritional studies.[3]

Lipidomics. Lipidomics is the systems-level analysis of lipids (fat molecules) and their interactions. It is a science still in its infancy but one that promises to revolutionize biochemistry. Lipids are grouped into eight categories that share common physical and chemical properties, and there are currently some 38,000 documented lipids[7].

Lipids that occur rarely or in small quantities are often the most effectual lipids in biological processes, meaning they are particularly important in disease diagnostics and in understanding pathology. Lipidomics can elucidate the pathology and treatment of many diseases such as cancer, diabetes, obesity, cardiovascular disease, arthritis, asthma, inflammatory bowel disease, Alzheimer's and others due to the associated disruption of lipid metabolic enzymes and pathways. A better understanding of lipidomics could significantly advance diagnostic medicine as well as provide novel treatment options[8].

1.2 Mass spectrometry

Mass spectrometry is a reproducible methodology based on the determination of the molecular mass, which is not dependent on the experimental conditions. Most important features of mass spectrometry are: reproducibility, sensitivity, accuracy.

Mass spectrometry is based on the production of gas phase ions. These ions can interact with an electric field and can be resolved following their electro-dynamic attitude, which is dependent on their mass-to-charge ratio.

Today a wide variety of mass spectrometers is available, all sharing the capability to assign mass-to-charge values to ions, although the principles of operation and the types of experiments that can be done on these instruments differ greatly.

Mass spectrometers have four essential parts:

a) system for sample introduction

b) source that produces gas phase ions from the sample

c) one or more mass analyser to separate ions

d) ion detector.
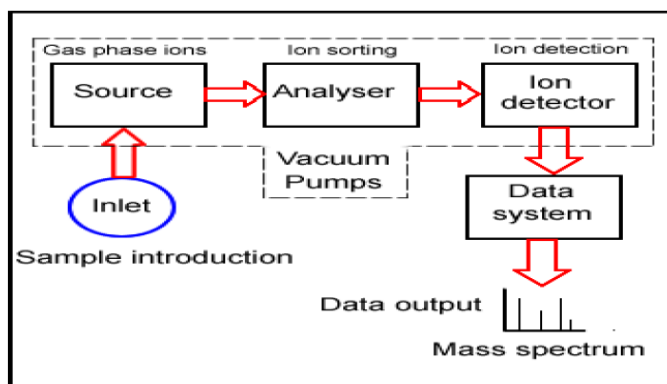
Figure 1: Working process of a mass spectrometer, from sample introduction system to the mass spectrum display.

Sample under investigation is introduced into the ion source of the instrument where the sample molecules are ionized. These ions are extracted into the analyser region of the mass spectrometer where they are separated according to their mass-to-charge ratios (*m/z*). The separated ions are detected, and the generated signals sent to a data system where the m/z ratios are stored together with their relative abundance for presentation in the format of an *m/z* spectrum *(*Fig. 1.1*).*

A class of mass spectrometers can be distinguished on its ionization system and on the type of analyser, that is an essential component to define the accessible mass range, sensitivity and resolution.

All mass spectrometers are operated at very low pressure to prevent collisions of ions with residual gas molecules in the analyser during the flight from the ion source to the detector.

The most widespread ionization methods in biochemical analyses are Electrospray Ionization (ESI) and Matrix Assisted Laser Desorption Ionization (MALDI). Mass analysers as quadrupoles (Q), ion traps (IT), time-of-flight (TOF), or combination of these in "hybrid instruments", are commonly used for their good resolution and sensitivity. TOF analysers are typically used in combination with MALDI sources (MALDI-TOF MS instruments). However, both MALDI sources and TOF analysers can be used in different configuration.

The coupling of Liquid Chromatography (LC) and tandem Mass Spectrometry (MS/MS) is a widely used analytical technique for quantitative

and qualitative analysis. Electrospray Ionization (ESI), Atmospheric Pressure Chemical Ionization (APCI), or Photo Ionization (APPI) allow the ionization of various semi-volatile, thermally labile, and polar to non-polar compounds, such as pharmaceuticals, pesticides, personal care products, steroids, explosives, drugs of abuse etc., in trace levels. Generated ions will be transferred after ionization through a vacuum interface into the mass analyser.

1.2.1 Electrospray (ESI)

Electrospray Ionization (ESI) has had a tremendous impact over the last few years on the use of mass spectrometry in biological research. ESI is well suited to the analysis of polar molecules ranging from less than 100 Da to more than 1,000,000 Da in molecular weight. In contrast to MALDI, in which the sample is a dried, crystalline admixture of protein/peptide sample and matrix, the peptides or proteins to be analysed by ESI are in aqueous solution. Proteins and peptides exist as ions in solution because they contain functional groups whose ionization is controlled by the pH of the solution.

During standard electrospray ionization sample is dissolved in a polar, volatile solvent and pumped through a narrow, stainless steel capillary (Fig. 1.2). A high voltage of 3 or 5 kV is applied to the tip of the capillary situated within the ionization source of the mass spectrometer and the sample emerging from the tip is dispersed into a spray of highly charged droplets, a process that is aided by a co-axially introduced nebulising gas flowing around the outside of the capillary.
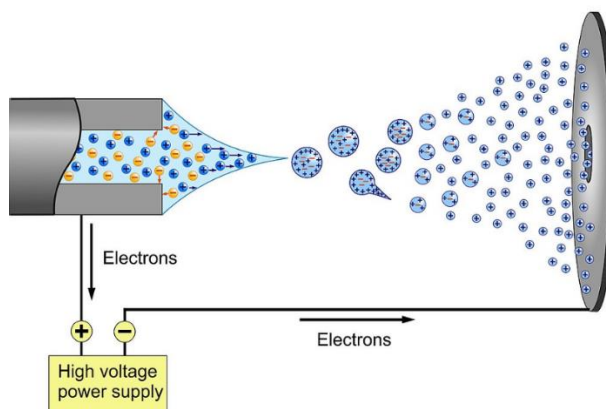
Figure 2: Diagram of electrospray ionization in positive mode

This gas, usually nitrogen, helps to direct the spray emerging from the capillary tip towards the mass spectrometer. The charged droplets diminish in size by solvent evaporation, assisted by a warm flow of nitrogen known as the drying gas which passes across the front of the ionization source. Eventually charged sample ions, free from solvent, are released from the droplets, some of which pass through a sampling cone or orifice into an intermediate vacuum region, and from there through a small orifice into the analyser of the mass spectrometer, which is held under high vacuum. Electrospray is known as a "soft" ionization method (such as MALDI) as the sample is ionized by the addition or removal of a proton, with very little extra energy remaining to cause fragmentation of the sample ions.

The peculiar aspect of this technique is the fact that it gives rise to multiply charged molecular-related ions such as $(M+nH)^{n+}$ in positive ionization mode and $(M-nH)^{n-}$ in negative ionization mode. Because of the form assumed by the signal, this source is not useful to the analysis of complex mixtures. For this reason, RP-HPLC system is generally coupled to ESI-source. ESI source generally works with quadrupole and ion trap analysers.

1.2.2 Quadrupole (Q) mass analyser

Quadrupole mass analyser (Figure 3) has now become one of the most widely used types of mass analyser because of its ease of use, small size and relative low cost. It consists of four parallel rods that have fixed DC and alternating RF potentials applied to them. Ions produced in the source of the instrument are then focused and passed along the middle of the quadrupoles. Their motion will depend on the electric fields so that only ions of a particular *m/z* will be in resonance and thus pass through to the detector following a stable trajectory.
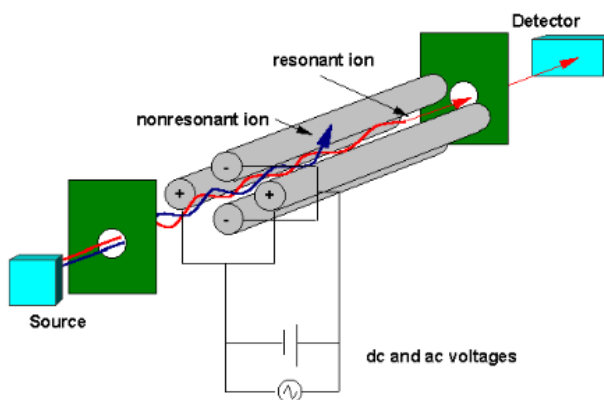
*Figure 3:*Quadrupole mass analyzer

All other ions do not have a stable trajectory through the quadrupole mass analyser and will collide with the quadrupole rods, never reaching the detector. The RF is varied to bring ions of different *m/z* into focus on the detector and thus build up a mass spectrum.

1.2.3 Tandem mass spectrometry

Tandem mass spectrometry (MS/MS) is used to produce structural information about a compound by fragmenting specific sample ions inside the mass spectrometer and identifying the fragment ions. Tandem mass spectrometry also enables specific compounds to be detected in complex mixtures because of their characteristic fragmentation patterns. Tandem mass spectrometry uses two stages of mass analysis, one to preselect an ion and the second to analyse fragments. Generally, in a tandem mass spectrometer the two analysers are separated by a collision cell where an inert gas (e.g. argon, xenon) collides with the selected sample ions and brings about their fragmentation.

1.2.4 Instruments for MS/MS analysis

Instruments for tandem mass spectrometry can be classified as *tandem in space* or *tandem in time* (Figure 4). *Tandem in space* means that ion selection, ion fragmentation and fragments analysis, are events that occur in three different regions of the spectrometer; instruments of this type are the triple quadrupole (ESI-QqQ) and hybrid instruments such as ESI-Q-TOF, ESI-QqLIT or MALDI-TOF-TOF. In *tandem in time* instruments, those three steps of analysis occur in the same region of the spectrometer but in different times; three-dimensional ion trap are instruments classified as *tandem in time*.
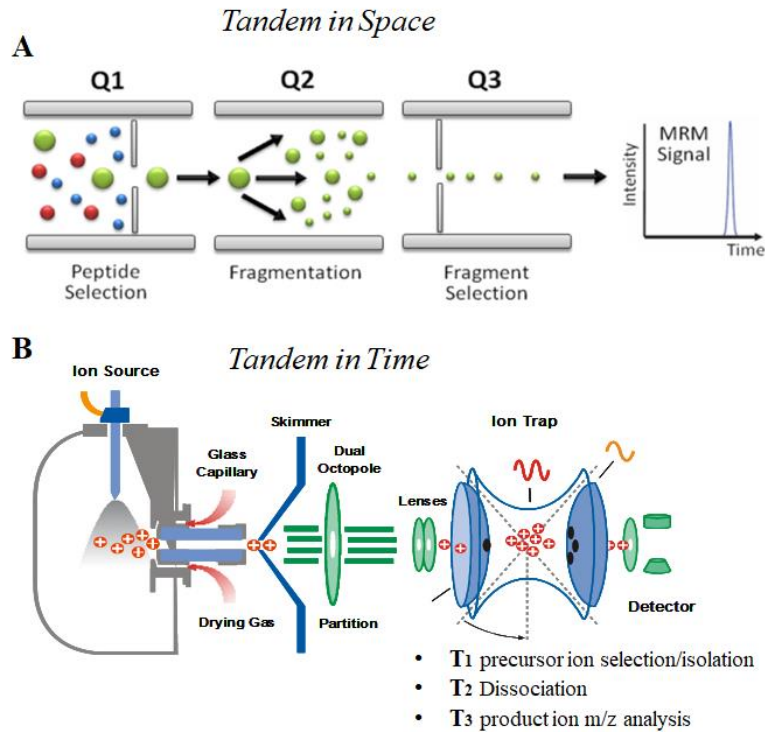
A commonly used instrument is the triple quadrupole, where two resolving quadrupoles mass filters are separated by a fragmentation cell. These are considered to be "beam-type" mass spectrometers because they are most often used with a continuous ion beam: when one of the mass filters is scanning, most of the ions are rejected by the instrument, resulting in poor sample use and low sensitivity, whereas when the mass filters are not scanning, but are transmitting a pre-selected ion, sensitivity and sample use are very high. MSMS mass spectrometer are Quadrupole ion Trap, Linear ion trap and triple quadrupole.

1.2.5 Quadrupole ion trap (IT)

The quadrupole ion trap is based on the same principle as the quadrupole mass filter, except that the quadrupole field is generated within a three-dimensional trap. The trap consists of three electrodes, a ring electrode and two hemispherical end caps electrodes. These electrodes allow to trap ion in a small volume (Figure 5). In IT, ions are dynamically stored in a three-dimensional

quadrupole ion storage device. The RF and DC potentials can be scanned to eject successive m/z ratios from the trap into the detector.
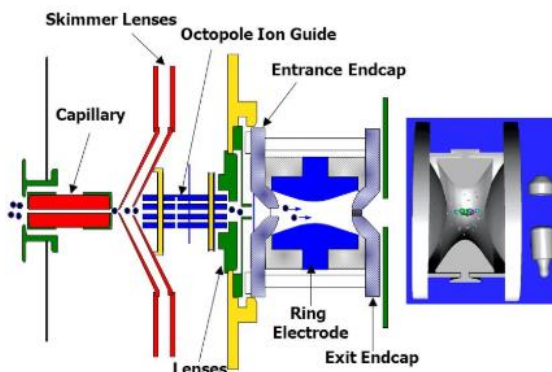
Figure 5: Graphical representation of a Quadrupole Ion Trap geometry

The advantages of the ion-trap mass spectrometer include compact size, and the ability to trap and accumulate ions to increase the signal-to-noise ratio of a measurement. For these reasons, ion traps are used in fragmentation experiments. Conventional ion trap mass spectrometers operate with a three-dimensional (3D) quadrupole field, which confers to the analysis very high efficiency as regard the time to fill the ion trap and to generate a complete mass spectrum, but presents some problems as regard the trapping efficiencies, primarily due to their small volume. Thanks to the introduction of linear ion traps, characterized by a greater ion accumulation capacity and greater trapping efficiency, these problems have been overcome.

1.2.6 Linear ion trap (LIT)

Single quadrupole systems contain only one mass filtering quadrupole while triple quadrupole systems consist of three quadrupoles. Q1 and Q3 are working as mass filters while Q2 is acting as collision cell. Quadrupoles can be used in scanning or filtering mode. During a mass scan, DC and RF voltages are ramped resulting in the acquisition of full scan mass spectra. Such spectra are typically used for qualitative data analysis. However, scanning a quadrupole suffers from low sensitivity and slow scan speed. Thus, quantitative studies are performed with quadrupoles working in filtering mode. The most selective mode to use a single quadrupole MS is called Selected Ion Monitoring (SIM). Hereby, a fixed set of DC and RF voltages is applied to the quadrupole and thus only a single m/z can pass. Ions with different *m/z* are filtered out. A fundamental understanding of the scan modes associated with

14

the TQMS is essential for understanding the MS and MS/MS capabilities of the instrument.[9].

Ion trapping can also be performed in linear 2D ion trap devices (LIT) (Le Blanc *et al*., 2003). The combination of triple quadrupole MS with LIT technology in a form of an instrument of configuration QqLIT, is particularly interesting because this instrument retains the classical triple quadrupole scan functions such as selected reaction monitoring (SRM), product ion (PI), neutral loss (NL) and precursor ion (PI) while also providing access to sensitive ion trap experiments. In addition, for peptide analysis, the enhanced multiply charged (EMC) scan allows an increase in selectivity, while the time- delayed fragmentation (TDF) scan provides additional structural information. The first commercially available instruments (Q-Trap TM) was manufactured by SCIEX and made available through the commercial area of Applied Biosystem.

The linear ion trap combines the advantages of a triple quadrupole with that of the three-dimensional ion trap guaranteeing very selective scanning of precursor ions and high sensitivity. Additional benefits of LITs include increased ion storage capacity and improved trapping efficiency. In addition, there is no inherent low-mass cut-off in the product ion spectra because fragmentation and ion-trapping steps are spatially separated.

Finally, in using Q3 as LIT it is possible to obtain MS^n spectra. Moreover, Q3 is operated at a very low pressure (3 * 10-5 Torr) that insures that the ions entering the Q3 trap are collisionally cool enhancing and enhancing efficiency of MS^n.

Figure 6 summarize the different scan mode in which this instrument can operate.
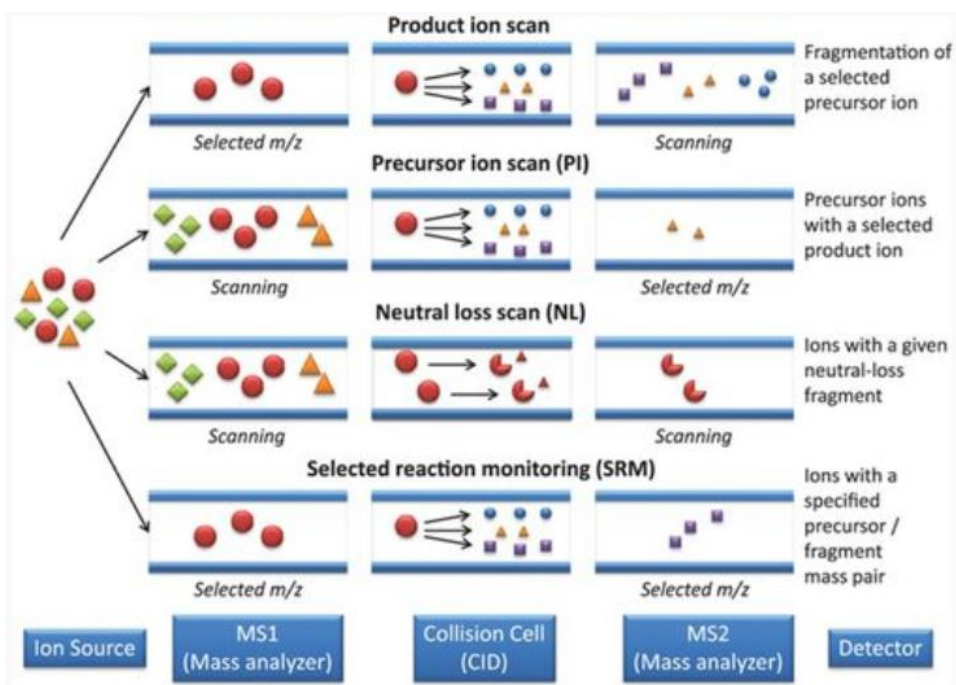
Figure 6: Four different scan mode in which a LIT can operate.

a) In the ***product or daughter ion scanning*** the Q1 is used to select user-specified sample ions arising from a particular analyte; usually the molecular-related (i.e. $(M+H)^+$ or $(M-H)^-$) ions. These chosen ions pass into the q2, are bombarded by the gas molecules which cause fragment ions to be formed, and these fragment ions are analysed (i.e. separated according to their mass to charge ratios) by the Q3. All the fragment ions arise directly from the precursor ions specified in the experiment, and thus produce a fingerprint pattern specific to the compound under investigation. This type of experiment is particularly useful for generating peptide sequence information.

b) In the ***precursor or parent ion scanning*** the Q1 allows the transmission of all sample ions, while the Q3 is set to monitor specific fragment ions, which are generated by bombardment of the sample ions with the collision gas in the q2. This type of experiment is particularly useful for monitoring groups of compounds contained within a mixture which fragment to produce characteristic fragment ions.

c) In the ***constant neutral loss scanning*** both the Q1 and Q3 scan and collect data across the whole *m/z* range, but the two are off-set so that the second analyser allows only those ions which differ by a certain number of

mass units (equivalent to a neutral fragment) from the ions transmitted through the first analyser. This scan mode is used, for example, to identify phosphorylated peptides in a peptide mixture due to the typical fragmentation of these peptides consisting in loss of a phosphoric acid residue that correspond a mass shift of -98 Da.

d) ***Selected Reaction Monitoring*** (**SRM**) is the most common mode of using a triple quadrupole MS/MS for quantitative analysis, allowing enhanced sensitivity and selectivity. The first quadrupole filters a specific precursor ion of interest. Ions generated in the ion source having a different *m/z* can not pass Q1. The collision cell is optimized to produce a characteristic product ion by collision of the precursor ion with a neutral collision gas, such as nitrogen. Generated product ions are transferred into the Q3 where only a specific *m/z* can pass. All other product ions are filtered out in Q3. In SRM mode, two stages of mass filtering are employed on a triple quadrupole mass spectrometer. In the first stage, an ion of interest (the precursor) is preselected in Q1 and induced to fragment by collisional excitation with a neutral gas in a pressurized collision cell (Q2). In the second stage, instead of obtaining full scan MS/MS where all the possible fragment ions derived from the precursor are mass analysed in Q3, only a small number of sequence-specific fragment ions (transition ions) are mass analysed in Q3 (Figure 7).
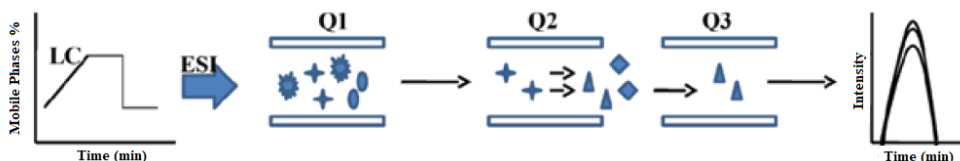
Figure 7:Typical diagram showing the selected reaction monitoring experiment. In this triple quad, Q1 and Q3 act like a mass filter whereas Q2 acts as a collision cell for selected peptide ion.[10]

Thus, SRM mode works like a double mass filter which drastically reduces noise and increases selectivity. Single quadrupole and triple quadrupole systems allow the detection of many SRM transitions, respectively. This enables quantitation of many targeted analytes in a single experiment. Typically, additional SRM transitions must be detected to perform identification of quantified compounds. Hereby, the most intense ion is called the '*quantifier*' and all additional ions are called '*qualifiers*'.

Selected reaction monitoring (SRM) coupled with stable isotope dilution mass spectrometry (SID-MS) using a triple quadrupole mass spectrometer is a

powerful method for quantitative measurement of target molecules. SRM methods, in principal, provide both absolute structural specificity for the analyte and relative or absolute measurement of analyte concentration when stable, isotopically-labelled standards are added to a sample in known quantities. When a synthetic, stable isotope labelled peptide is used as an internal standard, the concentration can be measured by comparing the signals from the exogenous labelled and endogenous unlabelled species. This can be done because they have the same physicochemical properties and differ only by mass. It has been a principal tool for quantification of small molecules in clinical chemistry for number of decades. MS-based quantitative assays have the necessary characteristics required for verification studies, namely: high specificity, sensitivity, multiplexing capability, and precision.

1.3 Mass Spectrometry-based omics

Dramatic technological advances in the biological sciences over the past few years have forged a new era of research including the emerging field of systems biology. Although the understanding of living organisms at the molecular system level is still in its infancy, it is evident that comprehensive investigations of the "omics cascade" with genomics, transcriptomics, proteomics, and metabolomics are important building blocks and will play a central role in this new science[11]. The integrative analysis of an organism's response to a perturbation on the transcriptome, proteome, and metabolome levels will lead to a better understanding of the biochemical and biological mechanisms in complex systems.

Many "omics" techniques have been developed for one goal: identification of specific target compounds involved in diagnosis, prognosis, therapeutic response prediction and population screening of human disease. For example, recent ''omics'' technologies have opened a new road to biomarker discovery and early detection of cancer. A biomarker is a biologically derived molecule in the body, which is indicative of the progress or status of a disease.

Mass spectrometry (MS) is one of the key analytical technology on which the emerging ''-omics'' approaches are based. It may provide detection and quantization of thousands of proteins and biologically active metabolites from a tissue, body fluid or cell culture working in a ''global'' or ''targeted'' manner, down to ultra-trace levels. It can be expected that the high performance of MS technology, coupled to routine data handling, will soon

bring fruit in the request for a better understanding of human diseases, leading to new molecular biomarkers, hence affecting drug targets and therapies.

1.3.1 Mass spectrometry-based metabolomics

Mass spectrometry-based metabolomics aims at the comprehensive and quantitative analysis of wide arrays of metabolites in biological samples having very diverse physicochemical properties and occurring at different abundance levels.

Consequently, comprehensive metabolomics investigations are primarily a challenge for analytical chemistry and specifically mass spectrometry has vast potential as a tool for this type of investigation. Metabolomics require special approaches for sample preparation, separation, and mass spectrometric analysis. It primarily focuses on metabolic fingerprinting, a technique that analyses all detectable analytes in each sample with subsequent classification of samples and identification of differentially expressed metabolites, which define the sample classes.

Metabolome refers to the complete set of small-molecule metabolites (such as metabolic intermediates, hormones and other signalling molecules, and secondary metabolites) to be found within a biological sample, in a high dynamic range of concentrations, such as a single organism. Although the metabolome can be defined readily enough, it is not currently possible to analyse the entire range of metabolites by a single analytical method. Each type of cell and tissue has a unique metabolic 'fingerprint' that can elucidate organ or tissue-specific information, while the study of biofluids can give more generalized though less specialized information. Commonly used biofluids are urine and plasma, as they can be obtained non-invasively or relatively non-invasively, respectively.

A metabolite is usually defined as any molecule less than 1 kDa in size. In plant-based metabolomics, it is common to refer to "primary" and "secondary" metabolites. A primary metabolite is directly involved in the normal growth, development, and reproduction. A secondary metabolite is not directly involved in those processes, but usually has important ecological function. Examples include antibiotics and pigments. By contrast, in human-based metabolomics, it is more common to describe metabolites as being either endogenous (produced by the host organism) or exogenous. Metabolites of foreign substances such as drugs are termed xeno-metabolites. Global

metabolic fingerprinting and quantitative metabolite profiling represent two complementary strategies currently applied for metabolomic investigations[11].

In most cases, the first step in a metabolite profiling experiment is to extract metabolites from the biological matrix. Extracted metabolites are often separated using GC/MS, LC/MS, or CE/MS. Critical to all MS-based approaches is the efficient desorption and ionization of metabolites, where the resulting gas phase ions can be separated by mass analysers such as quadrupole, time-of-flight, and ion trap. Ions are typically detected using a microchannel plate and photomultiplier tube and identified through comparison of exact mass, retention time, and fragmentation information with genuine standards and spectral databases[12].

Nowadays, high throughput LC/SRM/MS allows simultaneous analysis of different class of small molecules, metabolites, or drugs in a short time, therefore enabling a quantitative profiling of hundreds to thousands of samples for target-based metabolite profiling[13].

Among different types of mass spectrometers, a triple quadruple mass spectrometer is optimal for targeted metabolomics, based on its high sensitivity, high specificity, and excellent quantitation ability. Two stages of mass selection reside in a triple quadrupole mass spectrometer: precursor ion (MS1) and a fragment of the precursor ion (product ion, MS2) that produce a molecular weight and structure specific sensitive measurement for a given analyte. The potential power of utilizing a triple quadrupole mass spectrometry-based quantitation technique, namely, Selected Reaction Monitoring (SRM), for metabolomics has not been fully recognized until recently. A triple quadrupole mass spectrometry-based quantitation for small molecules has been heavily used by analytical chemists for analysing drug metabolites, hormones, pesticides, and herbicides with great precision (CV < 10%). To further increase selectivity and sensitivity of the triple quadrupole mass spectrometry-based quantitation assay, a front-end separation technique, such as LC, GC, or CE, is often added as the third dimension of separation. While a variety of separation techniques can be used to couple a triple quadrupole mass spectrometer, the polarity-based liquid chromatography (LC) stands out for its speed, simple sample pre-treatment, and numerous choices of types of commercially available columns based on different separation mechanisms such as reverse phase, normal phase, or hydrophilic interaction, etc[14].

Mass spectrometry-based metabolomics offers quantitative analyses with high selectivity and sensitivity and the potential to identify metabolites.

Combination with a separation technique reduces the complexity of the mass spectra due to metabolite separation in a time dimension, provides isobar separation, and delivers additional information on the physicochemical properties of the metabolites. However, mass spectrometry-based techniques usually require a sample preparation step, which can cause metabolite losses, and based on the sample introduction system and the ionization technique used, specific metabolite classes may be discriminated. Therefore, parallel application of several techniques, for example, GC-MS and LC-MS is desired to study the metabolome comprehensively[15].

1.3.2 Mass spectrometry-based proteomics

Recent successes illustrate the role of mass spectrometry-based proteomics as an indispensable tool for molecular and cellular biology and for the emerging field of systems biology. Proteomics was defined as the study of "the expressed protein complements of a genome at a specific time" by Wilkins in 1994. The terms "proteomics" and proteome" mirror the terms "genomics" and "genome". The drivers of genomic and proteomic analyses are the technological achievements of the past decade that enable the quantitative analysis of the DNA sequence, mRNA and protein expression inside cells and include tools such as DNA microarrays, two-dimensional gel electrophoresis (2D-GE) and mass spectrometry (MS).

Nowadays proteomics can rely on powerful analytical protein-separation technologies (chromatography, electrophoresis), that serve to simplify complex protein mixtures to compare apparent differences in protein levels between two samples. However, the most powerful analytical tools for proteomic analysis is mass spectrometry (MS), whose instrumentation has undergone huge changes over the past years, culminating in the development of highly sensitive, robust instruments that can reliably analyse biomolecules, particularly proteins and peptides. The introduction of ES and MALDI (1980's), in combination with the accessibility of genome sequence information, has revolutionized MS, thus allowing routine MS analysis of protein molecules. Two main strategies for protein ID by MS are currently used in proteomics: top-down and bottom-up proteomics. In top-down proteomics, intact proteins are introduced into a mass spectrometer and then subjected to gas-phase fragmentation[16].

However, the purpose to multiply charged product ions has always been a weak point of this approach, because it may prevent the determination of

product ion masses. With the introduction of the modern mass spectrometers with high mass measurement accuracy, this obstacle has been overcome (e.g., modern MALDI TOF/TOF instruments). Conversely, in bottom-up proteomics, the proteins are firstly separated by gel electrophoresis or chromatography, subsequently digested by specific enzymes (e.g., trypsin to cut lysine and arginine) and then introduced into the mass spectrometer. Bottom-up proteomics approach is represented by peptide mass fingerprinting (PMF) and tandem MS analysis. PMF has largely characterized the early years of the proteomic era; it relies on the acquisition of mass spectra from a tryptic digest of a protein sample and on the measure of tryptic peptide masses searched against a protein database such as UniProt, employing different database search engines and performing, for each protein, an in silico tryptic digest, hence generating a theoretical spectrum. The best overlap between the experimental and theoretical mass spectra then identifies the protein.

The 1D- or 2D-SDS-PAGE-LC-MS/MS based proteomics, often called shotgun proteomics, is now the main bottom-up proteomics technological approach. The workflow provides the protein separation by 1D-SDSPAGE according to MW, followed by in-gel tryptic digestion, peptide analysis by nano-LC-MS/MS and protein ID by database searching, as above described[17]. While the application of shotgun proteomics workflows to tissues, cells, and organelles usually results appropriate, the analysis of body fluids (e.g., serum, blood, plasma, intestinal fluids, urine samples) is particularly difficult because of the complexity and of the high dynamic range of contained analytes.

Recently, LC MS/MS in SRM ion mode has emerged as a promising technique for such precise quantification of targeted proteins. Originally applied to the measurement of small molecules (such as metabolites or drugs), where a capillary chromatography column is connected in-line to the electrospray ionization source of the mass spectrometer [18]. SRM exploits the unique capability of triple quadrupole (QQQ) mass spectrometers to act as mass filters and to selectively monitor a specific analyte molecular ion and one or several fragment ions generated from the analyte by collisional dissociation. The number of such fragment ions that reach the detector is counted over time, resulting in a chromatographic trace with retention time and signal intensity as coordinates. Several such precursor–fragment ion pairs, termed SRM transitions, can be sequentially and repeatedly measured at a periodicity that is fast compared to the analyte's chromatographic elution, yielding chromatographic peaks for each transition that allow for the concurrent quantification of multiple analytes. This multiplexing capability has led to the

term multiple reaction monitoring (MRM), which is frequently used as a synonym of SRM. When applied to proteomics, SRM measures peptides produced by the enzymatic digestion of a proteome as surrogates of the corresponding proteins. Molecular ions within a mass range focused around the mass of the targeted peptide are selected in the first mass analyser (Q1), fragmented at the peptide bonds by collision-activated dissociation (in Q2) and one or several of the fragment ions uniquely derived from the targeted peptide are measured by the second analyser (Q3). A suitably chosen set of SRM transitions therefore constitutes a specific assay to detect and quantify a target peptide and, by inference, a target protein in complex samples[19].

The application of SRM to proteomics has been slow and not without complications. Proteins are large molecules and by themselves at present are not compatible with the technique. The ensuing questions of how many and which of the many peptides generated by tryptic digestion of each target protein constitute optimal inputs for SRM assays are therefore of critical importance. In addition, peptides generally yield more complex fragment-ion patterns than metabolites or drugs, thus complicating the choice of appropriate SRM transitions[18].

1.3.3 SRM in clinical applications

The capability of SRM to quantify specific target compounds across a variety of samples appears particularly well suited for biomarker verification. Candidate biomarkers for their respective disease need to be verified across large sample sets to achieve enough statistical power, targeting easily accessible human specimens, such as serum or plasma[20]. Classically, biomarker verification relied on antibody-based assays, but recently SRM has emerged as an alternative because of its superior multiplexing capabilities of 50–100 analytes, the shorter and cheaper assay development and the capability to discriminate between protein isoforms.

Numerous examples of the application of SRM for the analysis of low molecular weight chemicals with biological relevance exist in the literature and these span a broad range of analytes including endogenous compounds, therapeutic agents and their metabolites, environmental toxicants and compounds of abuse or malicious intent. Recent examples of the use of SRM for analysing endogenous compounds in humans include the measurement of vitamins, steroids and neurotransmitters, multiple analyses of drugs and their metabolites are too numerous to list comprehensively, but include the measurement of therapeutic agents, such as warfarin, triazolam, nevirapine and

antibiotics, the measurement of drugs of abuse, such as heroin, cocaine and cannabinoids and the assay of performance enhancers, such as androgen and other stimulants. In addition, SRM has been utilised for the analysis of low molecular weight chemicals present in plants, fish and contaminated water courses, demonstrating the versatility and breadth of application of the technique for the routine quantification of low molecular weight products.[21]

Over the last ten years, several studies that applied SRM to protein analysis increased exponentially and the subject of such studies is progressively shifting from technological advances to biological or biomedical applications, a development that attests to the increasing maturity of the technology. A challenge in using SRM for candidate-biomarker verification is the required sensitivity for the quantification of low-abundance proteins, given a dynamic concentration range of plasma proteins over 12 orders of magnitude[18]. SRM has also been applied to study signalling pathways, for example, Wnt/β-catenin signalling, a system of high biological and biomedical importance given its deregulation in different types of cancer. Similarly, SRM has been used to quantitatively monitor linked to mammalian stem cell renewal and pluripotency in nuclear extracts from mouse embryonic stem cell[22].

Chapter 2

2.1 SRM method

2.1.1 Analytical approach

Targeted metabolomics and proteomics focus on the quantitative measurements of a specific subset of known analytes representative of biologically relevant processes by LC-MS/MS in Selected Reaction Monitoring scan mode. This approach was shown to be a valuable tool to identify altered pathways in pathological conditions and/or to define therapeutic modes of action.

Aim of this project was the development of robust and versatile analytical procedures for the direct determination and quantification of proteins or metabolites in different biological samples.

The methodological approach developed here are based on different steps: 1) identification of a set of proteins and metabolites of interest capable of satisfying a specific biological or clinic request 2) selection of transitions maximizing sensitivity and selectivity 3) optimization of SRM transitions by tuning acquisition parameters of the mass spectrometer 4) validation of the transitions in biological matrix to account to unspecific contributions of fragment ions background 5) quantification by SRM.

2.1.2 Identification of targeted analytes

The first step of a targeted experiment is the selection of a set of compounds of interest. Depending on the sensitivity and accuracy required, hundreds and eventually up to 1000 analytes can be targeted in a single LC-MS analysis after the transitions have been optimized. The selection of the metabolites and proteins set might be based on previous experiments or the scientific literature or to make possible in vitro and in vivo studies for specific disease.

In particular, for proteins analyses specific peptides (proteotypic peptides)[23]. that are unique for target protein and easily detectable by mass spectrometry should be chosen. The uniqueness of a peptide sequence in a proteome can be determined, in principle, from the genomic information, but the true complexity of proteomes is generally difficult to predict. The choice of peptides with favourable mass spectrometry properties is thus crucial, as it determines the sensitivity of the assay. Nowadays, information from prior experiments conducted on natural proteomes are available in different database

as, for example, PeptideAtlas[24], the Global Proteome Machine Database (GPMDB)[25], that can be used for identification of peptides that can be reproducibly detected and are thus likely associated with the most intense signals.

Numerous software tools for evaluating the best SRM peptides have emerged such as Skyline[26]. Among the parameters critical for predicting peptides were hydrophobicity, charge, energetic and structural properties.

In general, short hydrophilic and long hydrophobic peptides should be avoided, whereas fully tryptic peptides with an average length of ~10 amino acids, devoid of residues prone to artefactual or post-translational modifications should be targeted, like peptide susceptible to undergo glycosylation, phosphorylation, etc. might to lead to bias as they can be present in various forms[27].

2.1.3 Selection of optimal transitions

The combination of m/z settings for the first and third quadrupole is referred as "transition", it is important to select transition ions that maximize sensitivity and specificity of the SRM experiment. Several in silico methods are available, and continue to be refined, which facilitate the choice of the transitions. Transitions can be selected from (1) libraries of previously collected MS/MS spectra in online repositories, (2) computational tools that predict fragmentation, (3) analysis of synthetic peptide standards. Several software tools, for facilitating transition selection have been described. The current practice is to select two transitions in metabolomic analysis and at least two to five most intense transitions per peptides for proteomic investigations.

Skyline utilizes the ProteoWizard libraries [28]to allow analysis of data from all MS instrument platforms, thus providing a vendor-neutral resource for sharing and creation of both methods and results across instrument platforms. Skyline facilitates the generation and refinement of proteomic peptide lists from protein sequences or database entries, both by utilizing online MS/MS spectral repositories and by supporting the generation of custom-built libraries based upon sets of locally acquired tandem spectral data.

Skyline creates transition lists and vendor-specific instrument methods that can be imported directly into instrument control software for MS instruments from several vendors. Skyline also provides a platform for standardized analysis of SRM result files for peak integration and visualization and data quality assessment across multiple analyses. Finally, Skyline allows

the export of processed data in custom report formats compatible with subsequent statistical analyses, publication, and database deposition[29]. However, there is a physical limit to the number of transitions that can be measured in the same analysis. The SRM cycle time is the product of the number of transitions recorded in the cycle and the time spent on acquiring each transition signal (dwell time). A too long a cycle time, due to too many monitored transitions, hence an insufficient number of data points to reconstruct the chromatographic elution profile of the targeted peptide, compromising accurate quantification. Alternatively, this results in a low dwell time and a reduced signal-to-noise ratio, compromising the detection of low-abundance components. This generates the well-known trade-off between the number of transitions and the limit of detection of an SRM experiment[30].

2.1.4 Optimization of SRM transitions

In order to increase the limit of detection and quantification of the SRM assays mass spectrometry parameters that are molecule-dependent (such as declustering potential or cone voltage) or transition-dependent (for example, collision energy) needs to be optimized.

In primis, ionization devices and experimental conditions, such as flow rate, solvents and background, can influence charge state distributions and the process of dissolvation and dissociation of ion clusters supported by a voltage potential (referred to as 'declustering potential' (DP), 'fragmentor voltage' or 'ion transfer capillary offset voltage' depending on the manufacturer) must be optimized.

During fragmentation singly charged y ions are the predominant type of fragments generated by CID in a linear collision cell. Only small b ions are usually observed. Fragments with *m/z* values close to the precursor should be avoided as such transitions are usually noisy. Fragments with *m/z* values above the precursor generally display the highest selectivity, as the singly charged chemical background cannot result in fragments with higher *m/z* than the precursor. In contrast, tryptic peptide ions are predominantly doubly or triply charged with one charge at each terminus. Upon fragmentation, one charge is lost and therefore a part of the fragments has an *m/z* value bigger than the precursor value. A parameter that is of considerable importance is the collision energy, that is tuned to optimize the intensity of the fragment ions of interest: with increasing collision energy, a larger part of the precursor ions is fragmented, and fragment ion intensity increases until this increase is overcompensated by the losses due to secondary fragmentation events. The

optimal collision energy is approximately linearly correlated with the precursor mass for a given charge state. However, particular peptides or fragments deviate considerably from the predicted value.

The easiest and most systematic way of optimizing ionization and fragmentation conditions is to test possible transitions in direct infusion mode and ramp the parameters. This process is partly automated by add-ons for the acquisition software.

2.1.5 Validation of the transitions and matrix effect

Transitions extracted for an SRM assay should be validated by addressing the likelihood that the chosen transitions and their intensity distribution are associated with the target peptide or metabolite. Each transition selected for a specific target should be evaluated in the context of the actual biological matrix to account for unspecific contributions of the fragment ions deriving from co-eluting species with similar properties.

Plasma is probably the most complex biological matrix, so it is important to evaluate the matrix effect that can affect the goodness of the analysis and the quality of the result obtained. Whilst this is readily achievable for low molecular weight markers, it provides a major challenge when looking for proteins or peptides. It has been estimated that human plasma contains proteins that span 10 orders of magnitude of concentration and includes resident plasma proteins (high and medium abundance) and signalling proteins such as hormones and cytokines (low abundance).

Specifically, the matrix affects not only ion suppression, a major pitfall for the analyst, but also detection capability, repeatability and accuracy.

The mechanism of the matrix effect in LC–MS, well described by Trufelli et al.2011[31], is not fully understood. One of the most important aspects is competition between an analyte and a co-eluting matrix component during ionization. As a result, there is a decrease in analyte ionization (ion suppression) or an increase in this ionization (ion enhancement). The matrix effect depends strongly on the type of ionization source. In ESI droplets are produced and a greater number of additives (from eluents or sample matrices) may lower evaporation efficiency and the ability of analytes to reach the gas phase.

Reactions in the electrospray source depend on the properties of solvents and additives (like volatility, viscosity, pH or electrolyte concentration), the physicochemical properties of analytes (pKa, hydrophobicity, proton affinity or ion solvation energy) and the operating parameters of LC–MS instruments

(flow rate, temperature and voltage). Hence, all these variables make ionization a highly complex and changeable process. Furthermore, co-eluting components may produce similar ions in MS or MS/MS experiments, which leads to the erroneous interpretation of results, especially when these components are present at high concentrations in the extract and eluted in the same retention window as the target compounds.

Several strategies have been put forward to eliminate the problems resulting from matrix [32,33]. The most effective of them is exhaustive sample clean-up used immunoaffinity depletion of highly abundant plasma proteins in combination with peptide fractionation by strong- or mixed-mode cation exchange chromatography, which can help to remove interferents, and improves the LOQs to 1–10 ng ml− but it does run the risk of analyte loss. The second strategy is to improve chromatographic separation. This allows the analytes to be eluted in an appropriate period of time, in order to avoid co-elution with matrix components. A third approach is to dilute the final extract several times so that fewer matrix components will be injected into the analytical system

The matrix effect can be compensated for with appropriate calibration methods, firstly, by using calibration standards in pure solvents, and secondly, by using standards in the matrix [33].

Finally, the standard-addition method is a much cheaper and useful approach when no blank matrix is available for the calibration. A known amount of analyte (spike) is introduced into aliquots of sample extracts containing the target compound, so that any co-extracted impurity is accounted for in the calibration; however, the volume of standard added must be small enough to prevent sample extract dilution. By using this methodology, the unknown concentration initially present in the sample can be calculated by extrapolation.

At this point, for each analyte the recovery value (%) is estimated by measuring the concentration of analyte added to real samples (spiked samples) according to the following formula:

$$Recovery\ (\%) = \frac{c_1 - c_2}{c_3} \times 100$$

Where:

C1: analyte concentration measured after the addition

C2: analyte concentration measured before the addition

29

C3: added concentration

2.1.6 Quantification by SRM

SRM-based quantification can be coupled to different strategies for relative or absolute protein quantification.

The aim of relative quantification is to express the amount of a targeted compound in one state or sample relative to that of a second state or sample. Approaches to relative quantification that can be coupled to SRM include label-free or stable isotope–labelling methods. Isotope-labelling methods introduce stable isotope tags to proteins or peptides and rely on the concomitant mass spectrometry measurement of heavy and light or differently labelled peptides as an internal standardization, thus correcting for experimental and instrumental variability, downstream of the label-mixing step. Label-free approaches[34] rely on the direct evaluation of mass spectrometry signal intensities of naturally occurring peptides contained in a sample. Although in principle each of these quantification methods can be coupled to SRM analysis to be noted that methods based on chemically added tags normally change peptide-fragmentation patterns as well as the precursor-ion mass. In such cases, new SRM assays should be developed that are specific for the labelled peptides. In contrast, labels that conserve the chemical structure of peptides, such as metabolic labelling, also conserve their fragmentation patterns and thus the relative transition signal intensities. One can thus easily derive the assays for the heavy form of a peptide from the fragment-ion spectrum of the light peptide and the knowledge of the site and type of incorporation of the heavy label. A drawback of label-based methods is the introduction of additional sample-processing steps. An intermediate solution between the high precision of label-based approaches and the simplicity of label-free methods is the use of a single labelled peptide as a reference standard for all other measured endogenous peptides [35].

The gold standard for absolute quantification is stable-isotope dilution. The most commonly used approach relies on isotopically labelled reference peptides that are chemically identical to the light native peptides (AQUA peptides) [36]. The analytical precision of this method is high and can result in up to 5% errors in the estimation of the amount of peptide originally loaded onto the liquid chromatography system. However, issues such as incomplete digestion of the target protein, partial artefactual modification of the target peptide or partial loss of the synthetic peptide before addition can affect the

accuracy of this approach. For such reasons, the absolute quantification method is sometimes referred to as 'precise relative quantification[18].

In our SRM experiment, quantification was carried out by the external standard method with the realization of calibration curves, considering that areas subtended to peaks, for each transition, are proportional to concentration. The area under this curve is the extracted ion current (XIC) of a specific transition. XICs of peptides of interest are extracted from the original data file. Filters are used to correct baselines or to remove artefactual spikes. XICs are then smoothed with an average filter before performing a closing and an opening mathematical morphology operation with a small flat structuring element. The closing operation eliminates thin valleys and conserves the intensity of local maxima, while the opening operation eliminates thin peaks (i.e. remaining spikes) and conserves the intensity of local minima. Hence, detection of peak positions is performed on the closed profile, and the opened profile is used to eliminate remaining spikes. The peak boundaries are searched on the closed profile, and the peak area (i.e. the quantification value) is computed.

Another important and critical procedure in protein quantification is data normalization. Linearity determines the highest measurable concentration within the specified conditions. The lower limit of detection (LLOD, often referred to as LOD) and the LOQ are defined, respectively, as the concentration level at which the analyte can be reliably detected in the sample under consideration and as the level at which the analyte can be detected and measured with enough precision. Several methods are used to determine LOD and LOQ [23,37,38]. The simplest one consists in calculating LOD and LOQ for a given analyte as the amount of this analyte providing a signal corresponding to the mean value of repeated blank sample measurements +3 and +10 standard deviations, respectively.

These strategies are simple and cost-effective and have demonstrated high reproducibility and linearity.

2.2 References

(1)   Horgan, R. P.; Kenny, L. C. 'Omic'technologies: genomics, transcriptomics, proteomics and metabolomics. The Obstetrician & Gynaecologist 2011, 13, 189-195.

(2)   Kitano, H. Systems biology: a brief overview. Science 2002, 295, 1662-1664.

(3)   Joyce, A. R.; Palsson, B. Ø. The model organism as a system: integrating'omics' data sets. Nature reviews Molecular cell biology 2006, 7, 198.

(4)   Hartwell, L. H.; Hopfield, J. J.; Leibler, S.; Murray, A. W. From molecular to modular cell biology. Nature 1999, 402, C47.

(5)   Liolios, K.; Tavernarakis, N.; Hugenholtz, P.; Kyrpides, N. C. The Genomes On Line Database (GOLD) v. 2: a monitor of genome projects worldwide. Nucleic acids research 2006, 34, D332-D334.

(6)   Kuster, B.; Schirle, M.; Mallick, P.; Aebersold, R. Scoring proteomes with proteotypic peptide probes. Nature reviews Molecular cell biology 2005, 6, 577.

(7)   German, J. B.; Gillies, L. A.; Smilowitz, J. T.; Zivkovic, A. M.; Watkins, S. M. Lipidomics and lipid profiling in metabolomics. Current opinion in lipidology 2007, 18, 66-71.

(8)   Smith, R.; Mathis, A. D.; Ventura, D.; Prince, J. T. Proteomics, lipidomics, metabolomics: a mass spectrometry tutorial from a computer scientist's point of view. BMC bioinformatics 2014, 15, S9.

(9)   Lopes, A. S.; Santa Cruz, E. C.; Sussulini, A.; Klassen, A.: Metabolomic strategies involving mass spectrometry combined with liquid and gas chromatography. In Metabolomics: From Fundamentals to Clinical Applications; Springer, 2017; pp 77-98.

(10)   Kumar, S.; Mittal, P. Selected Reaction Monitoring: A Valid Tool for Targeted Quantitation of Protein Biomarker Discovery. J Anal Bioanal Tech 2017, 8, 2.

(11)   Dettmer, K.; Hammock, B. D. Metabolomics--a new exciting field within the" omics" sciences. Environmental Health Perspectives 2004, 112, A396.

(12)   Lee, D. Y.; Bowen, B. P.; Northen, T. R. Mass spectrometry—based metabolomics, analysis of metabolite-protein interactions, and imaging. Biotechniques 2010, 49, 557-565.

(13)   Wei, R.; Li, G.; Seymour, A. B. High-throughput and multiplexed LC/MS/MRM method for targeted metabolomics. Analytical chemistry 2010, 82, 5527-5533.

(14)   Villas-Bôas, S. G.; Mas, S.; Åkesson, M.; Smedsgaard, J.; Nielsen, J. Mass spectrometry in metabolome analysis. Mass spectrometry reviews 2005, 24, 613-646.

(15)   Dettmer, K.; Aronov, P. A.; Hammock, B. D. Mass spectrometry-based metabolomics. Mass spectrometry reviews 2007, 26, 51-78.

(16)   Girolamo, F. D.; Lante, I.; Muraca, M.; Putignani, L. The role of mass spectrometry in the "omics" era. Current organic chemistry 2013, 17, 2891-2905.

(17)   Evans, V. C.; Barker, G.; Heesom, K. J.; Fan, J.; Bessant, C.; Matthews, D. A. De novo derivation of proteomes from transcriptomes for transcript and protein identification. Nature methods 2012, 9, 1207.

(18)   Picotti, P.; Aebersold, R. Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. Nature methods 2012, 9, 555.

(19)   Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. Selected reaction monitoring for quantitative proteomics: a tutorial. Molecular systems biology 2008, 4, 222.

(20)   Rifai, N.; Gillette, M. A.; Carr, S. A. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nature biotechnology 2006, 24, 971.

(21)   Kitteringham, N. R.; Jenkins, R. E.; Lane, C. S.; Elliott, V. L.; Park, B. K. Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. Journal of Chromatography B 2009, 877, 1229-1239.

(22)   Hewel, J. A.; Liu, J.; Onishi, K.; Fong, V.; Chandran, S.; Olsen, J. B.; Pogoutse, O.; Schutkowski, M.; Wenschuh, H.; Winkler, D. F. Synthetic peptide arrays for pathway-level protein monitoring by LC-MS/MS. Molecular & Cellular Proteomics 2010, mcp. M900456-MCP900200.

(23)   Brönstrup, M. Absolute quantification strategies in proteomics based on mass spectrometry. Expert review of proteomics 2004, 1, 503-512.

(24)   Deutsch, E. W.; Lam, H.; Aebersold, R. PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. EMBO reports 2008, 9, 429-434.

(25)   Craig, R.; Cortens, J. P.; Beavis, R. C. Open source system for analyzing, validating, and storing protein identification data. Journal of proteome research 2004, 3, 1234-1242.

(26)   MacLean, B.; Tomazela, D. M.; Abbatiello, S. E.; Zhang, S.; Whiteaker, J. R.; Paulovich, A. G.; Carr, S. A.; MacCoss, M. J. Effect of collision energy optimization on the measurement of peptides by selected reaction monitoring (SRM) mass spectrometry. Analytical chemistry 2010, 82, 10116-10124.

(27)   MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 2010, 26, 966-968.

(28)   Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics 2008, 24, 2534-2536.

(29)   Liebler, D. C.; Zimmerman, L. J. Targeted quantitation of proteins by mass spectrometry. Biochemistry 2013, 52, 3797-3806.

(30)   Gallien, S.; Duriez, E.; Domon, B. Selected reaction monitoring applied to proteomics. Journal of Mass Spectrometry 2011, 46, 298-312.

(31)   Trufelli, H.; Palma, P.; Famiglini, G.; Cappiello, A. An overview of matrix effects in liquid chromatography–mass spectrometry. Mass spectrometry reviews 2011, 30, 491-509.

(32)   Kostiainen, R.; Kauppila, T. J. Effect of eluent on the ionization process in liquid chromatography–mass spectrometry. Journal of Chromatography A 2009, 1216, 685-699.

(33)   Frenich, A. G.; Vidal, J. L. M.; Moreno, J. L. F.; Romero-González, R. Compensation for matrix effects in gas chromatography–tandem mass spectrometry using a single point standard addition. Journal of Chromatography A 2009, 1216, 4798-4808.

(34)   Rinner, O.; Mueller, L. N.; Hubálek, M.; Müller, M.; Gstaiger, M.; Aebersold, R. An integrated mass spectrometric and computational framework for the analysis of protein interaction networks. Nature biotechnology 2007, 25, 345.

(35)   Zhang, H.; Liu, Q.; Zimmerman, L. J.; Ham, A.-J. L.; Slebos, R. J.; Rahman, J.; Kikuchi, T.; Massion, P. P.; Carbone, D. P.; Billheimer, D. Methods for peptide and protein quantitation by liquid chromatography-multiple reaction monitoring mass spectrometry. Molecular & Cellular Proteomics 2011, mcp. M110. 006593.

(36)   Gerber, S. A.; Rush, J.; Stemman, O.; Kirschner, M. W.; Gygi, S. P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proceedings of the National Academy of Sciences 2003, 100, 6940-6945.

(37)   Armbruster, D. A.; Tillman, M. D.; Hubbs, L. M. Limit of detection (LQD)/limit of quantitation (LOQ): comparison of the empirical and the statistical methods exemplified with GC-MS assays of abused drugs. Clinical chemistry 1994, 40, 1233-1238.

(38)   Zorn, M. E.; Gibbons, R. D.; Sonzogni, W. C. Weighted least-squares approach to calculating limits of detection and quantification by modeling variability as a function of concentration. Analytical Chemistry 1997, 69, 3069-3075.

Chapter 3

3.1 Introduction

Lipidomics is the study of the structure and function of the whole set of lipids (the *lipidome*) produced in a given organism as well as their interactions with other lipids, proteins and metabolites.

The importance in the study of *lipidome* lies in the many vital activities carried out within biological systems. Lipids act as:

- an energy reservoir,
- the major structural components of biological membranes (plasma membranes, membranes of intracellular organelles),
- lipid transporters[1],
- chemical messengers in signal transduction,
- interact with proteins to regulate their functions[2],
- important bioactive molecules in the body's immune system against viral and bacterial infections[3].

In 2005, the LIPID MAPS consortium, the International Lipid Classification, and National Nomenclature Committee grouped lipids into eight categories based on lipids' chemical and biochemical properties: fatty acyls (FA), glycerolipids, glycerophospholipids, sphingolipids, sterol lipids, prenol lipids, saccharolipids, and polyketides[4]. Lipids categories are further divided into subclasses that have diverse structural and chemical properties[5].

Lipidomics has experienced rapid progress, mainly because of continuous technical advances in instrumentation that are now enabling quantitative lipid analyses with an unprecedented level of sensitivity and precision. The still-growing category of lipids includes a broad diversity of chemical structures with a wide range of physicochemical properties[6].

Diet, genetic and disease context can alter levels of concentration and lipid composition in plasma. Plasma lipids are carried by lipoproteins in the body's circulation and changes in the number of lipoproteins, lipid composition and concentration, or a defect in lipid exchange among lipoproteins, reflect the functionality of lipoproteins and manifest signs of metabolic diseases.

Numerous studies demonstrated the lipidome of lipoproteins is altered during disease development, like other metabolomic parameters[7]. As an example, 27-hydroxycholesterol a cholesterol metabolite, promotes atherosclerosis via estrogen receptor alpha-mediated pro-inflammatory

events. Plasma 27-hydroxycholesterol is the most commonly found oxysterol in human plasma and it increases with hypercholesterolemia and its level is also correlated with coronary artery disease.[8]

3.1.1 Phospholipids and Phosphatidylcholine

Phospholipids are the most abundant class of membrane lipids. They are amphipathic molecules with the apolar part (consisting of two saturated or unsaturated fatty acids with 16 or 18 C atoms), linked to glycerol and the polar part (consisting of the phosphate group, esterified with the choline, the ethanolamine, the serine or inositol) also linked to glycerol (Figure 1 and Figure 2). When many phospholipids line up, they form a double layer that is characteristic of all cell membranes.
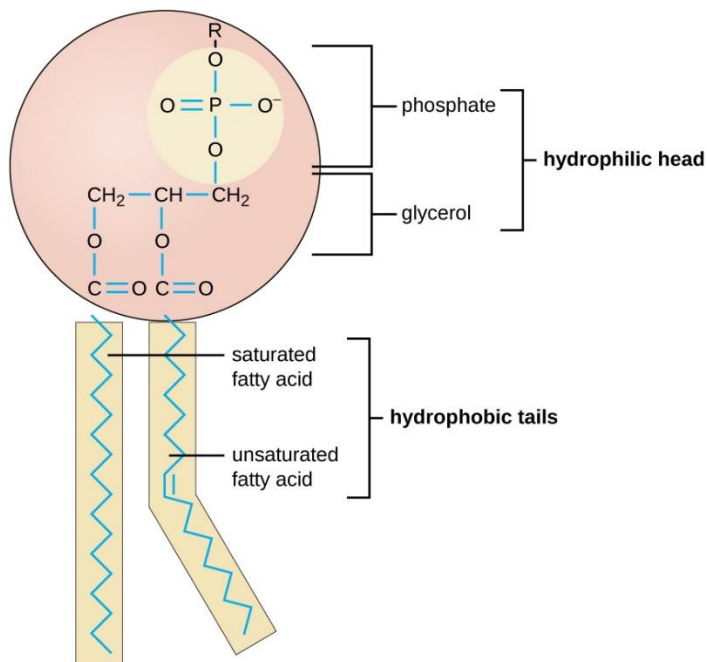
Figure 8: Schematization of a phospholipid structure.

In addition to the structural function, as the main constituents of the plasma membrane, phospholipids are organized to form a semipermeable barrier which can be freely passed through fat-soluble molecules but is impenetrable from those that are water-soluble. Moreover, they develop an important metabolic function being a source of lipid mediators that can be mobilized in response to specific stimuli[9].

Additionally, they are found in lung and joints, where they help lubricate cells. In pharmaceuticals, phospholipids are used as part of drug delivery systems, that promote the drug transport throughout the body to the area that it is meant to affect. Valium is an example of a medication that uses a phospholipid-based drug delivery system[10]. In the food industry, phospholipids can act as emulsifiers, are found in high concentrations in many other animal and plant sources, such as soybeans, sunflowers, cotton seeds, corn, and even cow brains [11,12].

3.1.2 Phosphatidylcholine biosyntesis

Phosphatidylcholines are the most abundant phospholipids in animal tissues and typically contain palmitic, stearic, oleic, linoleic, or arachidonic acid. They usually have saturated fatty acids in the sn-1 position and unsaturated fatty acids at sn-2.

The de novo pathways[13] for phospholipid synthesis use cytidine triphosphate (CTP) for activation of intermediate species. The principle pathway (Figure 2) of phosphatidylcholine biosynthesis uses cytidine diphosphate (CDP-choline)[14]. Many reactions of phospholipid synthesis occur in the endoplasmic reticulum. Choline is phosphorylated by ATP to phosphocholine, which reacts with CTP to form CDP-choline, from which phosphocholine is transferred to sn-1,2-diacylglycerol. The rate-limiting step in this pathway is catalyzed by CTP phosphocholine cytidylyltransferase, which is activated by fatty acids.

PC can also be generated endogenously in a second pathway via three sequential methylations of PE by phosphatidylethanolamine N-methyltransferase (PEMT)[15]. The PEMT pathway is quantitatively significant only in the liver where it contributes approximately 30% of total hepatic PC synthesis[15,16]. Subcellular fractionation revealed that both PEMT and the final enzyme of the CDP-choline pathway[16] reside on the endoplasmic reticulum (ER), underscoring the importance of this organelle in PC biosynthesis.

Phosphatidylcholine is degraded by phospholipases that cleave preferentially at specific bonds. The choline released is phosphorylated by choline kinase and reutilized in phosphatidylcholine synthesis. However, in liver mitochondria, choline is also oxidized to betaine (N-trimethylglycine).
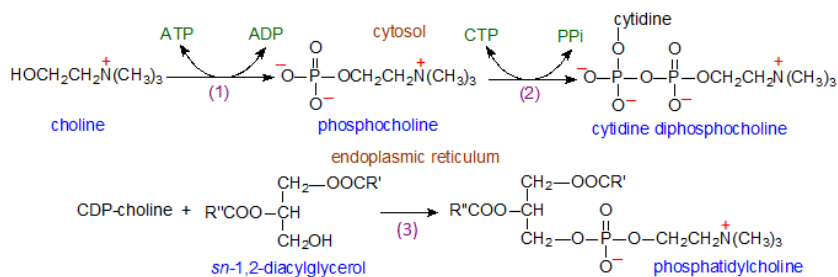
Figure 2: Main pathway for phosphatidylcholine in animals and plants.

3.1.3 Celiac Disease

Coeliac disease (CD) may be considered the most common chronic inflammatory condition since the estimated prevalence in Western Countries is near to one per cent[17]. CD is an autoimmune enteropathy caused by an abnormal immune response to dietary gluten that occurs in genetically susceptible individuals[18].

For the great majority of human beings, cereals represent an important source of nutrients, whereas for CD patients' certain cereal products represent poisons that not only destroy small intestinal mucosa, as reported in Figure 3[19], but also predispose to gastrointestinal malignancy[20].
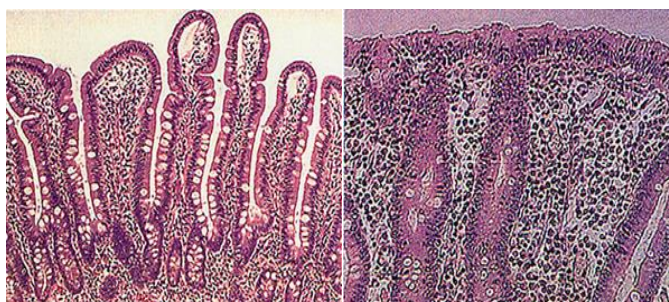
Figure 3: Degeneration of intestinal villi. Normal villi were represented in the picture on the left and CD villi on the right.

These products constitute the storage protein fraction of the endosperm of the grains which, in the wheat, is classified into four classes depending on their solubility: albumins that are soluble in water, globulins soluble in salt solution, gliadins soluble in alcohol solution and glutenins insoluble in neutral aqueous or saline solution and ethanol.

Gluten, which is provided by both gliadins and glutenins, is the product of a ball of wheat flour dough that has been exhaustively washed in tap water,

and the baking qualities of the wheat depends on its ability to trap carbon dioxide in dough.

The gluten peptides and the related prolamins are responsible for triggering mucosal lesions in CD patients[21]. Despite the recent progresses in understanding the molecular mechanisms of mucosal lesions, it remains unknown how increased amounts of gluten peptides can enter the intestinal mucosa to initiate the inflammatory cascade. Current knowledge indicates that different gluten peptides are involved in the disease processes as both 'toxic' and 'immunogenic'. The peptides defined as 'toxic' are able to induce mucosal damage either when added in culture to duodenal endoscopic biopsy or when administered in vivo, while those defined as 'immunogenic' are able to specifically stimulate HLA-DQ2- or DQ8-restricted T cell clones isolated from jejunal mucosa or peripheral blood of coeliac patients. These peptides can trigger two immunological pathways: one is thought to be a rapid effect on the epithelium that involves the innate immune response and the other represents the adaptive immune response involving CD4+ T cells in the lamina propria that recognize gluten epitopes processed and presented by antigen presenting cells.

3.1.4 Aim of the study

Celiac disease is a chronic inflammation of the small intestine. In presence of an inflammatory stimulus, enzymes belonging to the phospholipase class release arachidonic acid from membrane phospholipids such as: phosphatidylethanolamines (PE), phosphatidylcholines (PC), phosphatidylinositol (PI), phosphatidylserine (PS), phosphstidylglycerol (PG)(Figure 4)[22]. Arachidonic acid is the main precursor of eicosanoids, substances involved in the inflammatory response of the organism.
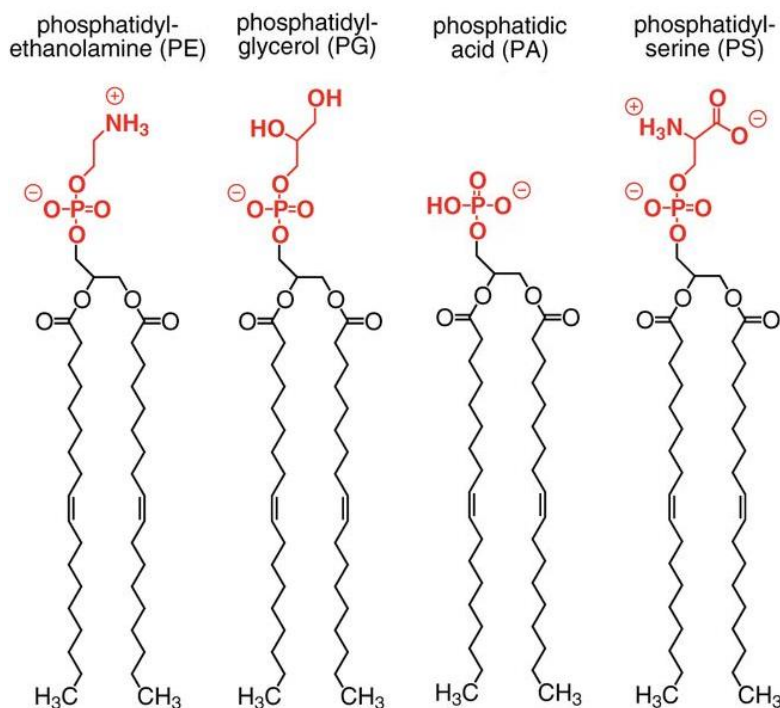
Figure 4: Structure of different classes of phospholipids: phosphatidylethanolamines (PE), phosphstidylglycerol (PG), phosphatidic acid (PA), phosphatidylserine (PS).

The diagnosis of celiac disease is often delayed in time and it may take several months or years for the patient and the health care professional to think of it as the cause of many non-specific symptoms. History and physical examination may give direction as to the diagnosis.

When the diagnosis is suspected, there is a two-step screening process to make the diagnosis. First a blood test for immunoglobulin A anti-tissue transglutaminase antibody (IgA TTG). Testing for antiendomysial antibody may also be considered. If the screening blood test is positive, then endoscopy and biopsy of the lining of the duodenum is requested.

The aim of this project concerns the development and optimization of a simple method, based on mass spectrometry in Multiple Reaction Monitoring (MRM) ion mode to quantify 83 phospholipids in sera samples.

In collaboration with Prof. L. Greco and R. Auricchio, from Dipartimento di Scienze Mediche Traslazionali from Federico II University of Naples, and Professor M. Ruoppolo from CEINGE of Naples the optimized method has been applied to sera from 23 CD children and 23 CTRL children.

In particular, it was explored the serum phospholipid profile of children at risk for celiac disease from the 4th month of life to look for a specific lipid profile that may predict the onset of celiac disease before the 6 years of age.

3.2 Material and Methods

3.2.1 Material

Chemials and reagents

D-α-phosphatidylcholine, dipalmitoyl; L-α- phosphatidylethanolamine, dioleoyl were purchased from Avanti Polar Lipids.

Acetonitrile (ACN) and methanol (MeOH) were purchased from Sigma-Aldrich. Acetic acid (AcA) was purchased from Baker. Ammonium acetate (AA) was purchased from Carlo Erba.

Samples cohorts

This study was carried out in collaboration with Prof. Luigi Greco and Renata Auricchio from Dipartimento di Scienze mediche traslazionali of University of Naples Federico II.

Frozen Serum samples from:

23 Children, at risk for celiac disease, who became celiac (CD), sampled at 4, 12 and > 24months

23 Children, at risk for celiac disease, who did not become celiac (CTRL), sampled at 4 and 12 months.

All children were gradually introduced to gluten from the 6th months of life. None had received gluten at 4 months.

3.2.2 Methods

Extraction

Aliquots of 100 µL human serum were diluted to 990 µL in chloroform/MeOH 17:1 v/v for metabolites extraction. After sonication the samples were centrifugated for 5 min at 13000 rpm. Once all the liquid phase has been removed, the pellets were diluted to 990 µL of chloroform/MeOH 2:1 v/v and centrifugated for 10 min at 13000 rpm. The supernatant was recovered and stored at -20 ° C until the LC-MRM/MS analysis.

Liquid chromatography and MRM mass spectrometry method set up

The samples were analysed by liquid chromatography (Eksigent expressHT ultra, HPLC) coupled to electrospray ionization on an AB sciex QTrap 4000, operating with Analyst 1.6.2 (AB SCIEX). Chromatographic separation was achieved on a Halo C18 1.0 mm x 50 mm, 2.7 µm column

using a flow rate of 40 µL/min at 40 °C during 12 min gradient (0 min 70% A, 2min 50% A, 9min 5% A, 11 min 50% A, 12 min 70% A).

5 µL of the extract was injected and each sample was analyzed twice. The auto sampler was cooled at 4 °C. Solvent composition was ACN/MeOH/AcA 97:2:1 v/v/v with AA 5mM for solvent A; and MeoH/AcA 99:1 v/v with AA 5mM for solvent B. The mass analysis was performed in both positive (ESI+) and negative (ESI-) mode.

For each phospholipids compound, precursor ions, product ions and optimal collision energies were established for each MRM transition using the Analyst software. The instrument was operated with the following settings: CUR 20, CAD 5, IS 4500, TEMP 380, GS1 25, GS2 24. Two methods were set up to investigate the presence of different classes of phospholipids the first able to work in positive polarity and the second one in negative. Table 1 and 2 collected the instrumental parameters used to set up the two MRM/MS methods for the identification and quantitation of 83 target molecules.

Table 1: Parameters to set up the MRM Method 1 for the identification of LPC and PC phospholipid classes. The best transitions and Declustering Potential (DP), Collision Energy (CE) and Ectraction Potential (CXP) are reported.

ANALYTES	Q1	Q3	DP	CE	CXP
LPC22:1	592,4	184,1	40	47	9
LPC22:0	594,4	184,1	40	47	9
LPC24:1	620,4	184,1	40	47	9
LPC24:0	622,4	184,1	40	47	9
LPC26:1	648,5	184,1	40	47	9
LPC26:0	650,5	184,1	40	47	9
PC28:2	674,5	184,1	40	47	9
PC28:1	676,5	184,1	40	47	9
PC28:0	678,5	184,1	40	47	9
PC30:2	702,5	184,1	40	47	9
PC30:1	704,5	184,1	40	47	9
PC30:0	706,5	184,1	40	47	9
PC32:2	730,5	184,1	40	47	9
PC32:1	732,6	184,1	40	47	9
PC32:0	734,6	184,1	40	47	9
PC34:2	758,6	184,1	40	47	9
PC34:1	760,6	184,1	40	47	9
PC36:2	786,6	184,1	40	47	9
PC36:1	788,6	184,1	40	47	9
PC36:0	790,6	184,1	40	47	9
PC40:4	838,6	184,1	40	47	9
PC42:5	850,6	184,1	40	47	9
PC(O-36:0)	776,6	184,1	40	47	9
PC(O-38:3)	798,6	184,1	40	47	9
PC(O-38:0)	804,6	184,1	40	47	9
PC(O-40:6)	820,6	184,1	40	47	9
PC(O-40:5)	822,6	184,1	40	47	9
PC(O-40:1)	830,7	184,1	40	47	9
PC(O-42:5)	864,6	184,1	40	47	9
PC(O-42:3)	854,6	184,1	40	47	9
PC(O-42:0)	860,7	184,1	40	47	9

Table 2: Parameters to set up the MRM Method 1 for the identification of PA, PE, PG, PS and PI phospholipid classes. The best transitions and Declustering Potential (DP), Collision Energy (CE) and Extraction Potential (CXP) are reported.

ANALYTES	Q1	Q3	DP	CE	CXP
PA 28:1	597,4	171,1	-180	-50	-13
	597,4	199,2	-180	-50	-13
	597,4	253,2	-180	-50	-13
	597,4	281,2	-180	-50	-13
PA 28:0	590,4	171,1	-180	-50	-13
	590,4	199,2	-180	-50	-13
	590,4	255,2	-180	-50	-13
	590,4	283,2	-180	-50	-13
PA 30:1	616,4	227,2	-180	-50	-13
	616,4	253,2	-180	-50	-13
PA 30:0	618,4	199,2	-180	-50	-13
	618,4	227,2	-180	-50	-13
	618,4	255,2	-180	-50	-13
	618,4	283,2	-180	-50	-13
PA 32:2	642,4	253,2	-180	-50	-13
PA 32:1	644,4	227,2	-180	-50	-13
	644,4	253,2	-180	-50	-13
	644,4	281,3	-180	-50	-13
PA 32:0	646,4	255,2	-180	-50	-13
PA 34:2	670,5	253,2	-180	-50	-13
	670,5	281,3	-180	-50	-13
PA 34:1	672,5	253,2	-180	-50	-13
	672,5	283,3	-180	-50	-13
	672,5	281,3	-180	-50	-13
	672,5	255,2	-180	-50	-13
PA 36:2	698,5	281,3	-180	-50	-13
PA 36:1	700,5	281,3	-180	-50	-13
	700,5	283,3	-180	-50	-13
PE 28:1	631,4	199,2	-160	-48	-11
	631,4	253,2	-160	-48	-11
PE 28:0	633,4	199,2	-160	-48	-11

	633,4	255,5	-160	-48	-11
PE 30:2	657,5	225,2	-160	-48	-11
	657,5	253,2	-160	-48	-11
PE 30:1	659,5	199,2	-160	-48	-11
	659,5	227,2	-160	-48	-11
	659,5	253,2	-160	-48	-11
	659,5	281,3	-160	-48	-11
PE 30:0	661,5	199,2	-160	-48	-11
	661,5	227,2	-160	-48	-11
	661,5	255,2	-160	-48	-11
	661,5	283,2	-160	-48	-11
PE 32:2	685,5	253,3	-160	-48	-11
PE 32:1	687,5	227,2	-160	-48	-11
	687,5	281,3	-160	-48	-11
	687,5	255,2	-160	-48	-11
	687,5	253,2	-160	-48	-11
PE 34:2	713,5	253,2	-160	-48	-11
	713,5	281,3	-160	-48	-11
PE 34:1	715,5	255,2	-160	-48	-11
	715,5	281,3	-160	-48	-11
PE 36:2	741,5	281,3	-160	-48	-11
PE 36:1	743,5	281,3	-160	-48	-11
	743,5	283,3	-160	-48	-11
PG 32:2	716,5	253,2	-130	-48	-11
PG 32:1	718,5	253,2	-130	-48	-11
	718,5	255,2	-130	-48	-11
PG 34:2	744,5	253,2	-130	-48	-11
	744,5	281,3	-130	-48	-11
PG 34:1	746,6	255,2	-130	-48	-11
	746,6	281,3	-130	-48	-11
PS 32:3	729,5	253,2	-90	-60	-33
PS 32:1	731,5	227,2	-90	-60	-33
	731,5	253,2	-90	-60	-33
	731,5	255,2	-90	-60	-33
	731,5	281,3	-90	-60	-33
PS 32:0	733,5	255,2	-90	-60	-33

PS 34:2	757,6	253,2	-90	-60	-33
	757,6	281,3	-90	-60	-33
	757,6	255,2	-90	-60	-33
PS 34:1	759,6	255,2	-90	-60	-33
	759,6	281,3	-90	-60	-33
PS 32:2	785,6	281,3	-90	-60	-33
PS 36:0	789,6	283,3	-90	-60	-33
PI 26:1	722,5	171,1	-85	-62	-10
	722,5	253,2	-85	-62	-10
PI 26:0	724,5	171,1	-85	-62	-10
	724,5	199,2	-85	-62	-10
	724,5	227,2	-85	-62	-10
	724,5	255,2	-85	-62	-10
PI 28:1	750,6	199,2	-85	-62	-10
	750,6	225,2	-85	-62	-10
	750,6	227,2	-85	-62	-10
	750,6	253,2	-85	-62	-10
PI 28:0	752,6	171,1	-85	-62	-10
	752,6	199,2	-85	-62	-10
	752,6	227,2	-85	-62	-10
	752,6	255,2	-85	-62	-10
	752,6	283,3	-85	-62	-10
PI 30:2	776,6	253,2	-85	-62	-10
	776,6	225,2	-85	-62	-10
PI 30:1	778,6	199,2	-85	-62	-10
	778,6	225,2	-85	-62	-10
	778,6	227,2	-85	-62	-10
	778,6	253,2	-85	-62	-10
PI 30:0	780,6	199,2	-85	-62	-10
	780,6	227,2	-85	-62	-10
	780,6	255,2	-85	-62	-10
	780,6	283,3	-85	-62	-10
PI 32:2	804,6	253,2	-85	-62	-10
PI 32:1	806,6	227,2	-85	-62	-10
	806,6	225,2	-85	-62	-10
	806,6	283,3	-85	-62	-10
	806,6	253,2	-85	-62	-10

PI 32:0	808,6	255,2	-85	-62	-10
PI 34:2	832,6	253,2	-85	-62	-10
	832,6	281,3	-85	-62	-10
PI 34:1	834,6	253,2	-85	-62	-10
	834,6	255,2	-85	-62	-10
	834,6	281,3	-85	-62	-10
	834,6	283,3	-85	-62	-10
PI 34:0	836,6	255,2	-85	-62	-10
	836,6	283,3	-85	-62	-10
PI 36:2	860,7	281,3	-85	-62	-10
PI 36:1	862,7	281,3	-85	-62	-10
PI 36:0	864,7	283,3	-85	-62	-10

As the Table 1 clearly show, all the monitored phosphatidylcholines report the same product ion (184,1 m/z). Indeed, for the identification of these molecules is peculiar the loss of choline due to the CID fragmentation in the collision cell.

Solutions of known concentration of the standards were prepared to calculate the calibration curve in the range of concentration 0,1-250 µg/L for both D-α-Phosphatidylcholine, Dipalmitoyl and L-α-Phosphatidylethanolamine, Dioleoyl.

The analyses were performed in triplicate for each point of the calibration curve. Recovery was estimated by measuring the concentration of standards added to real samples (spiked samples).

$$\textbf{\textit{Recovery}} \, (\%) = \frac{C1 - C2}{C3} * 100$$

where:
C1: analyte concentration measured after the addition
C2: analyte concentration measured before the addition
C3: added concentration of the standard.

3.2.3 Method validation

The Limit of Detection (LOD) and the Limit of Quantification (LOQ) are defined, respectively, as the concentration level at which the analyte can be reliably detected in the sample under consideration and as the level at which

the analyte can be detected and quantified with sufficient precision. Several methods are used to determine LOD and LOQ[23]. Determination of the signal-to-noise ratio is the simplest and is performed by comparing measured signals from samples with known low concentrations of analyte with those of blank samples and establishing the minimum concentration at which the analyte can be reliably detected. A signal-to-noise ratio between 3 or 2:1 is generally considered acceptable for estimating the detection limit, a typical signal-to-noise ratio is 10:1 is acceptable for estimating the quantification limit [24].

Possible matrix effects were evaluated by comparing standard and matrix matched calibration curves for each analyte. Standard solutions were prepared as described. The calibration curves were repeated three times. Matrix effects were evaluated by comparing five points standard and matrix-matched calibration curves.

Measurement uncertainty was estimated according to the law of error propagation. The uncertainty contributions considered in the combined uncertainty were the uncertainties of the preparation of native standard solutions for instrument calibration, the uncertainty contribution deriving from instrument calibration, the uncertainty coming from the precision of the analyses and the uncertainty of bias.

3.3 Results and discussion

3.3.1 Method set up

5 ul of a mixture of targeted metabolites standard solution at different concentrations was separated in a single run by a LC-MSMS. The total ion current (TIC) chromatogram as a function of retention time was registered from detector, showing a good peak shape for each analyte.

Identification and quantification are based on extracted ion chromatogram (XIC) signals of different transitions (Figure 5). Thanks to the external standard method, the two calibration curves for L-α-PC, dipalmitoyl and PE, dioleoyl were set up.
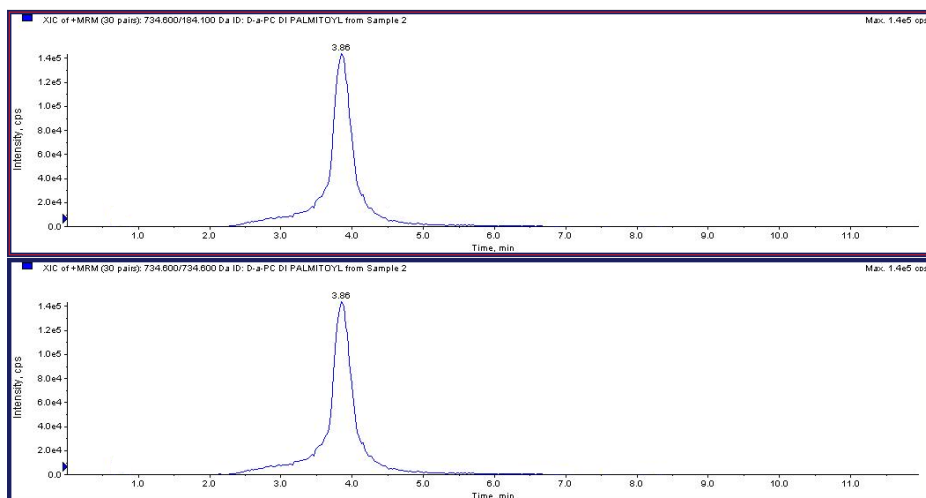
Figure 5: MRM Chromatograms for the selected transitions 734.6→184.1 and 734,6→734,6 recorded for D-α-Phosphatidylcholine, dipalmitoyl.
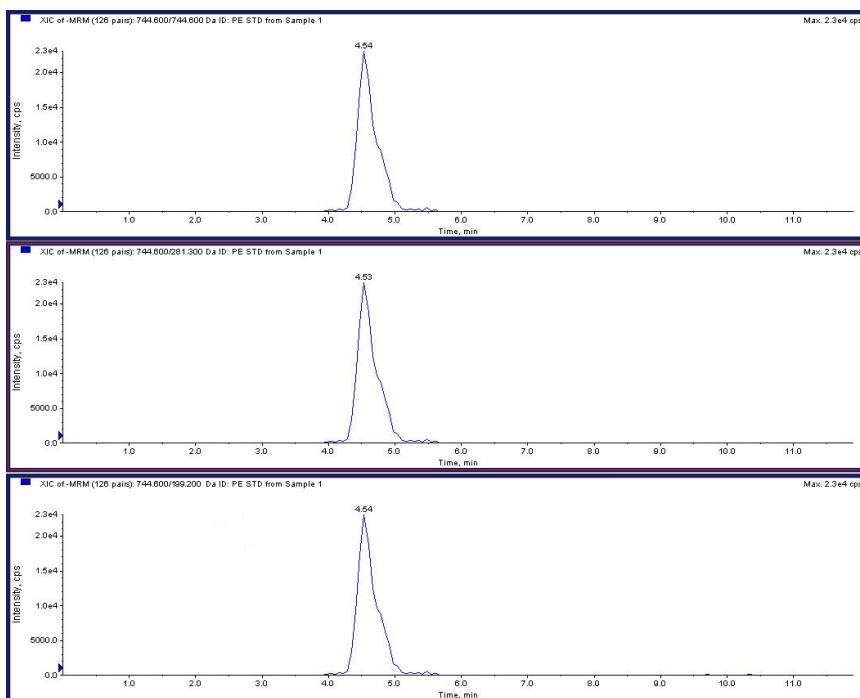
Figure 6: MRM Chromatograms for the selected transitions 744.6→199.2; 744.6→281,3; 744.6→744.6 recorded for L-α-Phosphatidylethanolamine, Dioleoyl.

Figures 5 and 6 showed the monitored transitions for the two standards; they are perfectly coeluted at 3.05 and 4.54 min respectively. For each standard a   quantifier transition was selected because it gave the best associated instrumental response and the absence of nonspecific signals (Dα-Phosphatidylcholine, Dipalmitoyl: "734.6→184.1", L-α-Phosphatidylethanolamine, Dioleoyl: "744.6→281.3").

The calibration curves were set up in a concentration range of 0,1-250 μg for both PC and PE; each point of the line was analysed in triplicate and the obtained curves are shown in Figure 6.
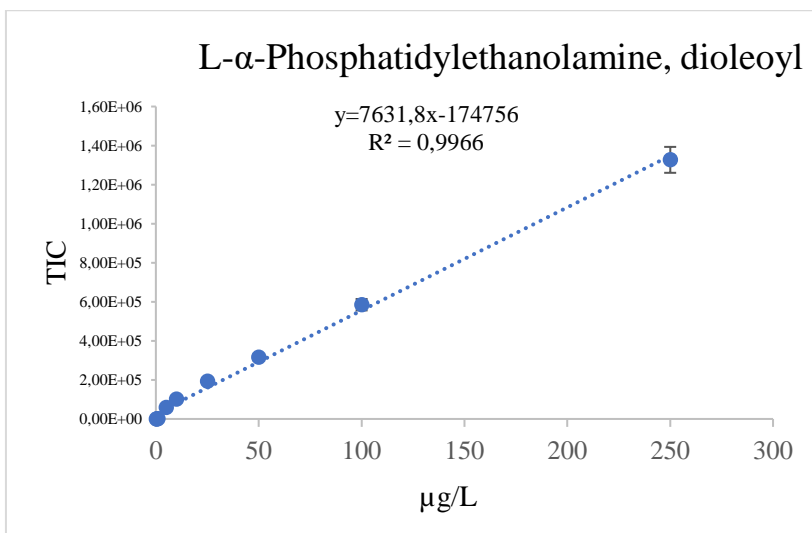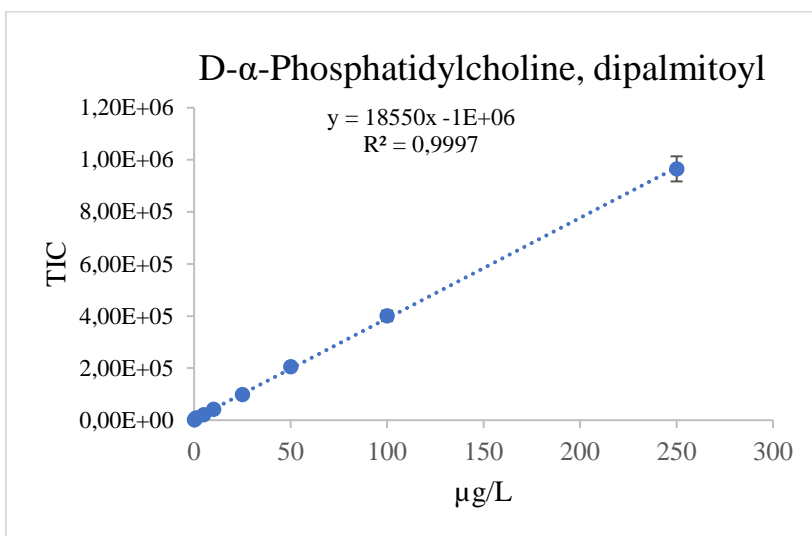
Figure 7 Calibration curves obtained from chromatogram peaks area integration of each transition as a function at different concentrations

For the quantitative analysis of LPC and PC the calibration curve obtained by PC was used; while for the PA, PE, PI and PS the curve obtained for PE was used.

The parameters of LOD, LOQ, recovery (%) and linearity are reported in Table 3.

Standard	Linear Range (pg/μL)	Recovery (%)	LOD (pg/μL)	LOQ (pg/μL)
PC	0,1-250	87	0,1	0,3
PE	0,1-250	88	0,1	0,3

3.3.2 Method validation on CD and CTRL infant sera samples

As described in Materials and Methods section, serum samples of children genetically predisposed to celiac disease have been treated.

The cohort of sera used for the validation of the MRM method is characterized by 23 genetically predisposed children for whom are available serum samples at 4 months, before the introduction of gluten in the diet, at 12 months, following weaning and then with the introduction of gluten and finally a variable Td time (between 15 and 63 months) corresponding to the time of diagnosis of celiac disease.

4-Months CD vs CTRL

The treated 4-month CD and CTRL serum samples were analysed in duplicate using the previously shown MRM method and Table 1 in the Supporting Data summarize the results as averages of the two replicates with dev std and CV (%) values.

Despite the high complexity of the analysed biological matrix, it was possible to identify and quantify many of the target molecules with a high reproducibility of the data. The data obtained were summarized in the following histograms one for each phospholipid class investigated (Figure 6). Data of children who developed celiac disease (CD) are indicated with blue bars while data of children who have not yet developed the pathology (CTRL) with orange bars. CTRL samples were used as control ones.

Each bar in Figure 8 represents a class of target molecules.

FFigura 8: *Comparison of children's serum samples at 4 months before gluten introduction. The average data for CD children is shown in blue and the average data for CTRL children is shown in orange.*

It can be seen how the PI 34: 1 and the molecules contained in the LPC, PC (3) and PS panels are increased in CD compared to CTRL already before the introduction of gluten in the child's diet.

Therefore, it could be considered as a potential molecular signature for celiac disease in genetically predisposed children. Obviously subsequent studies on a higher samples number are necessary to evaluate both the dosage normal range of the target molecules and the use of this biomarker in normal clinical practice even in subjects whose family history is unknown.

12-Months CD vs CTRL

CD and CTRL samples collected at 12 months were treated as described in the Materials and Methods section and analysed with the LC/MRM/MS method. Each sample was analysed in duplicate.

The averages of the obtained results are collected in Table 3 of Supporting Data. This set of samples allows the evaluation of the trend of the target molecules following the introduction of gluten in the daily diet. In Figure 8 the average of the concentration of each target analyte and the associated standard deviation is reported as histogram. The blue bars correspond to CD children while the orange bars correspond to CTRL children, patients who have not yet developed the pathology.
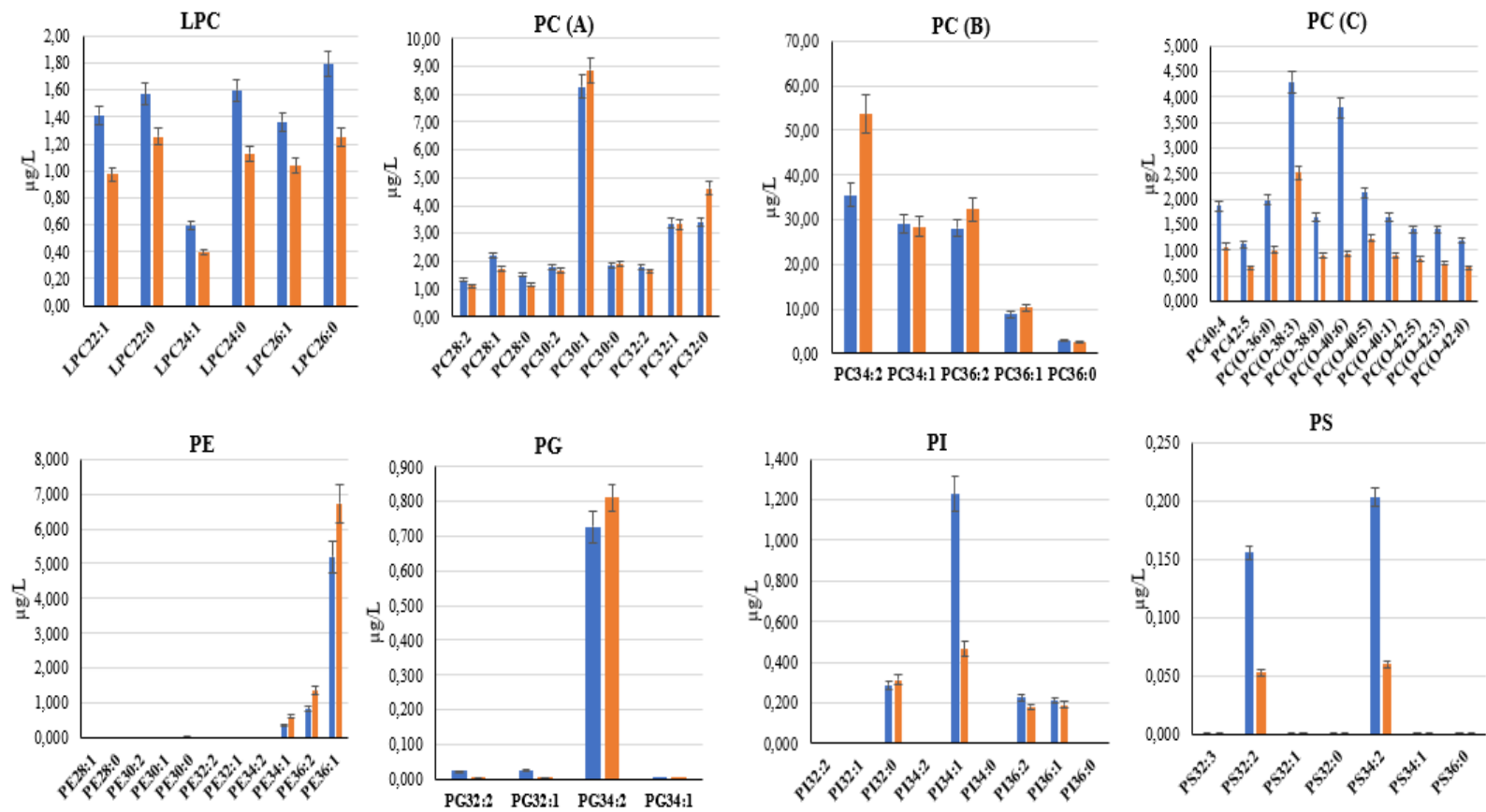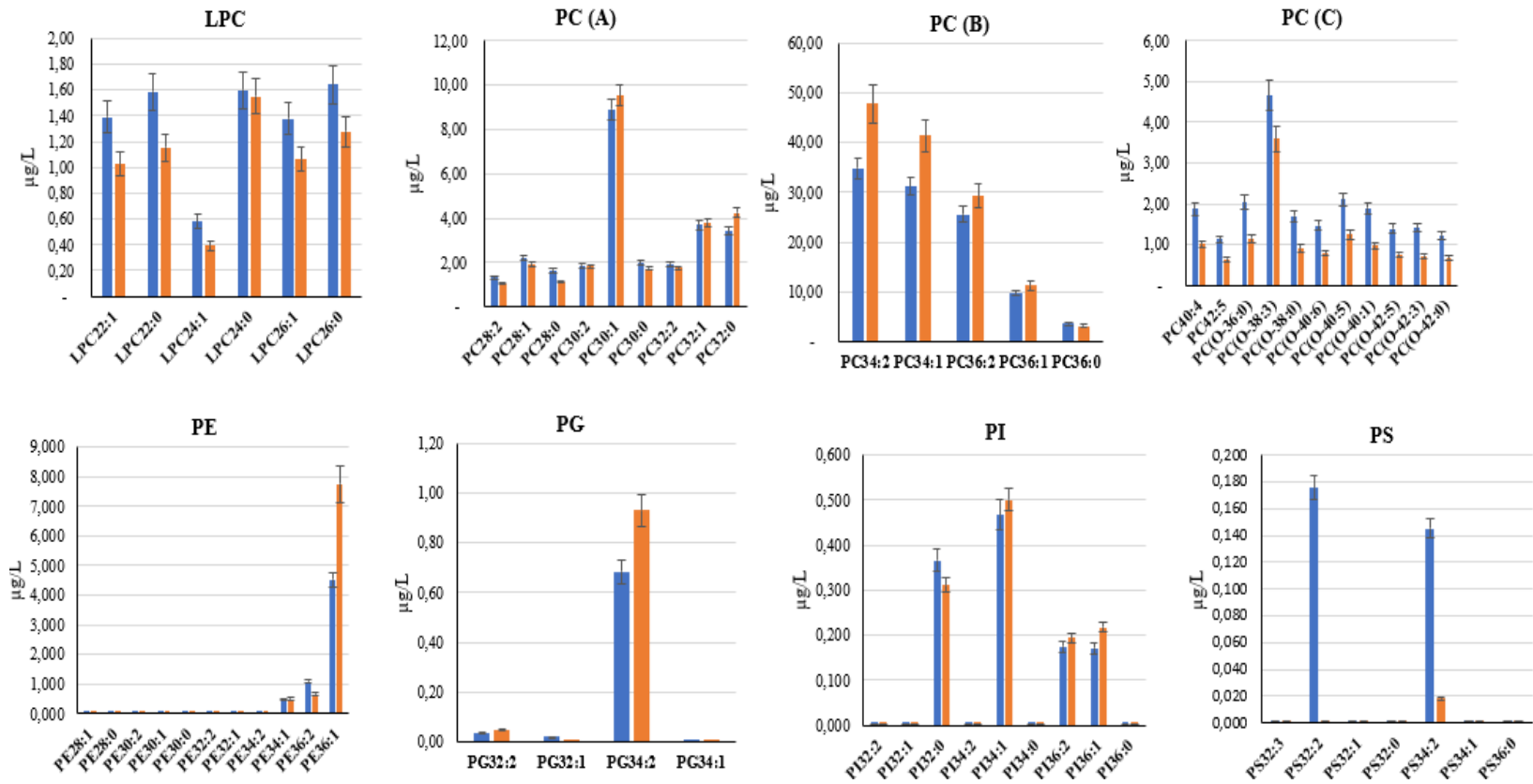
Figura 9: Comparison of children's serum samples at 12 months after gluten introduction. The average data for CD children is shown in blue and the average data for CTRL children is shown in orange.

As can be expected, with the introduction of gluten in the child's diet, there is a variation in the production of some classes of lipids. LPC and PC panels represent some lipid classes that increase with gluten introduction, while other lipids, on the other hand, are diminished, like PGs or PEs.

Interestingly that emerges from this study is that the change in some of these molecules is much more pronounced in individuals who have developed the disease (CD) than those who have not yet developed (CTRL).

4,12 and Time of Diagnosis (Td)

Finally, the last sera analysed were those collected between 15 and 63 months, called diagnosis time, or time span in which celiac disease was diagnosed for some of the selected subjects.

Table 4 of the supporting data contains the mean values, the standard deviation and the coefficient of variation (%) relative to the analyses carried out in duplicate on this series of sera.

As an example, data comparing the concentration of monitored lipids for CD subjects (at T4, T12 and Td) are summarized in the histograms of Figure 9. It is evident how, at the time of diagnosis, the monitored lipid values seem to settle around the values recorded at 12 months.
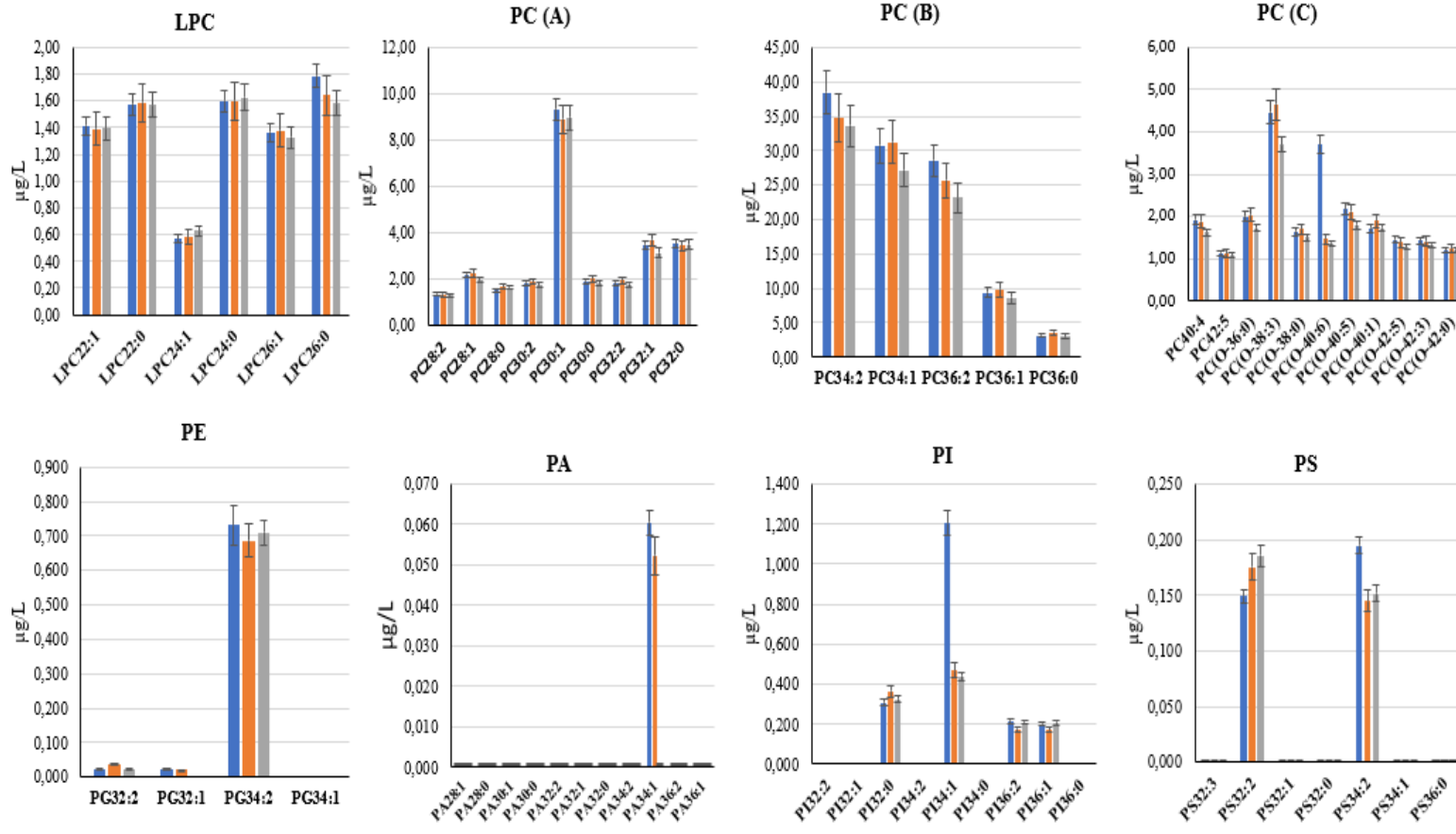
Figura 10: Comparison CD serum samples at 4, 12, and t >12 months after gluten introduction. In blue the mean results of 4-months samples, in orange the mean results of 12-months samples and in grey the mean value recorded at t < 12 months, the time of coeliac disease diagnosis.

3.4 Conclusions

The simplicity of the quantification method allowed the analysis of eighty-three phospholipids in a single run of twelve min makes the developed method well suited not only for research work about but also for the measurement of clinical like of patients affected from celiac disease.

The breakthrough of this technology is highlighted in its capability of providing a universal approach for developing assays of quantification for a wide spectrum of target molecules with minimum restrictions, and the ease of assembling multiplex detection in a single measurement.

The children of the two cohorts do share a similar genetic background, since they come from families with one celiac proband and bear the specific HLA haplotype (DQ2 or DQ8). So, it does appear that the genetic profile may not explain fully the great differences found between the two cohorts. This investigation let to possible selection of a panel of molecules that can be used as putative serum biomarkers for coeliac disease to be investigated in further studies.

3.5 References

(1)   Van Meer, G. Cellular lipidomics. *The EMBO journal* **2005**, *24*, 3159-3165.

(2)   Plakkal Ayyappan, J.; Paul, A.; Goo, Y. H. Lipid droplet-associated proteins in atherosclerosis. *Molecular medicine reports* **2016**, *13*, 4527-4534.

(3)   Morita, M.; Kuba, K.; Ichikawa, A.; Nakayama, M.; Katahira, J.; Iwamoto, R.; Watanebe, T.; Sakabe, S.; Daidoji, T.; Nakamura, S. The lipid mediator protectin D1 inhibits influenza virus replication and improves severe influenza. *Cell* **2013**, *153*, 112-125.

(4)   Fahy, E.; Cotter, D.; Sud, M.; Subramaniam, S. Lipid classification, structures and tools. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* **2011**, *1811*, 637-647.

(5)   Lydic, T. A.; Goo, Y.-H. Lipidomics unveils the complexity of the lipidome in metabolic diseases. *Clinical and translational medicine* **2018**, *7*, 4.

(6)   Brügger, B. Lipidomics: analysis of the lipid composition of cells and subcellular organelles by electrospray ionization mass spectrometry. *Annual review of biochemistry* **2014**, *83*, 79-98.

(7)   Holzer, M.; Birner-Gruenberger, R.; Stojakovic, T.; El-Gamal, D.; Binder, V.; Wadsack, C.; Heinemann, A.; Marsche, G. Uremia alters HDL composition and function. *Journal of the American Society of Nephrology* **2011**, ASN. 2010111144.

(8)   Umetani, M.; Ghosh, P.; Ishikawa, T.; Umetani, J.; Ahmed, M.; Mineo, C.; Shaul, P. W. The cholesterol metabolite 27-hydroxycholesterol promotes atherosclerosis via proinflammatory processes mediated by estrogen receptor alpha. *Cell metabolism* **2014**, *20*, 172-182.

(9)   Seddon, J. M.; Cevc, G.: *Lipid polymorphism: structure and stability of lyotropic mesophases of phospholipids*; Marcel Dekker, Inc: New York, NY, USA, 1993.

(10)   van Hoogevest, P.; Wendel, A. The use of natural and synthetic phospholipids as pharmaceutical excipients. *European journal of lipid science and technology* **2014**, *116*, 1088-1107.

(11)   Küllenberg, D.; Taylor, L. A.; Schneider, M.; Massing, U. Health effects of dietary phospholipids. *Lipids in health and disease* **2012**, *11*, 3.

(12)   Li, J.; Wang, X.; Zhang, T.; Wang, C.; Huang, Z.; Luo, X.; Deng, Y. A review on phospholipids and their main applications in drug delivery systems. *Asian journal of pharmaceutical sciences* **2015**, *10*, 81-98.

(13)   Majerus, P. W.; Ross, T. S.; Cunningham, T. W.; Caldwell, K. K.; Jefferson, A. B.; Bansal, V. S. Recent insights in phosphatidylinositol signaling. *Cell* **1990**, *63*, 459-465.

(14)   Cole, L. K.; Vance, J. E.; Vance, D. E. Phosphatidylcholine biosynthesis and lipoprotein metabolism. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* **2012**, *1821*, 754-761.

(15)   Vance, J. E.; Vance, D. E.: *Biochemistry of lipids, lipoproteins and membranes*; Elsevier, 2008.

(16)   Shields, D. J.; Lehner, R.; Agellon, L. B.; Vance, D. E. Membrane topography of human phosphatidylethanolamine N-methyltransferase. *Journal of Biological Chemistry* **2003**, *278*, 2956-2962.

(17)   Fasano, A.; Berti, I.; Gerarduzzi, T.; Not, T.; Colletti, R. B.; Drago, S.; Elitsur, Y.; Green, P. H.; Guandalini, S.; Hill, I. D. Prevalence of celiac disease in at-risk and not-at-risk groups in the United States: a large multicenter study. *Archives of internal medicine* **2003**, *163*, 286-292.

(18)   Mäki, M.; Mustalahti, K.; Kokkonen, J.; Kulmala, P.; Haapalahti, M.; Karttunen, T.; Ilonen, J.; Laurila, K.; Dahlbom, I.; Hansson, T. Prevalence of celiac disease among children in Finland. *New England Journal of Medicine* **2003**, *348*, 2517-2524.

(19)   Sollid, L. M. Coeliac disease: dissecting a complex inflammatory disorder. *Nature Reviews Immunology* **2002**, *2*, 647.

(20)   Schuppan, D. Current concepts of celiac disease pathogenesis. *Gastroenterology* **2000**, *119*, 234-242.

(21)   Corrao, G.; Corazza, G. R.; Bagnardi, V.; Brusco, G.; Ciacci, C.; Cottone, M.; Guidetti, C. S.; Usai, P.; Cesari, P.; Pelli, M. A. Mortality in patients with coeliac disease and their relatives: a cohort study. *The Lancet* **2001**, *358*, 356-361.

(22)   Niessen, W. M.: *Liquid chromatography-mass spectrometry*; Crc Press, 2006.

(23)   Sleno, L.; Volmer, D. A. Ion activation methods for tandem mass spectrometry. *Journal of mass spectrometry* **2004**, *39*, 1091-1112.

(24)   Calder, P. C. Polyunsaturated fatty acids and inflammatory processes: new twists in an old tale. *Biochimie* **2009**, *91*, 791-795.

Chapter 4

4.1 Introduction

Inflammation is typically associated with secretion of specific cytokines and chemokines which coordinate cell and tissue activities.[1] Toll-like receptors (TLRs) are receptors that have the prominent biological function to promote synthesis and release of cytokines to trigger inflammatory response[2]. TLR4 was the first identified mammalian TLR, belonging to cytomembrane type I transmembrane glycoprotein receptor, and can be activated by the lipopolysaccharides (LPS) of pathogenic microorganisms, lipoproteins and peptidoglycans, to promote cells to produce cytokines, chemokines, adhesion molecules and acute phase protein to regulate the inflammatory responses[3].
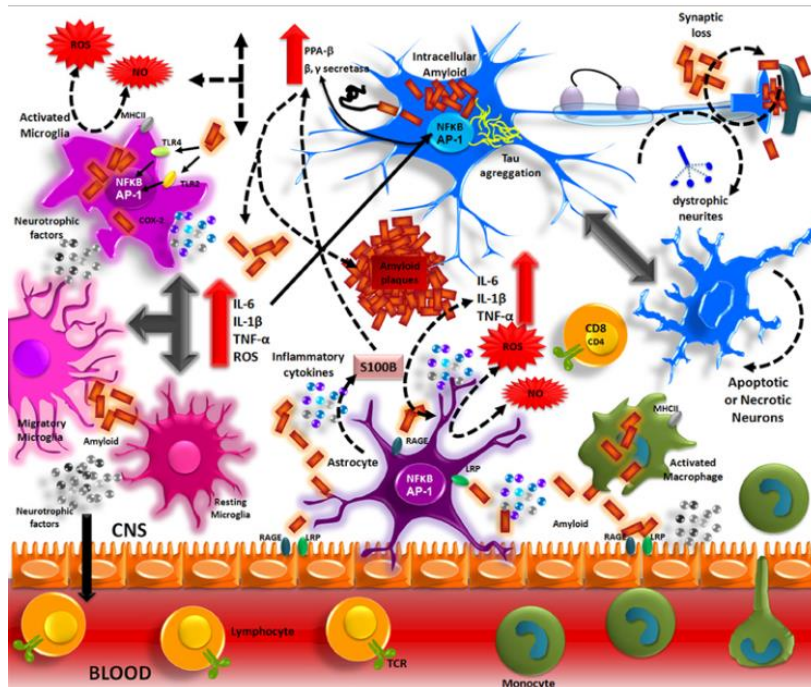
Figure 9:**Schematization of inflammatory process**. The Aβ peptide produced by APP processing, form aggregates that activate microglia through TLRs and RAGE receptors. These receptors in turn, activate NF-κ B and AP-1 transcription factors, which induce the reactive oxygen species (ROS) production and the expression of inflammatory cytokines (IL-1, IL-6, TNF). These inflammatory factors directly acting on the neurons and also stimulate the astrocytes, which amplify the pro-inflammatory signals, inducing a neurotoxic effect. The inflammatory mediators generate by resident CNS cells, induce the production of adhesion molecules and chemokines, which recruit peripheral immune cells[4].

Lipopolysaccharide (LPS) is one of the most powerful bacterial virulence factors in terms of proinflammatory properties[5]. Endotoxin, a cell wall component of Gram-negative bacteria, plays a central role in the pathogenesis of septic shock. By administering small doses of intravenous endotoxin to humans, a variety of acute inflammatory responses are induced which are qualitatively similar to those that occur during the early stages of septic shock.
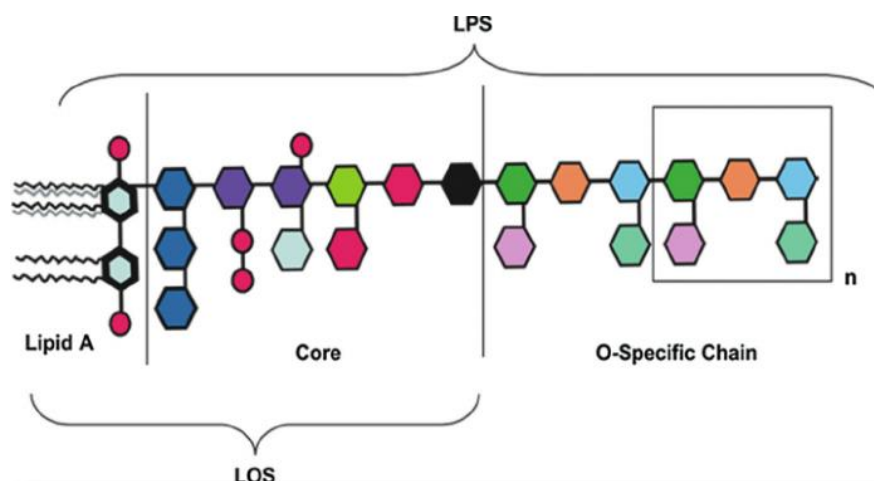
Figure 10: General chemical structure of LPS from Gram-negative bacteria.

In this study mass spectrometry in Multiple Reaction Monitoring (MRM) Ionization mode has been used to develop a method able to to quantify relevant proteins which were found to be specifically induced in response to *E.Coli* LPS inflammatory activation in THP I differentiated cells.

THP-1cells are monocytes derived from human monocytic leukaemia,and are currently predominantly used for the study of the pathogenesis of *Streptococcus pneumoniae*, *Mycobacterium marinum* and infection caused by others pathogenic agents[6,7].

Key advantages of MRM assays are the ability to target specific peptide sequences, including variants and modified forms, and the ability for multiplexing that allows analysis of hundreds of peptides during a single sample injection. The targeted approach of an MRM experiment provides an increase in sensitivity, selectivity and specificity[8], avoiding the problems related to antibody availability for all target molecules and cross reactivity of immunoenzimatic assays.

In this paper, we report the development of MRM/MS assay to quantify Tumour Necrosis Factor-α (TNF-α), Interferon-γ (INF-γ), Interleukin-8 (IL-8) and Interleukin-10 (IL-10) in THP I differentiated cells after stimulation with *E.Coli* LPS to induce inflammation.

A time course analysis has been performed after 2h, 4h, 6h, 9h and 24 h *E.Coli* LPS stimulation to investigate the trend of expression of the four target proteins.

4.2 Material and Methods

4.2.1 Chemicals and reagents

Interleukin 8, Interleukin 10, Tumor Necrosis Factor-α and Interferon-γ standard proteins, guanidine, Tris(idrossimetil)amminometane chloridrate (TrisHCl), dithiothreitol (DTT); ethylendiaminetetraacetate (EDTA), trypsin, iodoacetamide (IAM), ammonium bicarbonate (AMBIC), trichloroacetic acid (TCA) were purchased from Sigma-Aldrich. Formic acid (HCOOH), methanol, chloroform, acetonitrile (ACN) are from J.T. Baker. Pipette tips C18 (zip tip) and centrifugal filter units were purchased from Merck Millipore.

4.2.2 Cell culture

THP-1 monocyte cells were from ATCC, Manassas, VA. Cells were cultured in RPMI-1640 (Sigma-Aldrich, Milan, Italy), supplemented with 10% foetal bovine serum (HyClone, GE Healthcare Lifescience, Chicago, IL), L-glutamine and antibiotics, in a 5% CO_2 humidified atmosphere at 37 °C. THP-1 cells were differentiated by adding Phorbol 12-myristate 13-acetate (PMA) (Sigma Aldrich, Milan, Italy) to macrophages.

Lipopolysaccharides (LPS)

Lipopolysaccharides from *E. coli* was purchased from Sigma-Aldrich, Milan, Italy

Antibodies

Anti-IκB, anti-TNF-α and anti-β-actin primary antibodies where purchased from Santa Cruz Biotechnology (Dallas, Texas, USA).

4.2.3 Preparation of standard solutions and in solution digestion

Stock solutions were prepared by following the procedures recomanded by Sigma and stored at −20˚C until in solution digestion and peptide mixture analysis.

2μg of each standard protein were lyophylized and subsequently used to prepare the peptide mixture for the realization of the calibration curve by external standard method.

Standard proteins were submitted to reduction, alkylation and tryptic digestion. Samples were dissolved in denaturant buffer (6M Urea, 300

mM Tris pH 8.0, 10 mM EDTA) containing DTT (10-fold molar excess on the Cys residues) at 37 °C for 2 h and then iodoacetamide (IAM) was added to perform carboamidomethylation using an excess of alkylating agent (5-fold molar excess on thiol residues). The mixture was then incubated in the dark at room temperature for 30 minutes. The product was purified by chloroform/methanol/water precipitation. Supernatants were removed, and the pellets were dried. Digestion of proteins mixture was carried out in 10 mM AMBIC using trypsin at a 50:1 protein:enzyme mass ratio. The samples were incubated at 37°C for 16 h and after acidification (10% HCOOH) they were dried.

To eliminate any impurities the samples were suspended in 200 μL of 100 mM AMBIC, filtrated by centrifugal filter units (0.22 μm), and dried in a speed-vacuum concentrator. Peptide mixture was then resuspended in 0.1% HCOOH to a final concentration of 5 pmol/ μL.

Standard solutions for each selected protein were prepared by serial dilution from stock solutions and used for calibration curves. Standard mixtures for each analyte at the following concentrations: 1, 10, 25, 50, 100, 250, 500, 1000 fmol/μL for all the standard proteins were prepared. All standards were kept at -20˚C before LC-MRM/MS analysis.

4.2.4 Sample preparation: In solution digestion of LPS stimulated THP1cellular extracts

A total amount of 1 million of cells for both control THP1 samples and LPS stimulated THP1 cells were used for quantitative analysis of the selected proteins. The cellular extracts were lyophilized and then resuspended in denaturant buffer to reduction, carboamidomethilation and tryptic hydrolysis of proteins as previously described for standard proteins.

After tryptic hydrolysis, purified peptide mixture was resuspended in 50 μL of 0.1% HCOOH and stored at -20°C until the LC-MRM/MS analysis.

4.2.5 LC–MS/MS instrumentation and conditions: MRM targeted proteomic approach

In order to set up a targeted MRM method, Skyline software (3.7, 64-bit version MacCoss Lab Software, University of Washington, USA) was used for the *in silico* selection of peptides with unique sequence for

each selected protein. For each peptide, *m/z* precursor ion, *m/z* product ions and relative collision energy were provided by Skyline.

Peptides with zero missed cleavages were considered and the best two to five transitions per peptide were selected from the top ranked y-fragments. As a result, 9 peptides for the four selected proteins were selected and 29 transitions were monitored during a single analysis.

Table 1 summarized the amino acid sequence for all the selected peptide for each protein and reported the best transitions, *m/z* precursor ion- *m/z* product ions, with the corresponding Collision energy (V) value to optimize the MRM-MS method.

Table 1: List of selected protein reporting selected peptide sequence, precursor m/z, product ions m/z and Collision Energy (CE,V).

Protein	Peptide sequence	Precursor Ion m/z	Product Ions m/z	Collision Energy (V)
TNF-α	R.ANALLANGVELR.D [108, 119]	620,8540++	A[y10]1055,620+ L[y9]-984,5837+ L[y8]-871,4996+ A[y7]-758,4155+ N [y6]-687,3784+	20
	R.IAVSYQTK.V [158, 165]	455,2556++	A[y7]-796,4199+ V[y6]-725,3828+ S[y5]-626,3144+ Y[y4]-539,2824+	15
INF-γ	K.SVETIK.E [91, 96]	338,6974++	V[y5]-589,3556+ E[y4] -490,2871+ T [y3]-361,2445+	12
	K.FFNSNK.K [103, 108]	378,6874++	F[y5]-609,2991+ N[y4]-462,2307+	11
	R.DDFEK.L [112, 116]	327,1425++	D[y4]-538,2508+ F [y3]-423,2238+	11
IL-8	R.ELCLDPK.E [74, 80]	437,7206++	L [y6]-745,3913+ C[y5]-632,3072+ L[y4] -472,2766+	14
	K.ENWVQR.V [81, 86]	416,2090++	N[y5]-702,3682+ W[y4]- 88,3253+	11
IL-10	K.DQLDNLLLK.E [58, 66]	536,3059++	Q[y8]-956,5775+ L[y7]-828,5189+ D[y6]-715,4349+ N[y5]-600,4079+	20
	K.ESLLEDFK.G [67, 74]	490,7504++	S[y7]-851,4509+ L[y6]-764,4189+ L[y5]-651,3348+ E [y4]-538,2508+	18

Peptide mixture after purification was analyzed by LC-MS/MS analysis using a Xevo TQ-S (Waters) equipped with an IonKey UPLC Microflow Source coupled to an UPLC Acquity System (Waters). For each run, 1 µL peptide mixture was injected and separated on a TS3 1.0 mm × 150 mm analytical RP column (Waters, Milford, MA, USA) at 45°C with flow rate of 3 µl/min using 0.1% HCOOH in water (LC-MS grade) as eluent A and 0.1% HCOOH in ACN as eluent B. Peptides were eluted (starting 1 min after injection) with a linear gradient of eluent B in A from 7% to 95% in 55 min. The column was re-equilibrated at initial conditions for 4 min. The MRM mass spectrometric analyses were performed in positive ion mode. using an MRM detection window of 0.5-1.6 min per peptide; the duty cycle was set to automatic and dwell times were minimal 5 ms. Cone voltage was set to 35V.

4.2.6 Method validation Limit of detection and quantitation.

The **limits of detection** (LODs) were defined as the lower limit of concentration below which the sample could not be revealed and were determined by making 10 replicate measurements of blank samples. Calculations were made according to the following formula:

$LOD = 3 \times SD$

where SD was the standard deviation of the Blank.

The **limit of quantitation** (LOQ) is the lowest concentration at which the analyte can not only be reliably detected but also quantitated. The LOQ may be equivalent to the LOD or higher. Standard solutions were prepared by spiking known amount of IL-10, Il-8, INF-γ and TNF-α in the control sample untreated with LPS. The proteins were hydrolysed and analysed by the LC-MRM/MS procedure. LOQ was determined as the analyte concentration giving a signal to noise ratio (S/N) of 3 times higher than the LOD.

Matrix effect. Possible matrix effects were evaluated by comparing standard and matrix matched calibration curves for each analyte. Standard solutions were prepared as described. The calibration curves were repeated three times. Matrix effects were evaluated by comparing five points standard and matrix-matched calibration curves.

Specificity. The specificity of the assay was demonstrated by checking for interfering peaks at the retention time of the target analytes.

Selectivity. Selectivity was evaluated by analyzing 20 blank samples. The blank samples were prepared by executing the whole analysis procedure without test sample and omitting the standards addition. Selectivity was defined acceptable when no peaks were detected in the chromatogram of the procedural blank sample at the retention time of the analytes ± 0.1 minute or, if present, peaks did not exceed 30% of the height of the native analyte in the chromatogram of the lowest calibration level.

Linearity. The linearity of an analytical procedure is defined as its ability to obtain test results that are directly proportional to the concentration (amount) of analyte in the sample within a given range. Linearity was demonstrated directly on the standard solutions by dilution of a standard stock solution as described. Linearity was determined by a series of three injections of standards at different concentration levels. A linear regression equation was applied to the results. The linearity is evaluated graphically, by calibration curves, and by visually inspecting a plot of peak area as a function of analyte concentration.

Recovery. Recovery values were calculated from the measurement of standards spiked in test samples. The average recoveries and their relative standard deviations were calculated from all the results obtained for each standard via response factors, which were determined from calibration solutions as follows:

$$Recovery = \frac{c_1 - c_2}{c_3} \times 100$$

where:

C1: analyte concentration measured after the addition

C2: analyte concentration measured before the addition

C3: added concentration

Recovery values were not used for correction of analysis results. They were only used to monitor the yield of the sample preparation procedures.

4.3 Results

Despite the occurrence of ELISA tests to quantify protein content in cellular extracts or biofluids, the development of an MRM/MS method is proposed as a comparative method, with high selectivity and specificity for identification and quantification of a panel of target proteins in a single analysis trying to overcome the problems of cross reactivity in the use of antibodies.

Cells respond to an inflammatory stimulus primarily by activating signaling pathway which leads to cytokines and chemokines release which act as representatives for the occurring processes.

The MRM method proposed in this work fulfils the request of a targeted approach with multiplexing capability to evaluate the inflammatory response, as a fingerprint of the cells state, at different time points.

On the example of differentiated THP-1 cells stimulated with *E. Coli* LPS that the proposed MRM approach for quantification of TNF-α, INF-γ, IL-8 and IL-10 with high accuracy, sensitivity, and robustness is a powerful tool to determine the inflammation state of cells.

The best ionizing peptides *per* protein were easily determined and the best transitions precursor ion-product ions were indicated with the aid of the Skyline software, matrix interference could clearly be assessed by spiking samples and thus avoided. Therefore, the set-up of this MRM method did not require labeled peptides as references[9] and the quantitative analysis has been performed by the external standard method by using standard proteins.

4.3.1 THP-1 differentiation and LPS stimulation

To differentiate THP-1 monocytes, cells were plated into 6 well plates (5×10^5 cells/well), and then 2 nM of PMA was added to the culture medium for 96 hrs at 37 °C. To induce the inflammatory process, the macrophages were stimulated by adding 100 ng/mL of LPS from *E. coli* for 40 min at 37 °C. Afterwards, we analysed the activation of NF-kB (nuclear factor kappa-light-chain-enhancer of activated B cells), a key factor in regulating the immune response to infection, by following the phosphorylation of its inhibitor IκB. Upon phosphorylation and consequent degradation of IκB, the NF-κB complex is released and enters into the nucleus where it can 'turn on' the expression of specific

genes endowed with DNA-binding sites for NF-κB. The activation of these genes by NF-κB leads to the inflammatory or immune response. NF-κB turns on the expression of its own repressor, IκB. For this reason, IκB levels monitoring has to be performed in a specific time interval, in particular prior to its re-synthetisation (40 minutes at most)[10]. THP-1 cells were stimulated with LPS and, subsequently, western blot analyses have been performed by using antibodies specifically recognizing IκB. a significant decrease (4.4 fold) of intracellular levels of IκB upon treatment of THP-1 cells with *E. Coli* LPS was found, strongly suggesting an activation of the NF-κB pathway (**Figure 3**).
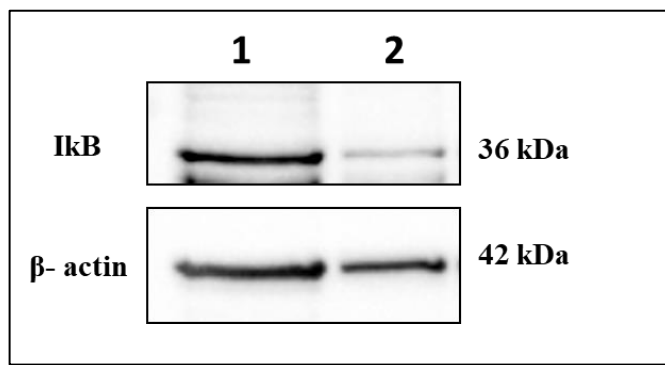
Figure 11: Western-blot analyses of intracellular levels of IκB activated by the treatment of differentiated THP-1 cells with LPS. Lane 1, total proteins extracted from untreated cells; lane 2, total proteins extracted from cells treated with 100 ng/mL of LPS (40 m minutes). Western-blot analyses have been performed by using antibodies specifically recognizing IκB. Endogenous β-actin was used as an internal standard.

Upon induction of inflammation process, in order to evaluate the release of cytokines in the conditioned medium, another marker of inflammation was followed, such as tumour necrosis factor α (TNF-α). TNF-α is a cytokine playing a key role in systemic inflammation and is one of the cytokines involved in the acute phase reaction. It is mainly produced by activated macrophages[11], but it is also produced by a broad variety of cell types including lymphoid cells, mast cells, endothelial cells, cardiac myocytes, adipose tissue, fibroblasts, and neurons. Large amounts of TNF-α are released in response to inflammation induced by cell exposure to lipopolysaccharide, bacterial products, and Interleukin-1 (IL-1)[12].

To evaluate the release of this protein, cells were incubated with 1 µg/mL of LPS from *E. coli* for 6 hrs and, subsequently, the conditioned

medium was analysed by western-blotting by using TNF-α as a marker of inflammation. It has to be noticed that, because of the low sensitivity of western-blotting methodology, no signal associated to TNF-α was observed (**data not shown**). For this reason, upon inflammation induction by treatment of cells with LPS (1 µg/mL) for 3 hrs at 37 °C, protein secretion was inhibited by co-treating cells with LPS (1 µg/mL) and brefeldin A (300 ng/mL) for further 3 hrs at 37 °C. By this way, TNF-α intracellular levels increase because of inflammation induction and concomitant block of protein secretion. In these experimental conditions, a significant increase of TNF-α levels (8.6 fold) has been observed upon inflammation induction (**Figure 4**).
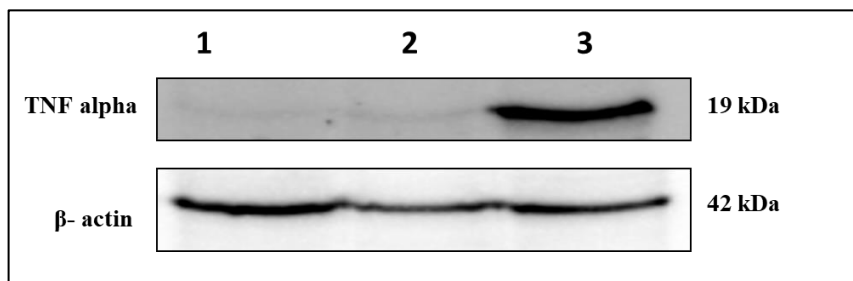
Figure 12: Analysis of TNF-α intracellular levels upon inflammation induction by treatment of differentiated THP-1 cells with LPS. Lane 1, total proteins extracted from untreated cells; lane 2, total proteins extracted from cells treated with 300 ng/mL of brefeldin (3 hrs); lane 3, total proteins extracted from cells co-treated with 300 ng/mL of brefeldin A (3 hrs) and 1 µg/mL of LPS (3 hrs). Western-blotting analyses were performed by using antibodies directed towards TNF-α. Endogenous β-actin was used as an internal standard.

A kinetic analysis by stimulating macrophage cells ($1x10^5$ cells/well) with 1 μg/mL of *E.coli* LPS for different lengths of time (2, 4, 6, 9 and 24 hrs) was performed. This time range allowed the exploration of both "primary marker of inflammation" due to the immediate response of the immune system to the action of LPS, and the "late marker of inflammation" that generally occurs since 10 hours after the stimulation. The supernatants of the different time points were collected and subjected to MRM analysis.

For the MRM analysis, the volume of conditioned medium analysed, has been normalized for the number of the cells ($1*10^5$ cells for each point).

4.3.2 MRM method set up

For each target protein the peptide showing the highest Total Ion Current for all the monitored transitions was assigned as *Quantifier* while the other/s monitored peptides were used as *Qualifiers* to enhance the specificity in protein recognition in a real sample.

Linearity was studied using TNF-α, INF-γ, IL-8, IL-10 standard proteins. Solution at a known concentration (fmol/μL) were analyzed and the calibration curves were realized by plotting the sum of the peak areas for all the monitored transitions for the quantifier peptide versus the concentration of standard protein (fmol/μL) and linear functions have been applied to the calibration curves. The analyses were performed in triplicate for each point of the calibration curves to evaluate the reproducibility of the developed method.

The coefficients of determination (R^2) were greater than 0.99 and the calculated CV values were under the 10% for all the analytes. Figure 6 the standard calibration curves obtained for the four analytes are reported (Figure 6).
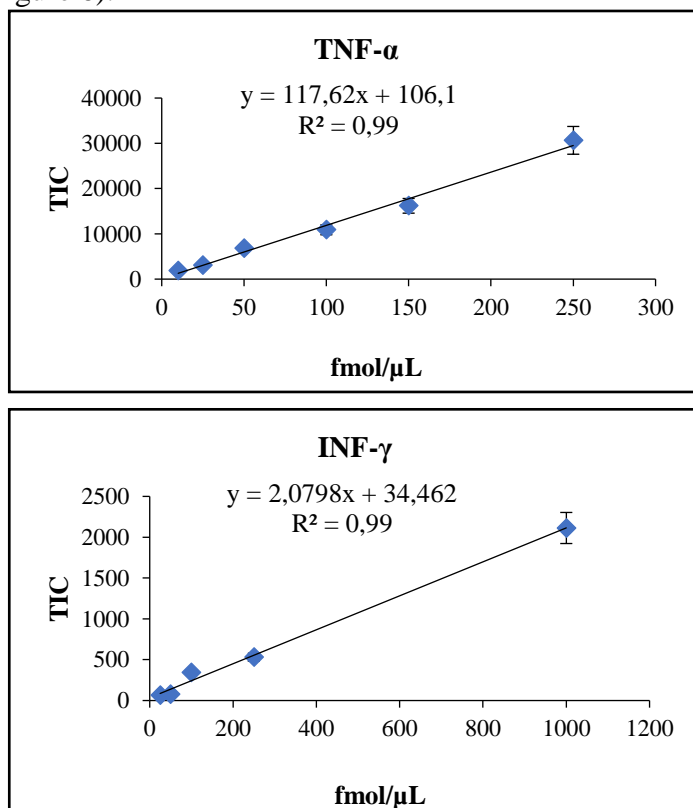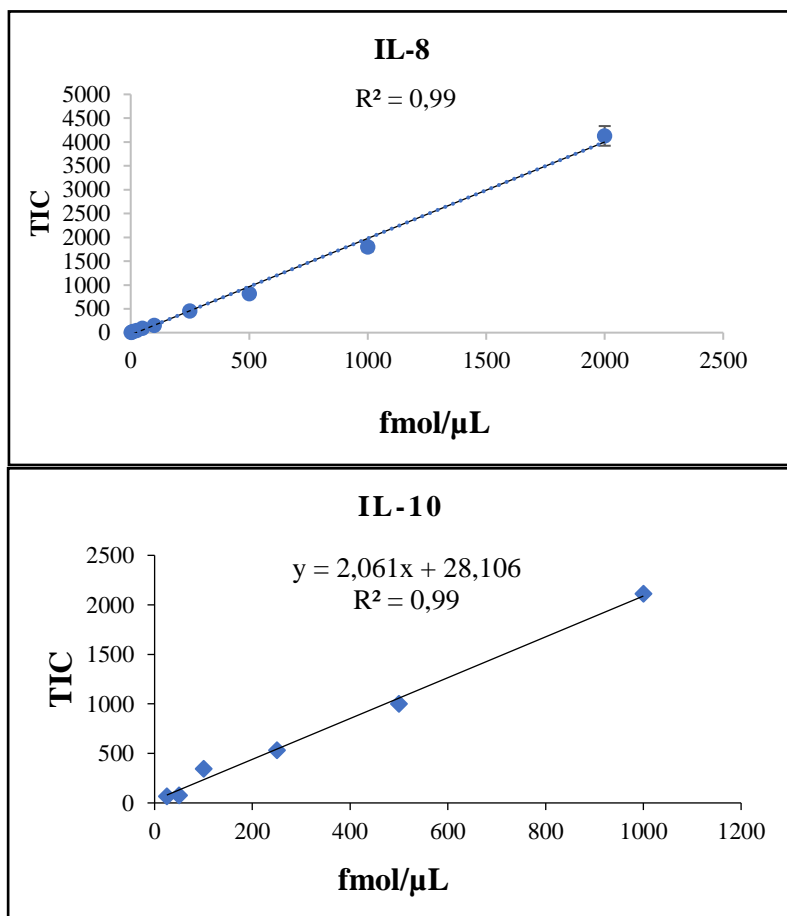
Figure 5: Calibration Curves for TNF-α, INF-γ, IL-8 and IL-10. Total ion current is reported as function of concentration expressed as fmol/µL.

In order to define the recovery in matrix and the LOQ, standard solutions were prepared by spiking known amount of standard proteins in mixture in THP I differentiated and non-stimulated cells samples.

The matrix effect was calculated as the percentage of the matrix matched calibration slope (B) divided by the standard calibration slope (A) yielding a matrix effect (B/A x 100) of 11–15% for the four analytes.

The limits of detection (LODs) were determined by making 10 replicate measurements of blank samples spiked with low concentrations of each analyte and calculated as: $LOD = 3SD_{Blank}$.

The LODs values resulted to be 4,8 fmol/µL for TNF-α, 2,5 fmol/µL for INF-γ, 1,2 fmol/µL IL-8 and 3,2 fmol/µL for IL-10.

The LOQ values for each target protein were 14,4 fmol/µL for TNF-α, 7,5 fmol/µL for INF-γ, 3,6 fmol/µL IL-8 and 9,6 fmol/µL for IL-10.

One million of THP I differentiated cells were used as control samples to define the endogenous concentration of the target proteins in non-stimulated conditions. This control sample was treated as described in Material and Methods section to obtain the tryptic peptide mixture for LC-MRM/MS procedure using the conditions previously defined.

The endogenous background of each analyte was corrected in Excel by calculating the non-spiked signal. LOQ, calculated as the analyte concentration giving a signal to noise ratio (S/N) of 10, resulted to be 40 fmol/µL for TNF-α, 11433 fmol/µL for INF-γ, 3 fmol/µL IL-8 and 760,97 fmol/µL for IL-10.

Recovery was calculated by comparing the extraction yield before and after the addition of the different standard analytes into the control samples. Quantitation was achieved by using the calibration curves. Recovery for each analyte was better than 92,3%.

4.3.3 MRM/MS method applied on LPS stimulated THP-1 cells

The MRM/MS optimized procedure was applied to 1 million THP-1 differentiated cells stimulated with *E. Coli* LPS to induce inflammation. The analyses were performed at different time point after LPS stimulation: 2, 4, 6, 9 and 24 hours.

The MRM analysis were performed in triplicate to evaluate the reproducibility of the method in complex matrixes for each time point to evaluate the reproducibility of the analysis.

The MRM TIC Chromatograms in **Figure 7** showed the capability of the developed method to identify with high specificity and selectivity the ELCLDPK peptide from IL-8 protein without matrix interference (anomalous peaks derived from the matrix effect).
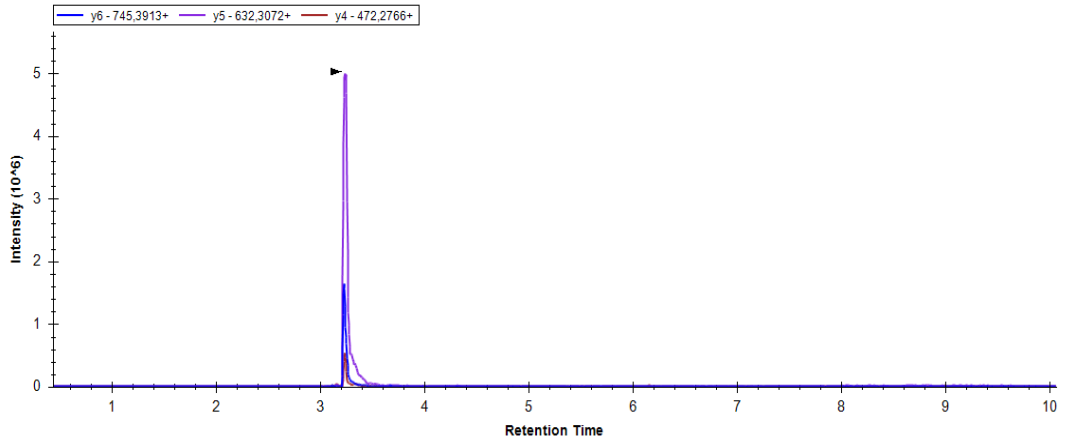
Figure 6: MRM TIC Chromatogram for ELCLDPK peptide from IL-8. Monitored transitions are perfectly coeluted at 3,2 min.

All the results obtained by the MRM analysis of THP I cells differentiated and stimulated with *E.Coli* LPS are summarized in the following histograms (Figure 8). The graphs report the expression trend of each target protein, as average of the three replicates, at different time points after the induction of inflammation with *E.Coli* LPS.
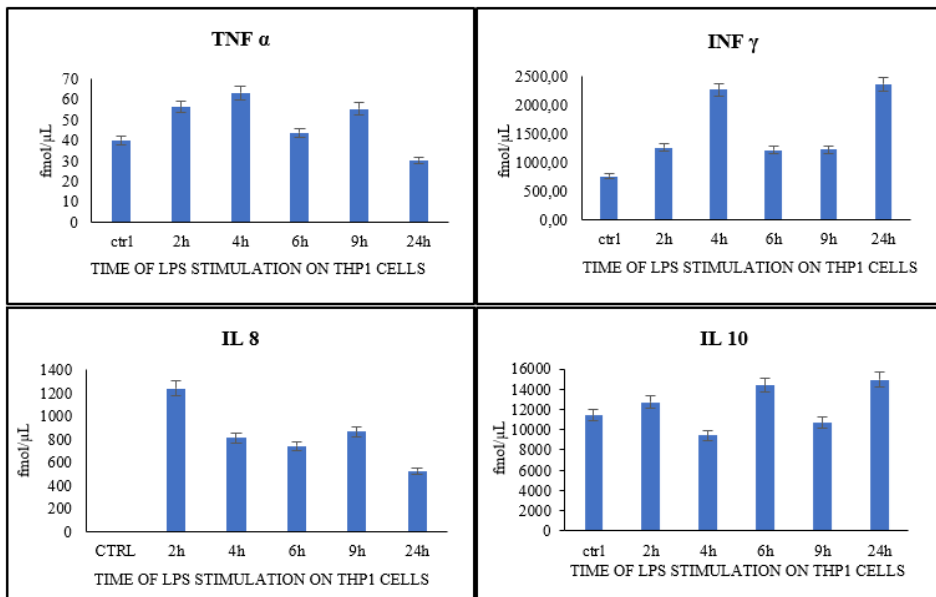
Figure 8: Graphical representation of target protein content in the time-course analysis. Protein content was expressed as fmol/µL for each sample of THP I cell stimulated with *E.Coli* LPS. Each sample was analysed in triplicate and the calculated standard deviation is 8-11%.

Each point of the time-course analysis was related to a sample of THP I cells not stimulated with *E.Coli* LPS (CTRL) in order to indicate the content of each target protein at the baseline level.

The following paragraphs show some information on the production and mechanism of action of selected proteins.

4.3.4 TNF-α

TNF-α is a proinflammatory cytokine which exerts multiple biological effect, it is a cell signalling protein (cytokine) involved in systemic inflammation and is one of the cytokines that make up the acute phase reaction. It is produced chiefly by activated macrophages, although it can be produced by many other cell types such as CD4+ lymphocytes, NK cells, neutrophils; mast cells, eosinophils, and neurons[13].

The primary role of TNF is in the regulation of immune cells. TNF, being an endogenous pyrogen, is able to induce fever, apoptotic cell death, cachexia, inflammation and to inhibit tumorigenesis and viral replication and respond to sepsis via IL1 & IL6 producing cells.

Dysregulation of TNF production has been implicated in a variety of human diseases including Alzheimer's disease, cancer, major depression, psoriasis and inflammatory bowel disease(IBD).

As it is possible to verify from the panel A of Figure 8, the production of TNF-α is increased immediately, after 2 h of inflammatory stimulation with LPS, and its production continues to increase up to 4h of stimulation of THP cells I. This confirming the pro-inflammatory action exerted by TNF-α in the recruitment of macrophages and giving rise to stimulation of complex signalling cascades that control various intracellular functions.

This cytokine acts in synergy with IFN-γ, stimulates the migration of immune cells to the infection site, contributing to the granuloma formation, capable of controlling the disease progression.

4.3.5 INF-γ

Interferons (IFNs) are substances originally identified at cellular culture supernatants infected by virus and that appeared to interfere directly in the viral replication, hence its denomination IFN-γ. It was synthesized mostly by T lymphocytes and NK cells after this cells activation with immune and inflammatory stimuli, rather than viral infection and orchestrates a vast array of pathological responses.[14] It upregulates expression of inflammatory mediators, especially the primary mediator such as TNF-α and interleukin (IL)-1, which induces alveolar macrophages to produce a large number of secondary inflammatory cytokines such as IL-6, IL-8[15].

As reported in Figure 8 (B), INF-γ production is straightly related to TNF- α. In fact, it follows the same expression pathway but delayed in time. The maximum INF- γ levels were recorded after 4h and after 24h of LPS stimulation according with the acute and late inflammation processes.

4.3.6 IL-8

IL-8, is secreted by monocytes/macrophages and epithelial cells. It promotes the accumulation and activation of inflammatory cells, the release of inflammatory mediators and aids in neutrophil chemotaxis to sites of inflammation to regulate the inflammatory response[16].

IL-8, a pro-inflammatory factor, has been shown to promote cancer progression not only by directly stimulating cancer cell proliferation, survival, and migration but also by increasing angiogenesis[17 18]

According to these data, IL-8 secreted levels by THP-1 cells were significantly increased after 2h of *E.Coli* LPS stimulation and these decrease slowly in time until it became about 1/3 after 24h.

4.3.7 IL-10

Due to its ability to inhibit the T lymphocyte production of cytokines, IL-10 was originally described as a cytokine synthesis inhibitory factor (CSIF). IL-10 acts inhibiting the production of pro-inflammatory cytokines (IFN-γ, TNF-α and IL-12) and the action of

antigen presenting cells, blocking the activation of T lymphocytes through the inhibition of expression of MHC class II molecules.

Unlike TNF-α and IFN-γ, IL-10 is considered primarily an inhibitory cytokine, important to the adequate balance between inflammatory and immunopathological responses[6,19].

Evidences of this behaviour are reported in panel D of Figure 8: the main increase in IL10 level by LPS is recorded after 24 h of LPS stimulation. Otherwise, the concentration of this protein remains almost constant or equal to the THP I non-stimulated sample.

To investigate the sensitivity of TNF alpha detection by using antibody, a pool of serum samples from patients affected by different inflammatory diseases was submitted to TNF-α test by using specific antibody as previously described. Figure 8 showed absence of positive signal of the target protein in serum sample (lane 3). Thus, different aliquot of TNF-α standard protein at known concentration were loaded in order to define the limits of the range of detection by using this methodology.

As shown in this figure, the limit od detection is around 6,25 ng corresponding to about 35 fmol/µL. Thus, taking in account that the limit of detection of TNF-α by MRM method was set up at 4,8 fmol/µl, this method seems to be promising for the detection and quantification of these inflammatory proteins in sera.
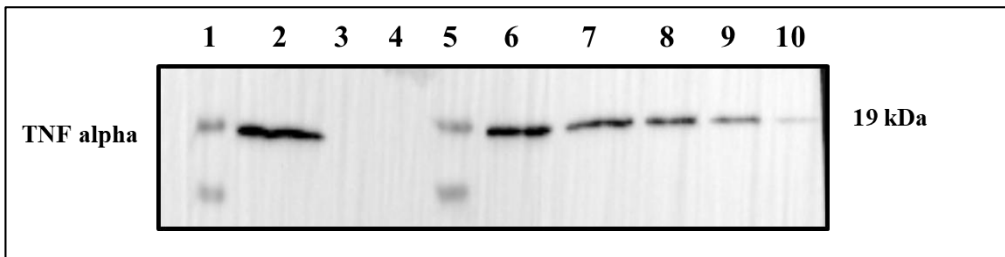
Figure 13: Analysis of TNF-α levels in the serum of patients with inflammatory diseases. Lane 1, pre-stained protein markers; lane 2, pure TNF-α protein (200 ng); lane 3, control serum; lane 4, serum from patients with inflammatory disease; lane 5, pre-stained protein markers; lane 6, pure TNF-α protein (100 ng); lane 7, pure TNF-α protein (50 ng); lane 8, pure TNF-α protein (25 ng); lane 9, pure TNF-α protein (12,5 ng); lane 10, pure TNF-α protein (6,25 ng). Western-blotting analyses were performed by using antibodies directed towards TNF-α protein

4.4 Conclusions

SRM and MRM strategies appear as an excellent match to the requirements of systems biology, not only allowing quantitative analysis of low-abundant proteins, but also delivering reliably quantitative data when proteins are analysed across multiple samples[20].

Cytokines are molecules that mediate mainly the intercellular communication in the immune system, being produced by different cell types. Cytokines have pleiotropic and regulatory effects and participate in the host defence and in inflammatory and tissue reparation processes.

The aim of this work was the development of a mass spectrometry method in MRM mode that allowed the simultaneous monitoring of four proteins TNF-α, INF-γ, IL8 and IL10. The developed method was used for the study of the expression trend of these proteins in THP I differentiated and stimulated samples to induce inflammation with E. coli LPS.

The method shows interesting characteristics: high specificity and selectivity, but also sensitivity comparable to immunoenzymatic assays, is a robust method with good reproducibility.

The differentiated THP cells, were treated with a well-known endotoxin to induce the inflammatory process, and in order to study how the system reacts to stimulation, a time-course analysis has been performed. Samples were collected at 2h, 4h, 6h, 9h, 24h after stimulation with LPS. Each sample was then analysed in triplicate to evaluate the reproducibility of the instrumental response.

The results obtained from the MRM analysis allowed us to draw an expression profile for each of the target proteins that goes well with the data reported in the literature.

The immune system's response to LPS is an increase in TNF-α production that triggers a cascade process mechanism that leads to the production of INF-γ and IL-8. Il-10 remains almost constant throughout the process and around 24 hours after stimulation with LPS there is an increase in its expression, supporting its action of anti-inflammatory cytokine.

4.5 References

(1)   Balkwill, F.; Burke, F. The cytokine network. *Immunology today* **1989**, *10*, 299-304.

(2)   Knobloch, J.; Chikosi, S.-J.; Yanik, S.; Rupp, J.; Jungck, D.; Koch, A. A systemic defect in Toll-like receptor 4 signaling increases lipopolysaccharide-induced suppression of IL-2-dependent T-cell proliferation in COPD. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **2015**, *310*, L24-L39.

(3)   Molteni, M.; Gemma, S.; Rossetti, C. The role of toll-like receptor 4 in infectious and noninfectious inflammation. *Mediators of Inflammation* **2016**, *2016*.

(4)   MERAZ RIOS, M. A.; Toral-Rios, D.; Franco-Bocanegra, D.; Villeda-Hernández, J.; Campos-Peña, V. Inflammatory process in Alzheimer's Disease. *Frontiers in integrative neuroscience* **2013**, *7*, 59.

(5)   Blais, D. R.; Vascotto, S. G.; Griffith, M.; Altosaar, I. LBP and CD14 secreted in tears by the lacrimal glands modulate the LPS response of corneal epithelial cells. *Investigative ophthalmology & visual science* **2005**, *46*, 4235-4244.

(6)   Zhang, J.; Hu, D.-K.; Wang, D.-G.; Liu, Y.; Liu, C.-B.; Yu, L.-H.; Qu, Y.; Luo, X.-H.; Yang, J.-H.; Yu, J. Effects of clinical isolates of Streptococcus pneumoniae on THP-1 human monocytic cells. *Molecular medicine reports* **2013**, *8*, 1570-1574.

(7)   Szulc-Kielbik, I.; Pawelczyk, J.; Kielbik, M.; Kremer, L.; Dziadek, J.; Klink, M. Severe inhibition of lipooligosaccharide synthesis induces TLR2-dependent elimination of Mycobacterium marinum from THP1-derived macrophages. *Microbial cell factories* **2017**, *16*, 217.

(8)   Liebler, D. C.; Zimmerman, L. J. Targeted quantitation of proteins by mass spectrometry. *Biochemistry* **2013**, *52*, 3797-3806.

(9)   Surinova, S.; Schiess, R.; Hüttenhain, R.; Cerciello, F.; Wollscheid, B.; Aebersold, R. On the development of plasma protein biomarkers. *Journal of proteome research* **2010**, *10*, 5-16.

(10)   Deptala, A.; Bedner, E.; Gorczyca, W.; Darzynkiewicz, Z. Activation of nuclear factor kappa B (NF-κB) assayed by laser scanning cytometry (LSC). *Cytometry: The Journal of the International Society for Analytical Cytology* **1998**, *33*, 376-382.

(11)   Olszewski, M. B.; Groot, A. J.; Dastych, J.; Knol, E. F. TNF trafficking to human mast cell granules: mature chain-dependent endocytosis. *The Journal of Immunology* **2007**, *178*, 5701-5709.

(12)   Walsh, L. J.; Trinchieri, G.; Waldorf, H. A.; Whitaker, D.; Murphy, G. F. Human dermal mast cells contain and release tumor necrosis factor alpha, which induces endothelial leukocyte adhesion molecule 1. *Proceedings of the National Academy of Sciences* **1991**, *88*, 4220-4224.

(13)   Jameson, J. L.; De Groot, L. J.: *Endocrinology-E-Book: Adult and Pediatric*; Elsevier Health Sciences, 2010.

(14)   Ana P. Ferreira-Duarte, A. S. P.-T., Gabriel F. Anhê, Antônio Condino-Neto, Edson Antunes   and Ivani A. DeSouza. MHC Class II Activation and Interferon-γ Mediate the Inhibition of Neutrophils and Eosinophils by Staphylococcal Enterotoxin Type A (SEA). *Front Cell Infect Microbiol* **2017**, *7: 518.*

(15)   Wang, P.; Han, X.; Mo, B.; Huang, G.; Wang, C. LPS enhances TLR4 expression and IFN-γ production via the TLR4/IRAK/NF-κB signaling pathway in rat pulmonary arterial smooth muscle cells. *Molecular medicine reports* **2017**, *16*, 3111-3116.

(16)   Baggiolini, M.; Walz, A.; Kunkel, S. Neutrophil-activating peptide-1/interleukin 8, a novel cytokine that activates neutrophils. *The Journal of clinical investigation* **1989**, *84*, 1045-1049.

(17)   Gabellini, C.; Trisciuoglio, D.; Desideri, M.; Candiloro, A.; Ragazzoni, Y.; Orlandi, A.; Zupi, G.; Del Bufalo, D. Functional activity of CXCL8 receptors, CXCR1 and CXCR2, on human malignant melanoma progression. *European journal of cancer* **2009**, *45*, 2618-2627.

(18)   Waugh, D. J.; Wilson, C. The interleukin-8 pathway in cancer. *Clinical cancer research* **2008**, *14*, 6735-6741.

(19)   Cavalcanti, Y. V. N.; Brelaz, M. C. A.; Neves, J. K. d. A. L.; Ferraz, J. C.; Pereira, V. R. A. Role of TNF-alpha, IFN-gamma, and IL-10 in the development of pulmonary tuberculosis. *Pulmonary medicine* **2012**, *2012*.

(20)   Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Molecular systems biology* **2008**, *4*, 222.

Chapter 5

5.1 Introduction

5.1.1 Rheumatoid Arthritis: etiology, gender distribution and social impact.

Rheumatoid arthritis (RA) is a polygenic and multifactorial syndrome, believed to be caused by a combination of genetic[1,2], and environmental factors. Once triggered, the pathological process has the synovial joints as the main target, in which both inflammatory (acute and chronic) and destructive phenomena are identifiable (Figure 1).
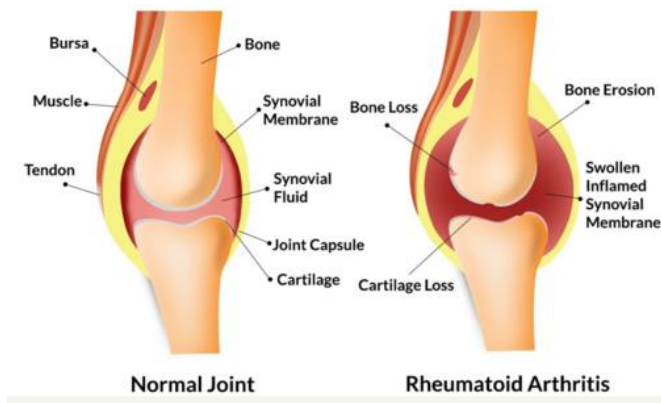
Figure 14: Comparison between a normal joint and one suffering from Rheumatoid Arthritis.

Many complex immunological and genetic interactions are involved in the outcome of the clinical disease. In the articular joints RA involves a multicellular inflammatory process; infiltration of lymphocytes and granulocytes into the articular cartilage, proliferation of synovial cells, leukocyte extravasation, and, neo-vascularization of the synovial lining surrounding the joints[3].

This proliferative process not only induces swelling, erythema, and pain of multiple joints, but also progresses to the destruction and loss of cartilage and bone architecture. Many cellular components (macrophages, dendritic cells, synovial cells, mast cells, neutrophils, T cells, and B cells), cell surface molecules (co-receptors, adhesion molecules, and integrins), signalling components (ZAP70, PTPN22, JAK, MAPK and Stat1), metabolic components, and humoral mediators (antibodies, cytokines, chemokines, metalloproteinases, serine proteases, and aggrecans) interact and aid in the

disease progression, leading to the digestion of extracellular matrix and the destruction of articular structures[4].

RA has a significant impact on the lives of those affected: about 90% of the subjects have some degree of disability within 20 years from the onset. In 22% of cases the patient is forced to abandon all types of work and 10% need continuous assistance. It is not only the quality of life that is compromised but also the life expectancy decreases on average 4 years in men and 10 in women if the disease is not adequately treated. The prevalence of RA is higher in females than males, the incidence is 4–5 times higher below the age of 50, but above 60–70 years the female/male ratio is only about 2[5]. The prognosis is worse in patients suffering from severe illness or presenting extra-articular symptoms. In fact, the social costs increase exponentially with the aggravation of the affliction. The control of the disease, therefore, preserves from the increase of the financial burden. Given the chronic and potentially disabling nature of the disease, it is essential to consider RA not only from a medical but also a social point of view.

It has been estimated that in Italy RA affects about 0.5% of the general population: 0.6% of women and 0.25% of men in adulthood. Although the prevalence of RA is significantly lower than other conditions, such as osteoarthritis, the frequent severity of the clinical picture and the high potential for disability make it a disease with a significant socio-economic impact in terms of costs, disability and loss of productivity[6].

5.1.2 Rheumatoid Arthritis diagnosis

The RA diagnosis arises through a specialized assessment of symptoms and clinical signs. These signs have a wide range of severity degrees, and there are several inflammatory and arthritic conditions that manifest similar symptoms. This makes RA diagnosis very difficult.

Therefore, the patient must undergo numerous medical, laboratory and radiographic tests.

Physical exam. The detailed chronology of the beginning and severity of symptoms as well as the family history of RA or other autoimmune disorders. It investigates the affected swollen, inflamed and painful joints and other joints, including the hands, upper limbs, knees and feet.

Health assessment questionnaires. The HAQ was developed as a comprehensive measure of outcome in patients with a wide variety of rheumatic diseases, including RA, osteoarthritis, juvenile RA, lupus,

scleroderma, ankylosing spondylitis, fibromyalgia, and psoriatic arthritis. It has also been applied to patients with HIV/AIDS and in studies of normal aging. It should be considered a generic rather than a disease-specific instrument. Its focus is on self-reported patient-oriented outcome measures, rather than process measures[7].

Laboratory tests. Laboratory tests can help to differentiate this pathology from other diseases and disorders that impair joint functioning. Several blood tests are necessary: *Erythrocyte Sedimentation Rate* (ESR), *C-Reactive protein* (CRP)[8] and *Rheumatoid Factor*.

Radiographic analysis. Several imaging studies that examine the affected joints may help in the diagnosis of RA. The presence of two or more eroded and destroyed joints is significant in rheumatoid arthritis. Radiography can be very useful in diagnosis but does not show any abnormality in the early stages of the disease (3-6 months). On the contrary, the use of articular echography is more and more relevant to check for decreased space within the joint's bones and damage to the cartilage and soft tissues, much more sensitive than traditional radiography (especially in the initial phase) and cheaper than Magnetic Resonance Imaging, MRI, in documenting the synovial membrane hypertrophy and the intensity of the joint and peri-articular inflammation.

DAS 28 test. The Disease activity score 28 (DAS28) combines 28 inflamed joints and 28 holds (hands, arms and knees) and the patient's perception of his inability. It also includes blood tests such as erythrocyte sedimentation rate. DAS28 has a complicated mathematical evaluation and now a simplified version is used combining 28 tender and swollen joint assessments of doctors' and 'counts' patients, and C-Reactive protein levels.

5.1.3 Serum biomarkers and RA diagnosis

A biomarker is a characteristic that can be objectively measured as an indicator of normal or pathologic biological processes, or as an indicator of response to therapy[9]. A biomarker is a diagnostic indicator of a clinical endpoint and often is a reporter of early-stage disease development. Biomarker selection is likely to be most beneficial when focused on early-stage mechanisms and changes in the disease process.

Biomarkers have the potential to aid clinical diagnosis when symptoms are present or to provide a means of detecting early signs of disease when they are not. Some biomarkers can serve as early surrogates of eventual clinical outcomes or guide therapeutic decision making by enabling identification of

individuals likely to respond to a specific therapy. Using biomarkers might reduce the costs of drug development by enabling individuals most likely to respond to be enrolled in clinical trials, thereby minimizing the number of participants required[10].

In early RA stages a biomarker may help the diagnosis particularly during acute flare-ups where arthritic joints become hot swollen and tender. Radiographic image features correlate very poorly with symptoms and a biomarker test may assist with evaluation and selection of patients for join replacement surgery. The biomarker response may also be of benefit in assessment of post implantation responses. Persistent pain, functional impairment, infection and implant loosening are a few of the common effects of surgery, and a biomarker may be of use in determining the physiological aspects of these problems as opposed to purely mechanical ones[11]. In the diagnosis of RA, rheumatoid factor values have been historically used to support the medical examination and radiographic data.

Rheumatoid factor

Rheumatoid factors (RFs), a class of immunoglobulins (Igs) that have different isotypes and affinities, were first detected more than 70 years ago, but there is still much to discover about the mechanisms underlying their production, physiological role, and pathological effects. In 1948, Rose described these antibodies in patients with rheumatoid arthritis (RA), and in 1952 they were finally christened RFs because of their association with RA.

Rheumatoid factor (RF) is an antibody against the Fc portion of IgG, but in clinical practice is measured as IgM. The role of RF in disease is not clear but may be related to the binding by RF on B-cells to immunoglobulins attached to antigens for antigen presentation, resulting in amplification of the humoral response. The sensitivity of RF for rheumatoid arthritis varies, depending on the patient population. RF is associated with more severe disease with greater extra-articular features. In general, the sensitivity of RF for RA is around 50 to 85%, increasing over time, as some people with rheumatoid arthritis are initially RF negative and later seroconvert. However, RF alone cannot make a diagnosis of rheumatoid arthritis, as roughly 5% of a young healthy population can be RF positive, and this rate increases with age. Although mostly associated with rheumatoid arthritis, RF can also be seen in Sjögren's syndrome, mixed connective tissue disease, mixed cryoglobulinemia, SLE, and polymyositis, as well as in non-rheumatic disorders such as chronic infections, inflammatory disorders, and malignancy. Only 50% of early rheumatoid demonstrate RF positivity.

Furthermore, around 15% of patients with RA will never have RF positivity. Therefore, this test should only be ordered where the diagnosis of RA is strongly suspected: namely, in the setting of obvious joint swelling, synovitis, or effusions. Another setting would be the finding of erosive changes radiologically[12]. RFs can also interfere with other laboratory tests, including those designed to detect anticardiolipin antibodies (especially if IgM levels are in the low positive range), anti-2GPI antibodies, anti-HCV antibodies, antirubella antibodies, thyroid assays, and tests for carbohydrate antigen 19–9 and various cytokines[13].

Anti-CCP antibodies

Anti-cyclic citrullinated peptide (anti-CCP) antibodies have been around for decades but have recently been gaining popularity. These were first described as markers for RA in 1964. It was because of very exacting technical requirements that anti-CCP antibodies never gained wide-spread acceptance, despite their specificity for RA. Some research suggests that citrullinated extracellular fibrin in synovium is a significant autoantigen involved in the immune response in rheumatoid arthritis. It has similar sensitivity for RA (50 to 85%) but higher specificity (90 to- 95%)[14-17]

Up to two-thirds of patients with RA harbor ACPA, as has been known for many years[18]. However, the underlying mechanism linking ACPA with a more severe disease progression with increased joint destruction is incompletely elucidated. Recent data suggest that ACPA influences bone resorption by directly activating osteoclasts[19]. Serological status has become an important factor in diagnosis and prognosis of the disease. Patients with seropositive RA share certain genetic and environmental risk factors and have been shown to have a more severe disease course. Less is known about seronegative RA, but evidence of genetic associations for ACPA-negative RA has been found[20]. The clinical presentation and disease course of seronegative RA are usually reported as less severe than for seropositive RA, although studies are conrtoversial[21,22].

5.1.4 Aim of the study

Blood tests would offer a powerful, minimally invasive method for early diagnosis of RA. However, no reliable biomarkers for RA are presently available. The aim is to develop a multiple reaction monitoring (MRM)-based quantification method to select a panel of candidate biomarkers.

5.2 Materias and Methods

5.2.1 Samples

80 serum samples of patient with RA were provided from Complejo Hospitalario Universitario de Santiago de Compostela (CHUS), Spain. Age, sex, ACPA and RF status were known for all patients. According to the RF and ACPA values these samples were divided into four distinct sub-groups of 20 sera samples each: RF+/ACPA+, RF+/ACPA-, RF-/ACPA+ and RF-/ACPA-.

5.2.2 Materials

13C/15N-labeled lysine and arginine, formic acid (FA), methanol (MeOH), acetonitrile (ACN), were purchased from Thermo Fisher (Waltham, USA). Urea, thiourea, dithiothreitol (DTT), iodoacetamide (IAM), sodium ammonium bicarbonate (AMBIC), trypsin, trifluoroacetic acid (TFA), water (H2O), Stage Tip C18 47mm, Empore were purchased from Sigma-Aldrich (Saint Louis, USA).

5.2.3 Selection of target peptides

Previous studies on this and other RA cohorts, based on 8plex iTRAQ and LC-MALDI/TOF/TOF analysis, showed a statistical relevant modulation (p-value<0.05) of 10 proteins. According to LC-MALDI/TOF/TOF identification and Skyline selection, 26 proteotypic peptides were selected to be appropriate for use in MRM-based assays.

Quantitative analysis was performed by internal standard method: 26 synthetic isotopically labelled peptides were purchased from Thermo Fisher (Waltham, USA).

5.2.4 Design of multiple reaction monitoring (MRM) methods

Skyline software v1.3 (MacCoss Lab Software, Seattle,WA, USA)[23] was used to set up and optimize the best peptides for the target proteins with several filters. Proteotypic peptides with the highest spectral counts, only fully tryptic peptides, with no missed cleavages, with a length between 8 and 30 amino acids and devoid of methionine and cysteine residues, if possible, were chosen for MRM assays development. In addition, sequences that may cause incomplete digestion, such as continuous sequences of arginine (R) or lysine

(K) and a proline (P) at the C-terminal side of R or K were also excluded since partial tryptic hydrolysis at the peptide bond is often observed in MS/MS.

The top transitions were selected for method development based on the presence of abundant y ions at m/z greater than that of the precursor. In the absence of high-m/z y ions, the most abundant fragment b ions were selected.

All the monitored peptides, transitions for the light and heavy peptides and collision energy that were included in the MRM method developed herein, are detailed in Supporting Data Table 1.

5.2.5 Sample Processing

100 µg of protein content were treated with 6 M urea/2M Thiourea.

2 µL of diluted serum was denaturized by adding 1,25 µL of 50mM DTT and 3 µL of 25 mM AMBIC for 1h at 37 °C. The alkylation step was then performed by adding 6,5 µL of IAM solution (100 mM) and further incubated at room temperature for 45 min in the dark.

The samples were diluted with 10 µL of 25 mM AMBIC prior to the addition of 0,5 µL of trypsin solution (1 µg/ µL) for overnight digestion at 37 °C. The amount of trypsin in the digest was calculated for a 1:20 enzyme to substrate ratio. Digests were then acidified with 10 µL of 2% TFA and combined with 10 µL of solutions containing varying SIS peptide concentrations (250, 100, 50, 25 fmol/ µL), depending on the ionization capability of every synthetic peptide.

The sample clean-up was performed by using homemade C18 StageTip. After filter activation (50 µL x2 times MeOH, 50 µL x2 times of 80% ACN/0.5% TFA and 50 x 3 0,5% TFA), the sample was loaded two times on the filter and centrifuged for 2 min at 4000 rpm. Salts were eluted with 50 µL 0,5% TFA and 50 x2 0,5% formic acid. The elution was performed with 30 µL x2 80% ACN/0,5% FA.

Samples were dried under vacuum and rehydrated in 20 µL of 98% H_2O/2% ACN/ 0.1% FA prior to injection, to give a final serum protein concentration of 0,5 µg/µL.

5.2.6 Calibration Curves and SIS Peptide Mixtures

In order to determine in which concentration range the instrumental response is linear for all the selected peptides and which is the best concentration of heavy peptides to be used in the serum samples, an equimolar mixture of all 26 peptides was prepared.

An 11-point dilution curve (from 0.1 to 12000 fmol/μL) containing all the heavy peptides in FA was prepared. These stock solutions were diluted 1:10 and spiked into 10 μg of a pool of digested human plasma protein to prepare the reverse calibration-curve in the range of concentrations from 0.01 to 1200 fmol/μL. The endogenous amount of target peptides was detected by injecting three times the only digested pool of serum (Blank). Each concentration level was processed in technical triplicate.

All the reverse calibration curves, reporting the peak area ratio heavy/light as function of the heavy concentration (fmol/μL) are reported in Supporting Data Figure 1.

Limit of detection (LOD) was calculated as three times the standard deviation of the average area of the blank sample and the limit of quantification (LOQ) was calculated as 10 times the standard deviation value.

5.2.7 LC-MS/MS Methods

Digested samples (2 μL of injection volume) were separated by reversed phase on a nano LC Eksigent system equipped with a Trapping column TRAP Acclaim pepmap 100A nanoviper C18 100μm x 2cm x 5μm and a column Acclaim pepmap 100A C18 75μm x 15cm x 5μm.

Mobile phases A and B consisted of 98% H_2O, 2% ACN, 0.1% FA and 5% H_2O, 95% ACN, 0.1% FA respectively. The flow rate was 300 nL/min throughout the following 70 min multistep gradient: 0 min: 5%B, 5 min: 15%B, 45 min: 35%B, 46 min: 95%B, 56 min: 95%B, 57 min, 5%B, 70 min: 5%B. The nano LC Eksigent system was interfaced to an AB Sciex 5500 QTRAP mass spectrometer via a nanospray ESI source. The capillary voltage was set to 2600 V and the temperature was fixed to 150°C, the Curtain Gas (Cur) was 30 Bar and Ion Source Gas (GSI) was set to 25 l/min. The EP and CXP potentials were set to 10 and 15 V respectively and both Q1 and Q3 were set to unit resolution.

Data were acquired in the positive dynamic MRM mode, using a cycle time of 900 ms for a minimum dwell time of 15 ms. The equivalent three optimized transitions were used to monitor the isotopologue of each peptide: light (endogenous) peptide.

5.2.8 Data Processing

The .*raw* data was processed, and the integration was performed with Skyline software 25 version 4.1.0.11796. Quantitation was performed via

regression analysis of peptide standard curves (1/x weighting), realized for all transitions that were found to be interference free and detectable across the entire concentration range. Statistical analysis was performed using GraphPad Prism 7 version 7.04. Distributed data were analysed using the non-parametric Mann–Whitney U test and statistical significance was set at a p value below 0.05.

Kruskal–Wallis and Mann–Whitney U test and the box plot graphs were performed using GraphPad Prism 5.01 software (La Jolla, CA, USA).

5.3 *Results*

Rheumatoid arthritis is a chronic disease that causes pain, swelling and joint stiffness, limiting the range of movement and function of the affected joints. Although the joint is the part of the organism most involved, inflammation can also develop in internal organs (such as lungs, kidneys, heart, nervous system, blood vessels, eyes).

Although they owe their name to their first detection in RA patients, RFs are found in patients with other autoimmune and nonautoimmune diseases, as well as in healthy subjects. It has been demonstrated that low-affinity RFs appear to be key player in immune responses to many infectious organisms, and high-affinity RFs indicate more severe and persistent disease in patients with RA. RFs are probably the result of the immune response to inflammation (depending on genetic background) and may have regulatory effects on Ig production by controlling B cell activation[13]. Both rheumatoid factor (RF) and anti-cyclic citrullinated peptide antibodies (ACPA) are associated with poor radiologic outcomes in patients with rheumatoid arthritis (RA). general, RA patients positive for RF or ACPA (SPRA) are considered to manifest an aggressive disease course compared with seronegative RA patients (SNRA). However, the relationship between seropositivity and measures of disease severity other than radiologic outcome is disputed. In addition, it plays an important role in the pathogenesis of RA and is significantly associated with radiographic progression [24]. In order to obtain a panel of proteins that can be used, in the common clinical routine, to evaluate both the course of the disease and the effectiveness and the response of the body to pharmacological treatments, it is important to have an optimized method for the patients follow up.

The mayor advantage of the MRM approach is the ability to monitor and quantify, at the same time, all the target peptides with high selectivity and specificity. In contrast, the immunoenzymatic tests (as ELISAs), are more time consuming and do not allow the chance to multiplex the analysis of all the selected proteins.

5.3.1 *Assay development*

Quantitative analysis was conducted by Internal Standard Method, by using as internal standards 26 target peptides with an isotopically labelled Arginine o Lysine. ($\Delta m = +10$ Da, $\Delta m = +8$ Da).

First, assay concentration ranges were established by running dilution curves of the SIS peptides in a human plasma digest. The range of concentration used for the calibration curves was 0.5-1200 fmol/µL. The analyses were performed in triplicate and the results are summarized in Table 1 in which the calculated LOD, LOQ are also reported.

Table 1: Reverse curve parameters for all selected peptides used to set up the MRM/MS method.

Labelled Peptide	LOD [fmol/µL]	LOQ [fmol/µL]	Curve parameters	
			Slope (x10+3)	Intercept (x10+3)
ALG	5,44±0,45	10,68±0,77	7,48E+00	-2,86E+00
TLD	6,50±0,66	7,25±0,47	4,85E+03	6,05E+04
DLL	1,24±0,89	3,92±0,41	2,06E+02	3,31E+01
ATF	135,03±10,2	150,50±9,83	8,64E+01	-6,05E+04
IPS	10,40±0,99	12,01±1,1	1,35E+03	1,65E+01
TSS	17,29± 1,88	18,06±1,92	1,85E+02	-2,42E+05
GVT	26,91±2,33	43,21±3,66	3,82E+01	3,79E-01
LED	11,55±1,15	11,89±1,02	7,61E+03	-7,92E+06
WEL	0,60±0,05	3,33±0,14	1,28E+01	-2,51E-02
VIP	6,71±0,31	7,07±0,43	8,07E+02	2,20E+06
EAQ	1,48±0,09	5,31±0,05	8,24E+01	3,63E-02
YWG	7,42±0,66	8,75±0,65	4,81E+03	-5,50E+07
LLN	1,08±0,09	19,61±1,8	1,48E+00	-2,29E-03
YTF	4,52±0,45	8,82±0,99	2,84E+05	2,10E+08
HLS	16,11±1,78	17,82±1,44	5,10E-01	-9,52E-03
VLE	3,25±0,33	3,59±0,49	4,92E+04	4,88E+06
YVG	7,20±0,98	9,05±1,01	7,95E-01	-3,56E-02
EQL	8,78±0,56	10,28±0,78	1,85E+05	2,81E+07
SDVV	0,76±0,05	15,42±1,45	8,22E-02	7,09E-03
EHV	0,21±0,001	2,56±0,07	1,31E+06	5,49E+07
EQL	12,10±1,11	20,05±1,98	1,03E-02	-1,16E-03
SDVM	2,55±0,54	4,34±0,37	7,02E+06	-8,59E+07
EIG	0,29±0,06	0,89±0,05	6,16E-02	-2,33E-02
AVL	37,57±2,66	65,71±5,23	8,25E+04	2,87E+06
VGY	11,89±1,44	17,53±1,77	1,14E+00	-2,31E-01
VTS	5,01±0,61	5,25±0,48	1,58E+05	3,40E+06

Linearity of results for at least 4 orders of magnitude, reproducibility of analyses and low values of CV (<20%) for the standard heavy peptides spiked in human serum were observed.

To minimize the matrix effect that can negatively affect the quantification of the target proteins, a known amount of SIS peptides was added to the samples following enzymatic digestion (Supporting Data Table 2), but before

the desalting phase. In this way, endogenous and heavy peptides are subject to the same purification treatment and therefore to the same loss.

5.3.2 Analysis of RA serum samples

As described in Material and Methods, 80 sera samples were analysed. These samples were divided into 4 subgroups related to the RF and ACPA tests value as follows: RF+/ACPA+, RF+/ACPA-, RF-/ACPA+ and RF-/ACPA-.

The digested samples were analysed in duplicate by using the MRM developed method and .raw data were analysed with Skyline software to evaluate the perfect coelution of the monitored transitions for endogenous and heavy peptides and to check the peak area integration as reported in Figure 2. Two technical replicates were measured, leading to peak area ratios between the endogenous and heavy peptide (SIS) with good coefficients of variation (% CVs < 20).
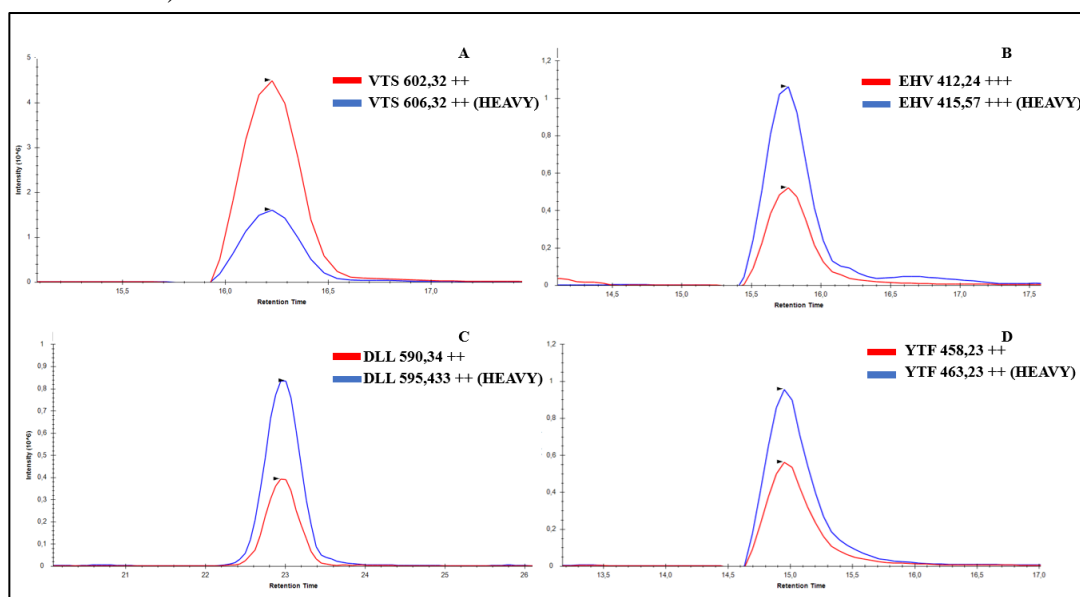
Figure 15: Target peptides VTS (**A**), EHV (**B**), DLL (**C**) and YTF (**D**). Coelution of Light (red) and heavy (blue) transitions at 16,3, 15,7, 23,0 and 15,0 min respectively.

Quantitative analyses on real sera samples were conducted using SIS peptides in the single point calibration method. Single-point calibration consists of calculating the ratio between a known amount of SIS peptide and the analyte. The analyte/SIS peptide ratio is then multiplied by the concentration of the SIS peptide spiked in the sample. The molecular weight of the protein was used to calculate protein concentrations, as previously

94

described[25]. This calibration method assumes that the calibration curve in the range that includes the concentration of the SIS peptide and the analyte is linear, that the slope is 1, and that the intercept goes through the origin as previously observed in the spiked sera[26].

The obtained results for the 80 sera in the following tables are shown (Table 2A-D). The protein concentration is expressed as µg/mL showing values in good agreement with published data[26,27].

Table 2 A-D: Quantitative data for the four sub-groups RF-/ACPA- (**A**), RF+/ACPA+ (**B**), RF+/ACPA- (**C**) and RF-/ACPA+ (**D**). Standard deviation of approximately 10% was recorded.

		RF-/ACPA-																			
2A	**Sample ID**	**128**	**140**	**150**	**153**	**161**	**163**	**171**	**177**	**190**	**191**	**198**	**203**	**207**	**210**	**212**	**214**	**221**	**224**	**227**	**231**
Protein	**Peptide**	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml
1	ALG	26,0	25,3	25,8	12,2	15,7	40,0	38,6	25,2	9,2	11,2	20,6	33,9	27,8	16,8	12,1	14,1	13,1	21,3	18,9	14,7
	TLD	135,6	42,6	44,4	33,5	88,8	94,4	73,2	64,9	27,5	28,8	51,2	65,6	48,5	27,5	18,8	29,0	24,2	38,3	37,1	23,9
	DLL	18,2	17,3	21,8	9,9	13,9	39,4	32,0	21,9	12,1	16,1	28,4	31,1	25,9	14,0	9,9	14,7	12,4	20,2	18,7	11,4
2	ATF	4753,8	997,7	2034,8	399,5	2300,1	1092,4	1406,8	2226,1	1184,6	676,4	560,7	999,1	1125,6	934,5	1066,1	1672,2	763,3	1174,8	935,7	1620,9
	IPS	5288,5	1564,3	3129,2	604,7	1791,1	1675,7	2231,5	3966,3	1223,1	854,8	769,7	1826,8	2896,8	1510,7	1885,1	3439,4	1548,2	2475,9	1806,9	2664,6
	TSS	4337,0	587,2	1240,8	796,8	2096,1	868,3	837,4	1435,3	1028,7	694,5	642,7	751,6	954,9	710,1	774,7	1036,3	463,1	821,0	552,9	804,5
3	GVT	1286,4	344,9	613,5	249,3	795,6	810,3	330,3	582,1	16,8	425,7	350,6	545,2	584,9	317,9	384,1	379,8	102,0	470,3	317,5	133,2
	LED	263,7	100,3	143,6	40,0	139,6	183,1	75,3	137,8	1,7	104,8	96,0	140,6	157,3	91,8	113,2	124,3	32,4	132,4	124,6	29,4
4	WEL	1803,4	576,1	1485,7	529,2	1030,3	1425,2	1423,5	1175,6	677,7	657,6	710,3	867,2	1184,9	895,9	939,2	1001,4	618,5	1209,8	1025,8	1021,3
	VIP	448,1	110,7	288,4	134,4	176,4	296,1	315,7	228,0	149,1	142,2	174,5	157,1	232,6	159,8	158,2	194,9	116,4	247,5	178,4	175,7
	EAQ	311,7	128,3	300,5	111,1	174,4	261,2	237,4	213,6	118,6	117,1	127,3	142,5	197,4	134,0	137,9	172,8	95,8	195,3	166,2	157,1
5	YWG	172,9	29,2	101,6	23,7	47,0	50,4	80,7	42,9	35,8	27,6	19,2	42,3	42,4	43,3	48,1	43,0	33,4	46,7	36,4	41,6
	LLN	2703,7	367,3	1212,2	356,8	936,8	872,3	945,0	711,0	371,3	253,0	168,9	366,8	223,1	352,9	352,9	369,3	299,9	446,8	273,0	301,9
6	YTF	324,9	156,1	348,7	130,3	283,6	300,2	359,8	248,4	119,0	128,9	91,2	216,8	258,0	211,3	150,9	162,7	120,0	265,1	187,8	214,1
	HLS	926,4	366,3	793,0	352,1	781,1	696,8	898,4	597,8	367,3	463,5	378,1	477,7	560,3	472,3	352,2	339,3	252,5	542,3	342,6	439,2
	VLE	252,1	161,5	271,9	125,2	232,7	289,1	308,8	222,5	143,0	150,1	124,2	190,6	222,8	172,2	127,0	145,4	102,1	234,2	154,5	184,2
7	YVG	677,8	278,5	418,9	234,4	370,7	806,3	350,1	357,7	162,7	199,0	288,7	522,6	340,9	170,3	228,6	274,2	128,2	291,8	130,5	293,9
	EQL	1054,0	360,1	511,1	293,1	797,0	1000,6	381,1	456,3	188,5	202,4	245,0	540,2	277,7	185,5	215,8	312,1	141,3	295,5	152,9	368,5
	SDVV	907,8	707,3	1044,9	174,9	656,3	1544,4	770,0	437,5	305,7	355,5	478,4	1187,9	692,8	428,5	552,4	662,2	373,1	648,5	418,1	933,6
8	EHV	291,8	144,7	252,7	120,5	179,2	291,5	248,4	165,3	137,6	114,7	92,7	238,2	190,5	141,7	146,9	161,6	89,2	191,4	100,9	114,2
	EQL	58,2	1602,8	3747,8	2933,3	7633,3	9802,0	6475,6	4893,3	2179,8	1551,6	1159,1	2868,4	1752,7	1495,3	1677,7	2391,5	1380,4	2838,9	1277,4	1367,0
	SDVM	454,6	256,6	507,0	382,6	286,2	489,1	427,6	676,8	184,3	166,8	120,9	441,8	390,3	280,5	302,2	336,0	182,2	372,0	236,7	263,9
9	EIG	159,4	142,4	231,7	108,6	167,7	309,1	178,9	153,5	77,4	120,6	157,0	213,6	156,5	111,3	85,9	112,0	36,9	147,4	104,1	77,4
	AVL	6710,7	3007,5	5406,2	3256,0	6383,7	7023,6	4494,3	3511,3	1612,1	2890,7	3611,5	5209,0	3622,7	2336,3	1620,7	2344,0	847,7	2950,8	2072,1	1627,8
10	VGY	1119,4	1352,5	2543,0	1078,4	1324,1	2571,3	986,4	1105,6	320,4	859,3	1283,5	2128,4	1087,3	900,9	1080,9	1120,8	693,6	1363,4	440,3	1111,9
	VTS	1196,7	1399,7	2755,2	1229,8	1525,9	2865,1	1001,0	1119,4	299,8	932,5	1511,1	2478,5	1084,5	942,9	1178,0	1178,8	666,5	1355,0	430,0	1171,0

		RF+/ACPA+																			
2B	Sample ID	130	131	132	135	137	138	139	144	148	149	156	158	159	160	162	165	167	169	170	176
Protein	**Peptide**	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml
1	**ALG**	48,0	41,0	21,7	28,7	31,7	13,8	37,2	20,8	38,7	13,5	14,8	15,4	12,4	19,5	6,6	23,6	25,7	20,3	24,6	33,5
	TLD	119,7	122,6	51,0	72,3	108,6	32,1	40,1	51,2	82,0	32,3	42,6	28,8	19,5	28,9	13,0	38,1	47,8	42,4	43,2	62,6
	DLL	48,2	43,8	21,5	27,5	32,1	13,5	18,0	21,8	41,5	16,9	19,0	12,8	9,4	17,4	7,3	24,1	24,3	20,7	24,4	31,4
2	**ATF**	1406,2	2635,2	2311,3	1768,2	701,5	173,4	1326,8	1585,7	2638,2	815,5	877,4	811,6	729,2	770,3	464,2	1450,6	1352,6	950,4	2679,5	1690,2
	IPS	2200,4	4052,7	3822,5	2248,5	1003,5	603,0	2421,0	2654,9	4547,4	982,7	1410,7	1454,1	1665,6	1882,6	581,1	2703,3	2817,3	2088,3	5529,5	3807,8
	TSS	663,2	1346,2	1285,7	1256,7	1163,8	469,4	811,6	1081,8	1614,1	884,0	714,9	663,1	535,6	674,3	315,6	807,6	827,9	544,9	1495,5	1083,6
3	**GVT**	533,1	658,5	619,7	497,1	498,0	284,0	447,4	567,2	787,2	427,1	331,9	277,3	215,1	292,2	186,8	382,1	500,5	412,9	562,7	864,7
	LED	124,6	155,0	140,0	103,4	106,3	60,9	120,7	142,1	209,2	124,6	100,4	75,3	58,7	100,6	49,9	114,0	159,7	133,5	168,0	292,4
4	**WEL**	679,0	1394,5	823,9	932,0	1360,8	492,7	1114,9	1159,6	1592,1	946,5	738,2	718,4	792,2	821,9	338,4	985,0	998,9	973,5	1358,9	1552,0
	VIP	121,0	283,8	280,7	205,4	279,0	125,1	190,5	204,1	299,4	185,3	162,6	155,1	175,0	158,9	70,4	194,0	191,1	188,0	259,9	248,3
	EAQ	124,7	214,7	257,9	174,7	218,6	86,6	181,0	173,8	241,7	153,6	114,7	123,0	133,6	118,8	51,5	148,1	145,3	150,5	195,1	209,1
5	**YWG**	15,8	80,7	74,6	41,0	54,5	13,9	35,9	48,9	59,8	46,0	38,3	26,8	40,4	37,6	26,7	53,5	63,3	47,0	56,4	77,1
	LLN	185,8	1125,5	995,3	686,8	1020,9	145,5	536,7	533,3	735,0	227,3	231,7	311,5	259,4	309,6	181,6	369,5	522,4	381,2	458,3	600,9
6	**YTF**	193,5	336,5	319,9	258,5	338,2	112,6	218,2	219,2	290,6	161,9	137,1	138,6	176,9	149,5	59,9	208,0	227,8	219,0	295,7	284,3
	HLS	418,2	787,8	790,6	640,0	937,5	296,1	579,4	504,3	662,0	541,6	417,9	364,4	414,8	355,0	150,4	505,0	484,1	489,7	635,8	550,2
	VLE	122,1	343,6	287,1	189,2	290,5	93,9	189,9	195,0	264,3	181,9	144,5	126,7	148,5	128,8	54,8	170,3	188,4	191,7	264,7	239,7
7	**YVG**	500,2	621,1	421,8	529,5	905,3	365,9	413,2	441,0	442,5	163,3	260,8	268,3	232,2	408,4	129,8	448,2	459,0	432,4	366,5	541,6
	EQL	601,4	676,1	515,9	663,1	1149,2	380,6	470,5	484,1	545,2	156,0	274,6	265,6	246,8	356,3	109,0	409,4	475,4	413,8	411,7	656,9
	SDVV	1249,5	1409,5	1066,0	1181,9	1514,8	666,6	810,1	1000,1	901,6	340,7	537,6	494,8	630,9	873,8	206,6	1007,6	470,3	952,0	843,5	1668,7
8	**EHV**	113,9	303,8	318,9	155,2	273,2	152,0	202,7	238,0	232,5	116,4	130,6	89,5	91,2	183,0	43,3	217,2	174,9	139,9	126,7	272,0
	EQL	1833,6	4924,0	4223,0	3863,3	8066,3	2416,5	4992,7	6269,1	5556,6	1553,2	2259,9	1499,9	1012,5	2250,1	457,0	2113,3	2854,6	2217,3	1807,2	4135,9
	SDVM	232,0	575,7	616,7	259,3	368,5	239,1	323,9	425,7	424,4	186,3	202,4	146,4	184,5	360,2	68,2	403,7	974,1	264,1	247,9	673,4
9	**EIG**	209,9	275,2	173,2	184,9	464,1	119,4	213,9	202,0	209,1	79,3	118,1	97,8	95,1	163,8	62,7	163,1	168,4	202,6	160,3	323,2
	AVL	4286,4	7050,9	3852,8	4282,5	12553,8	2535,6	4895,2	4551,0	5859,3	1868,5	2435,8	2375,7	2385,5	3824,5	1313,2	3460,5	3441,9	4065,2	3103,2	6445,0
10	**VGY**	1808,2	2165,9	1596,9	1863,9	2026,6	1143,5	942,0	1322,3	1548,4	826,7	1218,2	968,0	1120,8	1348,9	432,0	1731,5	1776,7	1057,8	1945,9	2626,6
	VTS	2046,4	2498,4	1659,8	1980,6	2388,4	1399,7	973,8	1383,1	1601,0	884,8	1366,1	953,4	1232,2	1560,9	451,5	1923,9	1987,3	1055,3	2053,0	2861,5

2C	Sample ID	RF+/ACPA-																			
		133	136	141	168	175	199	206	243	276	286	289	290	293	303	314	322	329	332	337	347
Protein	Peptide	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml
1	ALG	39,2	20,2	24,4	19,7	14,9	13,0	28,6	30,9	47,7	27,3	165,6	1474,1	3,0	12,3	5,2	8,7	1,5	8,4	9,1	66,1
	TLD	33,8	36,0	45,2	31,1	23,4	20,9	52,7	49,4	47,2	51,6	23,5	24,5	33,6	24,3	21,8	51,2	81,2	59,9	65,7	85,9
	DLL	20,0	17,3	22,8	19,6	11,8	11,0	25,3	29,6	23,8	24,7	16,2	12,9	19,4	12,8	12,7	30,1	43,8	28,5	34,4	44,9
2	ATF	1802,6	1636,1	1574,5	1283,5	861,0	760,8	1107,6	1962,8	768,4	617,2	765,3	471,9	1114,7	882,8	1134,0	966,6	1365,9	1242,7	1764,4	891,3
	IPS	3354,3	3428,4	3706,9	2982,1	1865,7	1366,2	1874,0	4133,9	1316,8	962,5	1489,7	912,0	2145,1	1736,9	2110,0	2027,3	2809,3	2590,8	3268,4	1607,0
	TSS	1135,2	984,8	1055,4	1356,7	957,3	828,6	675,7	2034,3	708,2	760,5	772,4	508,6	1182,2	817,4	934,5	819,0	962,3	835,1	1245,8	614,9
3	GVT	629,7	387,3	372,4	153,6	325,5	382,8	469,2	574,6	502,5	576,3	201,2	247,0	442,0	459,5	397,8	449,4	511,7	150,6	594,0	444,5
	LED	238,7	164,3	144,7	5,1	88,8	90,5	120,8	154,4	124,4	159,2	38,4	64,0	90,9	116,9	106,6	109,4	214,1	26,7	223,1	185,1
4	WEL	1580,5	1069,5	1147,9	2001,5	1172,3	1050,6	1515,0	2233,1	957,5	1357,0	1038,9	921,8	1561,9	1498,2	1093,2	2145,0	1945,0	2067,8	1812,4	1202,8
	VIP	281,7	145,5	206,1	308,1	199,2	165,0	239,8	330,8	196,5	243,2	172,6	159,1	262,7	214,3	148,8	295,8	223,8	252,8	242,4	149,9
	EAQ	253,4	143,2	177,0	232,1	164,0	144,0	139,0	240,2	129,2	163,7	115,3	118,1	220,8	216,3	110,3	244,1	185,8	223,7	203,3	138,7
5	YWG	85,6	38,6	55,2	97,5	57,3	33,8	34,7	78,5	47,0	55,7	31,4	35,3	46,2	55,9	39,1	71,1	53,2	47,0	63,5	36,4
	LLN	659,1	298,9	395,1	434,0	276,1	200,9	159,2	361,5	255,0	233,6	193,2	221,9	215,5	261,4	167,3	303,3	289,8	273,5	318,2	165,7
6	YTF	322,8	255,0	231,7	264,7	180,8	157,9	158,3	305,6	140,6	166,9	155,7	122,8	229,5	256,3	161,8	301,0	298,9	365,7	295,1	225,3
	HLS	674,1	497,4	437,0	690,1	450,2	399,5	432,3	821,6	385,1	425,2	408,8	336,4	572,9	633,3	370,5	704,4	670,4	818,9	623,4	454,5
	VLE	249,7	223,6	216,0	221,8	169,8	136,0	134,8	261,4	137,5	162,1	119,2	113,7	210,6	233,1	129,2	231,1	234,4	299,5	250,7	175,9
7	YVG	348,3	296,2	342,7	440,8	202,5	188,0	475,6	370,1	517,7	529,1	197,6	158,6	373,1	237,4	232,5	579,4	377,1	178,2	530,7	546,4
	EQL	375,4	404,9	370,1	289,7	186,2	169,9	416,8	303,2	382,9	436,5	163,1	127,3	315,8	194,5	224,8	445,4	427,1	193,8	632,0	620,5
	SDVV	1187,3	1034,4	1026,2	970,6	499,2	438,7	1062,9	1194,6	1090,3	1316,5	435,3	370,8	781,0	462,3	657,8	1397,2	1273,7	773,5	1536,1	1720,2
8	EHV	147,3	211,4	115,5	235,5	133,0	84,0	194,8	237,2	192,3	175,0	119,8	94,4	176,3	181,1	137,5	211,4	207,6	114,3	256,8	238,1
	EQL	1875,0	2689,8	1477,4	2269,7	1277,4	719,4	2462,0	2930,6	2067,1	1716,4	1602,8	1280,2	2036,6	2035,1	1608,4	2001,4	3430,0	1801,0	3700,5	3016,1
	SDVM	336,6	467,5	266,4	473,5	247,8	154,5	414,5	502,6	382,5	392,3	204,6	132,7	278,8	305,9	252,8	398,7	520,4	261,5	643,1	571,2
9	EIG	144,2	211,7	144,2	92,8	120,3	105,3	273,0	163,1	181,7	298,0	130,5	77,9	123,3	142,2	95,7	166,8	172,6	147,8	254,5	327,6
	AVL	2660,3	4343,1	2922,6	5071,3	2866,9	2653,4	5777,3	3642,3	4358,2	6543,9	2951,0	1797,9	2663,0	3117,1	2058,1	4114,4	3069,3	2849,2	4899,7	7373,9
10	VGY	1731,8	1196,7	984,0	1142,5	555,5	1196,0	1279,5	1189,5	1716,7	2832,2	463,7	449,7	713,6	1486,7	1144,4	2126,1	2520,1	1195,3	2092,3	1374,2
	VTS	1880,4	1212,1	975,6	1166,2	563,7	1239,4	1342,0	1223,1	1587,1	2644,6	500,1	471,0	708,4	1529,6	1295,3	2172,3	2571,8	1167,5	2038,8	1366,6

2D	Sample ID	125	134	143	152	172	174	230	240	256	258	261	269	280	288	304	313	336	341	356	357
											RF-/ACPA+										
Protein	Peptide	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml	µg/ml
1	ALG	18,0	10,6	27,1	12,7	31,7	11,9	28,0	75,0	83,2	24,3	10,9	279,5	0,0	1,2	1,3	1,7	1,4	10,0	4,0	6,9
	TLD	29,5	16,7	42,9	20,4	56,5	21,8	55,2	36,6	21,4	69,9	26,9	34,9	44,9	27,3	35,0	52,2	33,9	52,2	39,7	121,3
	DLL	14,6	8,7	22,3	8,2	29,4	11,9	30,5	17,1	13,5	42,7	14,6	17,9	19,3	12,0	19,8	23,3	16,4	26,7	19,1	61,4
2	ATF	1810,1	623,1	2056,6	1178,2	1390,0	865,0	1958,3	608,3	940,6	1241,9	516,3	1097,2	1673,6	498,8	1000,8	2532,2	1195,3	1368,9	1366,8	3300,6
	IPS	1991,4	1014,3	3006,5	1876,4	2028,1	1237,4	3130,0	1210,2	1761,2	2472,5	878,0	2234,9	2810,9	848,1	1974,7	4428,2	2368,6	2538,7	2355,3	5223,0
	TSS	954,7	552,1	1279,9	673,8	678,8	571,7	1107,7	400,9	612,8	885,9	787,6	770,9	1399,9	806,9	992,6	1618,1	997,6	1012,6	885,3	1613,2
3	GVT	355,9	234,2	430,2	300,8	604,4	30,2	70,5	137,1	531,6	531,3	442,4	403,5	662,8	324,6	511,4	528,6	527,1	459,6	162,2	911,2
	LED	89,2	66,8	128,4	73,4	155,9	2,3	13,8	22,6	127,6	138,9	91,0	127,6	123,6	71,5	134,6	165,5	167,7	164,0	47,1	444,6
4	WEL	1524,0	755,6	1739,2	935,0	1269,4	1056,2	1900,7	810,1	1172,4	1853,9	1235,4	1322,5	1931,7	458,8	1463,0	2220,8	2078,2	1515,7	1565,5	2899,4
	VIP	195,6	100,2	257,1	123,5	172,6	137,5	266,9	122,0	172,3	295,5	179,0	210,1	376,3	77,5	206,5	299,1	230,0	167,0	167,2	277,5
	EAQ	175,8	91,7	218,9	105,1	138,4	120,2	191,8	111,1	169,2	243,2	171,7	164,5	252,7	181,3	175,8	288,0	225,3	184,6	172,0	324,6
5	YWG	51,5	30,8	63,4	50,2	63,0	33,0	96,1	34,8	36,7	74,2	43,0	38,3	64,8	5,7	59,5	77,6	49,3	22,3	42,7	38,2
	LLN	256,8	179,6	401,4	366,6	415,6	181,2	660,7	200,7	251,6	389,0	255,2	192,1	383,9	11,5	299,8	424,3	296,2	84,4	282,6	253,4
6	YTF	241,6	120,4	280,6	164,8	178,3	117,7	222,5	236,3	195,9	264,6	128,8	204,1	257,7	185,3	207,6	292,1	261,8	257,2	216,2	449,9
	HLS	601,6	291,7	634,6	399,9	542,5	332,1	661,5	711,3	490,9	708,9	507,9	544,5	748,9	495,9	538,9	726,7	608,3	612,3	439,0	880,4
	VLE	217,4	106,0	254,4	139,7	185,0	117,1	245,8	255,7	171,5	243,2	203,4	192,2	209,6	184,3	179,6	298,0	222,1	217,8	182,0	387,5
7	YVG	278,8	130,5	293,6	243,9	613,1	216,5	665,8	194,9	396,4	743,2	288,7	344,1	482,6	159,1	356,4	385,1	255,3	295,1	235,1	535,2
	EQL	292,9	121,9	283,5	223,5	578,3	179,4	485,8	166,2	345,7	646,5	162,0	320,2	346,0	44,6	318,5	444,7	273,2	320,5	262,2	813,6
	SDVV	717,6	313,8	827,2	614,0	1336,1	915,8	4647,0	342,1	953,8	1623,6	450,2	823,3	980,3	611,0	457,5	1320,3	851,8	1010,4	789,1	2241,9
8	EHV	174,4	81,5	221,6	137,8	297,9	117,2	164,0	136,2	128,1	380,2	183,3	116,7	308,3	117,1	147,6	299,2	206,8	176,3	168,3	294,2
	EQL	1726,9	729,3	2293,5	1437,1	3233,3	1416,3	1654,0	1508,0	1277,7	3653,7	1256,4	1111,2	3150,8	497,5	1540,6	3730,8	2168,3	2128,4	1747,9	3177,0
	SDVM	376,5	170,4	462,1	306,6	614,6	330,0	361,5	198,1	207,3	607,7	229,1	208,2	477,4	194,2	650,8	668,4	461,9	424,8	404,7	842,5
9	EIG	109,1	92,6	184,2	122,3	340,8	90,3	256,1	110,7	194,1	352,9	132,6	178,9	154,2	119,0	139,8	199,9	136,2	217,6	158,6	420,8
	AVL	2311,6	1893,5	3114,0	2442,6	7192,6	1746,9	5563,5	2767,0	4287,8	5893,1	2501,5	3770,8	2823,7	2450,1	2817,8	4110,5	2720,3	4091,8	2781,9	7125,1
10	VGY	894,8	794,3	1883,9	1099,2	2228,4	632,2	1733,9	868,1	1610,1	2500,1	499,9	914,7	746,9	132,8	593,7	2847,2	1863,5	2119,0	852,8	2303,2
	VTS	928,7	804,0	1863,9	1096,4	2368,4	553,8	1809,4	906,8	1637,3	2548,6	508,7	915,3	833,4	153,0	634,5	2722,0	1847,6	2133,9	861,1	2323,1

As it is clear from Tables 2A-D, with the same protein but considering different peptides, different concentration values of the target protein are obtained. This phenomenon is due to several factors such as the non-total protein hydrolysis with trypsin, the different intrinsic ease of ionization and fragmentation efficiency of peptides[28] and, above all, the suppression of ionization effect due to the matrix complexity.

This unique method allows to quantify the ten target proteins simultaneously in a single analysis, overcoming the problems of multiplex analyses that occur with enzyme immunoassays, such as ELISAs. In addition, without fractionation or enrichment strategies, MRM assays allow for the quantitation of protein in the low µg/mL or high ng/mL concentration range, whereas immunoaffinity depletion of abundant blood proteins and minimal protein fractionation can enable quantitation in the low ng/mL range.

5.3.3 Prism analysis

In this section the statistical analysis results are reported. Non-parametric Mann-Whitney tests, on peak area ratio (endogenous/heavy) obtained from sera sample analyses, were performed by using Prism GraphPad version 7.

The RF-/ACPA- sub-group, also called serum negative condition (SNRA) condition, was selected as Control condition, as published data reports, because a low percentage of subjects with rheumatoid arthritis is classified as serum negative. SNRA patients manifested more active disease at baseline but showed a better response to treatment compared with SPRA. SNRA does not appear to be a benign subtype of RA[24].

In addition, patients with this condition present a long-term minor damage compared to those who are seropositive for RF only and these in turn presenting a lesser degree of detriment than those in which the RA is diagnosed with positive ACPA[14].

Several peptides were identified that show a statistically significant modulation ($p < 0.05$) as indicated in Figures 3, 4 and 5. All the figures show the ratio of the area of the endogenous peptide to its standard in relation to the conditions of study.

In Figure 3, the comparison of the RF+/ACPA+ condition versus the control (RF-/ACPA-) is shown. 9 peptides, belonging to proteins 1, 7, 9 and 10, are significantly increased in RF+/ACPA+.
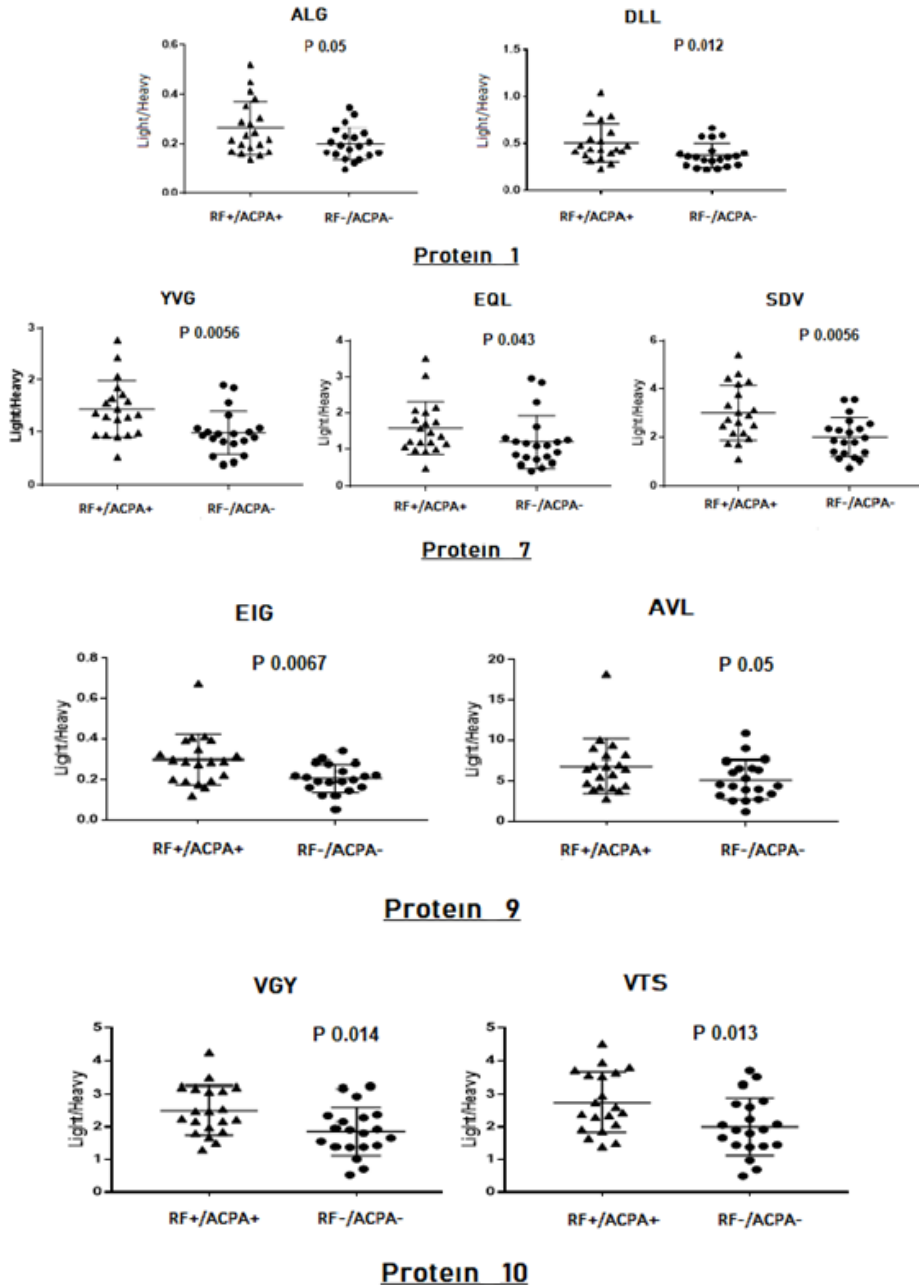
Figura 3: Modulated peptides (P-value <0.05) in the comparison of RF+/ACPA+ condition versus the control (RF-/ACPA-). Peptides: ALG, DLL (protein 1), YVG, EQL, SDV (protein 7), EIG, AVL (protein 9), and VGY, VTS (protein 10).

In the RF+/ACPA- condition (Figure 4), there are 7 statistically modulated peptides. These peptides belong to proteins 4, 5, 6, 7 and 9. All are increased, except those belonging to protein 5.
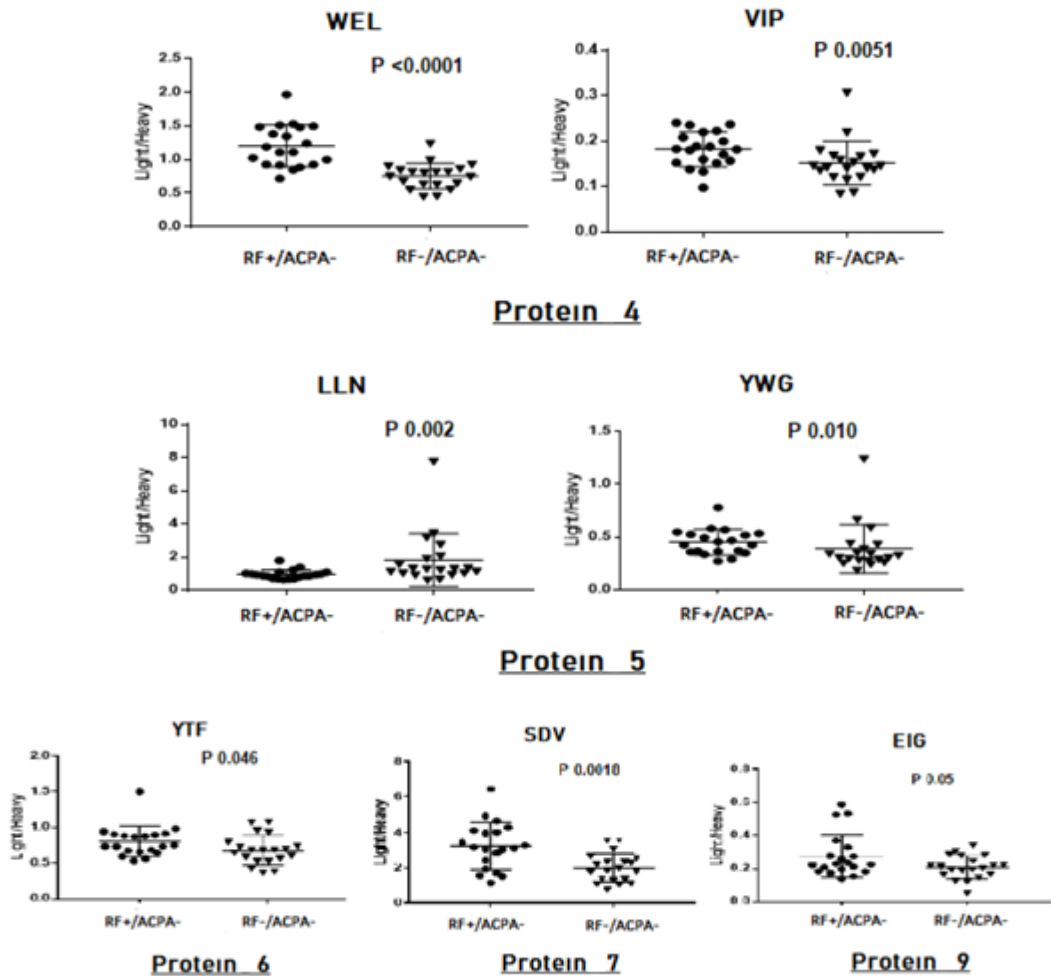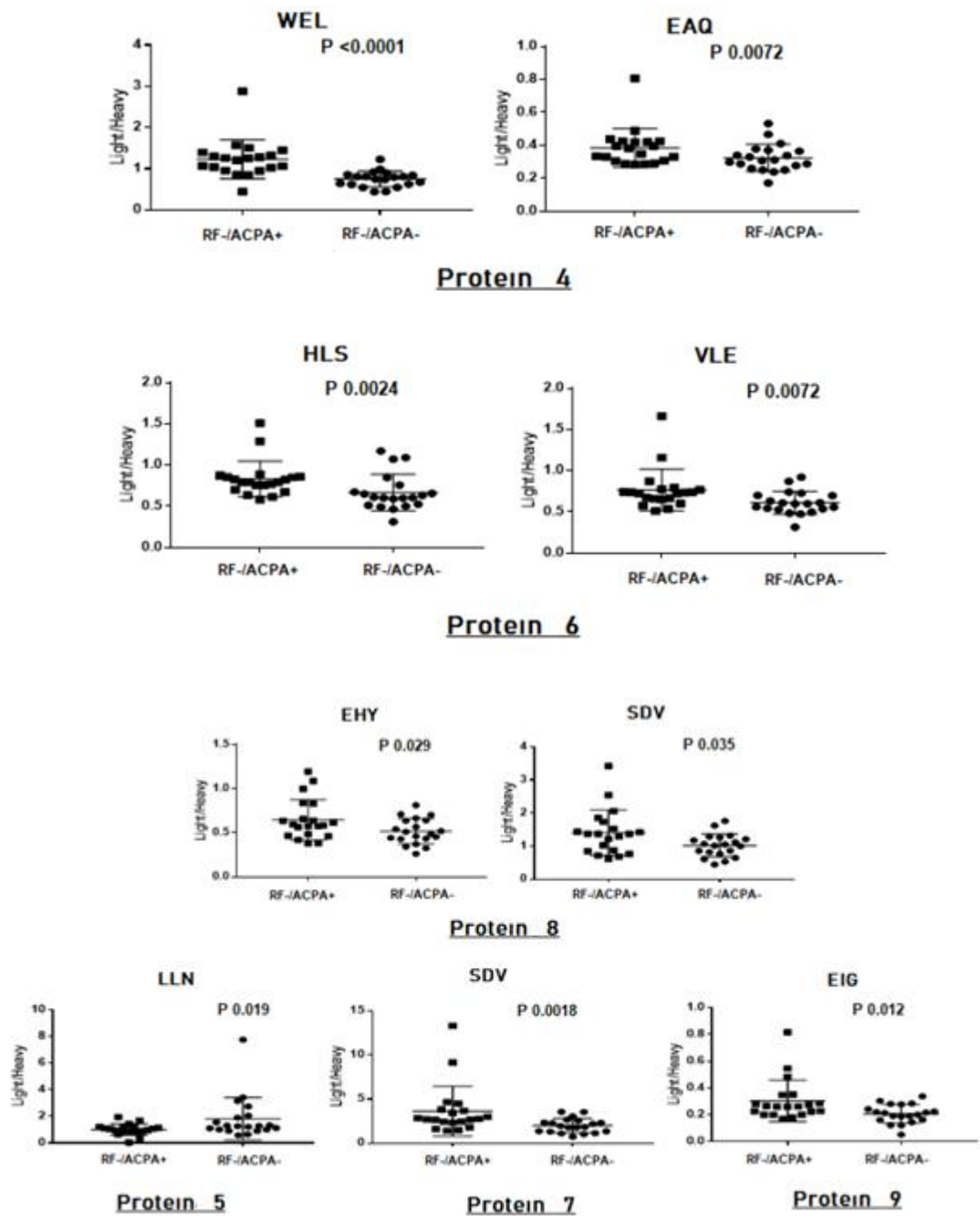
Figura 4: Modulated peptides (P-value <0.05) in the comparison of the RF+/ACPA- condition versus the control (RF-/ACPA-). Peptides: WEL, VIP (protein 4), LLN, YWG (protein 5), YTF (protein 6), SDV (protein) and EIG (protein 9).

Finally, Figure 5 shows the 9 statistically enhanced peptides obtained in the comparison of the RF-/ACPA+ condition and the control. These belong to proteins 4, 5, 6, 7, 8 and 9 and, as in the previous case, they are all increased except for the peptides of protein 5.

Figura 5: Modulated peptides (P-value <0.05) in the comparison of the RF-/ACPA+ condition versus the control (RF-/ACPA-). Peptides: WEL, EAQ (protein 4), HLS, VLE (protein 6), EHY, SDV (protein 9), LLN (protein 5), SDV (protein 7) and EIG (protein 9).

The statistical analysis allowed us to observe that, of the 26 peptides analysed, significant differences were found in 19 of them (Table 3).

Table 3: This table summarized all the statistically significant peptides (p<0,05) for the three conditions RF+/ACPA+, RF-/ACPA+ and RF+/ACPA- against the selected control condition RF-/ACPA-.

Protein	Peptide	RF+/ACPA+	RF+/ACPA-	RF-/ACPA-
1	ALG	X		
1	DLL	X		
4	WEL		X	X
4	EAQ			X
4	VIP		X	
5	LLN		X	X
5	YWG		X	
6	HLS			X
6	VLE			X
6	YTF		X	
7	YVG	X		
7	EQL	X		
7	SDVV	X	X	X
8	EHV			X
8	SDVM			X
9	EIG	X	X	X
9	AVL	X		
10	VGY	X		
10	VTS	X		

All proteins are overexpressed except for protein 5, which appears diminished in the second and third comparisons. This protein is involved in the vitamin A transport regulation through blood plasma from the deposit located in the liver to peripheral tissues. This vitamin is essential for growth, cell differentiation and bone development, so its low expression could influence the lack of bone recovery in patients suffering degradation due to RA. In addition, the SDV and EIG peptides, belonging to proteins 7 and 9 respectively, appear increased under all conditions. This could indicate that those two proteins are closely related to each other.

All the proteins that are modulated participate in the immune response of the organism and are characterized by increasing their levels during the acute phase response. Their functions are also related to inflammatory processes, regulation of angiogenesis, cell proliferation and maintenance of soft and bone tissues. In

addition, some of them are responsible for transporting other proteins through the bloodstream or in the protein storage, hence an alteration in their levels can be very harmful to the body.

Sokolove et al.[17], evaluate the interaction and possible relationship of RF and ACPA with RA. Comparing the group RF+/ACPA+, with RF-/ACPA-, they observed a higher activity of the disease, as well as higher levels of ESR and CRP than normal, which can be related to inflammatory processes, in this case derived from the RA. In addition, the finding of high levels of several cytokines compared to control supports the possible relationship between these biomarkers and the disease.

Aletaha et al.[29] revealed that the lowest DAS is associated to seropositive patients only for ACPA, while the highest DAS reaches the patients RF+/ACPA+ or RF+/ACPA-. These data suggest that high DAS are related to the positivity of rheumatoid factor. In addition, they compared the groups avoiding the positive double and, apart from obtaining again greater activity in the RF+ group, greater activity was seen in the ACPA- group than in the ACPA+, which supports that the RF+ is associated with higher DAS and, also, the possibility that ACPA+ is associated with less activity.

5.4 Conclusion

The social impact of RA as personal disability, inability to work and costs of health urges us to look for serum biomarkers that, together with radiographic data, tests of the Rheumatoid Factor and ACPA allow to early diagnose this pathology in order to promptly intervene with a therapy that is, as much as possible, ad personam and that allows the follow-up of the patient.

The present work fits perfectly into the research context of potential new protein biomarkers for RA diagnostics.

MRM mass spectrometry guarantees high specificity and selectivity, increased by using SIS peptides and allows to overcome the limits of multiplex analysis of ELISA enzymatic assays, maintaining, however, comparable sensitivities.

In this study 80 serum samples of patients with RA were analysed. The samples were divided into four sub-groups based on the RF and ACPA values.

Targeted analysis involved the monitoring and quantification of a panel of ten proteins that, from previous large-scale proteomic studies, were significantly modulated. These proteins are related to RA because they are involved in the acute phase of inflammation, in bone mineralization and in the immune system response.

Validation of the protein panel in larger cohorts of healthy patients and patients with other diseases is necessary to confirm the specificity of these proteins for RA. This validation is pending to be carried out by ELISA immunoassays.

5.5 References

(1)   Karouzakis, E.; Rengel, Y.; Jüngel, A.; Kolling, C.; Gay, R. E.; Michel, B. A.; Tak, P.; Gay, S.; Neidhart, M.; Ospelt, C. DNA methylation regulates the expression of CXCL12 in rheumatoid arthritis synovial fibroblasts. Genes and immunity 2011, 12, 643.

(2)   Huber, L. C.; Brock, M.; Hemmatazad, H.; Giger, O. T.; Moritz, F.; Trenkmann, M.; Distler, J. H.; Gay, R. E.; Kolling, C.; Moch, H. Histone deacetylase/acetylase activity in total synovial tissue derived from rheumatoid arthritis and osteoarthritis patients. Arthritis & Rheumatism 2007, 56, 1087-1093.

(3)   Orr, C.; Vieira-Sousa, E.; Boyle, D. L.; Buch, M. H.; Buckley, C. D.; Cañete, J. D.; Catrina, A. I.; Choy, E. H.; Emery, P.; Fearon, U. Synovial tissue research: a state-of-the-art review. Nature Reviews Rheumatology 2017, 13, 463.

(4)   Firestein, G. S.; McInnes, I. B. Immunopathogenesis of rheumatoid arthritis. Immunity 2017, 46, 183-196.

(5)   Kvien, T. K.; Uhlig, T.; ØdegÅrd, S.; Heiberg, M. S. Epidemiological aspects of rheumatoid arthritis. Annals of the New York Academy of Sciences 2006, 1069, 212-222.

(6)   Rossini, M.; Rossi, E.; Bernardi, D.; Viapiana, O.; Gatti, D.; Idolazzi, L.; Caimmi, C.; DeRosa, M.; Adami, S. Prevalence and incidence of rheumatoid arthritis in Italy. Rheumatology international 2014, 34, 659-664.

(7)   Fries, J. F.; Spitz, P.; Kraines, R. G.; Holman, H. R. Measurement of patient outcome in arthritis. Arthritis & Rheumatism 1980, 23, 137-145.

(8)   Ramamoorthy, R. D.; Nallasamy, V.; Raghavendra Reddy, N. E.; Maruthappan, Y. A review of C-reactive protein: A diagnostic indicator in periodontal medicine. Journal of pharmacy & bioallied sciences 2012, 4, S422.

(9)   Group, B. D. W.; Atkinson Jr, A. J.; Colburn, W. A.; DeGruttola, V. G.; DeMets, D. L.; Downing, G. J.; Hoth, D. F.; Oates, J. A.; Peck, C. C.; Schooley, R. T. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clinical Pharmacology & Therapeutics 2001, 69, 89-95.

(10)   Robinson, W. H.; Lindstrom, T. M.; Cheung, R. K.; Sokolove, J. Mechanistic biomarkers for clinical decision making in rheumatic diseases. Nature Reviews Rheumatology 2013, 9, 267.

(11)   Ahmed, U.; Thornalley, P. J.; Rabbani, N.: Factors influencing the development and effectiveness of biomarkers in rheumatoid arthritis and osteoarthritis. Future Medicine, 2015.

(12)   Colglazier, C. L.; Sutej, P. G. Laboratory testing in the rheumatic diseases: a practical review. South Med J 2005, 98, 185-191.

(13)   Ingegnoli, F.; Castelli, R.; Gualtierotti, R. Rheumatoid factors: clinical applications. Disease markers 2013, 35, 727-734.

(14)   Kroot, E. J. J.; De Jong, B. A.; Van Leeuwen, M. A.; Swinkels, H.; Van Den Hoogen, F. H.; Van't Hof, M.; Van De Putte, L. B.; Van Rijswijk, M. H.; Van Venrooij, W. J.; Van Riel, P. L. The prognostic value of anti–cyclic citrullinated peptide antibody in patients with recent-onset rheumatoid arthritis. Arthritis & Rheumatism: Official Journal of the American College of Rheumatology 2000, 43, 1831-1835.

(15)   Taylor, P.; Gartemann, J.; Hsieh, J.; Creeden, J. A systematic review of serum biomarkers anti-cyclic citrullinated peptide and rheumatoid factor as tests for rheumatoid arthritis. Autoimmune Diseases 2011, 2011.

(16)   Gonzalez-Lopez, L.; Rocha-Muñoz, A. D.; Ponce-Guarneros, M.; Flores-Chavez, A.; Salazar-Paramo, M.; Nava, A.; Cardona-Muñoz, E. G.; Fajardo-Robledo, N. S.; Zavaleta-Muñiz, S. A.; Garcia-Cobian, T. Anti-cyclic citrullinated peptide (anti-CCP) and anti-mutated citrullinated vimentin (anti-MCV) relation with extra-articular manifestations in rheumatoid arthritis. Journal of immunology research 2014, 2014.

(17)   Sokolove, J.; Johnson, D. S.; Lahey, L. J.; Wagner, C. A.; Cheng, D.; Thiele, G. M.; Michaud, K.; Sayles, H.; Reimold, A. M.; Caplan, L. Rheumatoid factor as a potentiator of anti–citrullinated protein antibody–mediated inflammation in rheumatoid arthritis. Arthritis & rheumatology 2014, 66, 813-821.

(18)   Schellekens, G. A.; de Jong, B. A.; van den Hoogen, F. H.; Van de Putte, L.; van Venrooij, W. J. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. The Journal of clinical investigation 1998, 101, 273-281.

(19)   Boeters, D. M.; Nieuwenhuis, W. P.; Verheul, M. K.; Newsum, E. C.; Reijnierse, M.; Toes, R. E.; Trouw, L. A.; van der Helm-van, A. H. MRI-detected osteitis is not associated with the presence or level of ACPA alone, but with the combined presence of ACPA and RF. Arthritis research & therapy 2016, 18, 179.

(20)   Daha, N. A.; Toes, R. E. Rheumatoid arthritis: Are ACPA-positive and ACPA-negative RA the same disease? Nature Reviews Rheumatology 2011, 7, 202.

(21)   Ajeganova, S.; Huizinga, T. W. Rheumatoid arthritis: seronegative and seropositive RA: alike but different? Nature Reviews Rheumatology 2015, 11, 8.

(22)   Nordberg, L. B.; Lillegraven, S.; Lie, E.; Aga, A.-B.; Olsen, I. C.; Hammer, H. B.; Uhlig, T.; Jonsson, M. K.; van der Heijde, D.; Kvien, T. K. Patients

with seronegative RA have more inflammatory activity compared with patients with seropositive RA in an inception cohort of DMARD-naïve patients classified according to the 2010 ACR/EULAR criteria. Annals of the rheumatic diseases 2017, 76, 341-345.

(23)   MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 2010, 26, 966-968.

(24)   Choi, S.; Lee, K.-H. Correction: Clinical management of seronegative and seropositive rheumatoid arthritis: A comparative study. PloS one 2018, 13, e0199468.

(25)   Percy, A. J.; Simon, R.; Chambers, A. G.; Borchers, C. H. Enhanced sensitivity and multiplexing with 2D LC/MRM-MS and labeled standards for deeper and more comprehensive protein quantitation. Journal of proteomics 2014, 106, 113-124.

(26)   LeBlanc, A.; Michaud, S. A.; Percy, A. J.; Hardie, D. B.; Yang, J.; Sinclair, N. J.; Proudfoot, J. I.; Pistawka, A.; Smith, D. S.; Borchers, C. H. Multiplexed MRM-based protein quantitation using two different stable isotope-labeled peptide Isotopologues for calibration. Journal of proteome research 2017, 16, 2527-2536.

(27)   Kocełak, P.; Owczarek, A.; Bożentowicz-Wikarek, M.; Brzozowska, A.; Mossakowska, M.; Grodzicki, T.; Więcek, A.; Chudek, J.; Olszanecka-Glinianowicz, M. Plasma concentration of Retinol Binding Protein 4 (RBP4) in relation to nutritional status and kidney function in older population of PolSenior Study. Advances in medical sciences 2018, 63, 323-328.

(28)   Bruins, A. P.; Covey, T. R.; Henion, J. D. Ion spray interface for combined liquid chromatography/atmospheric pressure ionization mass spectrometry. Analytical Chemistry 1987, 59, 2642-2646.

(29)   Aletaha, D.; Neogi, T.; Silman, A. J.; Funovits, J.; Felson, D. T.; Bingham III, C. O.; Birnbaum, N. S.; Burmester, G. R.; Bykerk, V. P.; Cohen, M. D. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis & Rheumatism 2010, 62, 2569-2581.

5.6 Supporting Data

SD Table 1: MRM/MS method. List of selected proteins reporting selected peptide sequence, precursor m/z, product ions m/z, Declustering Potential (DP, V) and Collision Energy (CE, V).

Protein	Peptide sequence		Precursor ion m/z	Product ion m/z	Ion type	DP (V)	CE (V)
1	ALG	Light	576,81	774,41	+2y7	73.2	29.6
				661,33	+2y6	73.2	29.6
				546,3	+2y5	73.2	29.6
		Heavy	581,81	784,42	+2y7	73.2	29.6
				671,33	+2y6	73.2	29.6
				556,31	+2y5	73.2	29.6
1	TLD	Light	1019,05	1166,68	+2y10	105.4	45.5
				1053,59	+2y9	105.4	45.5
				823,5	+2y7	105.4	45.5
		Heavy	1024,05	1176,69	+2y10	105.4	45.5
				1063,6	+2y9	105.4	45.5
				833,51	+2y7	105.4	45.5
1	DLL	Light	590,34	838,48	+2y7	74.2	30.1
				725,39	+2y6	74.2	30.1
				342,2	+2b3	74.2	30.1
		Heavy	595,34	848,49	+2y7	74.2	30.1
				735,4	+2y6	74.2	30.1
				342,2	+2b3	74.2	30.1
2	ATF	Light	612	813,48	+3y7	75.7	30.9
				714,41	+3y6	75.7	30.9
				617,36	+3y5	75.7	30.9
		Heavy	615,33	823,49	+3y7	75.7	30.9
				724,42	+3y6	75.7	30.9
				627,37	+3y5	75.7	30.9
2	IPS	Light	523,31	835,47	+2y7	69.3	27.7
				748,44	+2y6	69.3	27.7
				649,37	+2y5	69.3	27.7
		Heavy	527,31	843,48	+2y7	69.3	27.7
				756,45	+2y6	69.3	27.7
				657,38	+2y5	69.3	27.7
2	TSS	Light	808,95	1010,59	+2y9	90.1	38

				896,55	+2y8	90.1	38
				783,46	+2y7	90.1	38
		Heavy	812,96	1018,6	+2y9	90.1	38
				904,56	+2y8	90.1	38
				791,48	+2y7	90.1	38
3	GVT	Light	913,99	1055,56	+2y9	97.7	41.8
				472,32	+2y4	97.7	41.8
				1143,58	+2b11	97.7	41.8
		Heavy	919	1065,57	+2y9	97.7	41.8
				482,33	+2y4	97.7	41.8
				1143,58	+2b11	97.7	41.8
3	LED	Light	797,4	777,45	+2y7	89.2	37.6
				664,37	+2y6	89.2	37.6
				577,33	+2y5	89.2	37.6
		Heavy	801,4	785,46	+2y7	89.2	37.6
				672,38	+2y6	89.2	37.6
				585,35	+2y5	89.2	37.6
4	WEL	Light	544,76	773,4	+2y6	70.8	28.5
				660,31	+2y5	70.8	28.5
				500,28	+2y4	70.8	28.5
		Heavy	549,77	783,41	+2y6	70.8	28.5
				670,32	+2y5	70.8	28.5
				510,29	+2y4	70.8	28.5
4	VIP	Light	885,96	1020,51	+2y9	95.7	40.8
				933,48	+2y8	95.7	40.8
		Heavy	890,97	1030,52	+2y9	95.7	40.8
				943,49	+2y8	95.7	40.8
4	EAQ	Light	570,82	812,49	+2y7	72.7	29.4
				699,4	+2y6	72.7	29.4
				329,15	+2b3	72.7	29.4
		Heavy	574,82	820,5	+2y7	72.7	29.4
				707,42	+2y6	72.7	29.4
				329,15	+2b3	72.7	29.4
5	YWG	Light	599,82	849,48	+2y8	74.8	30.5
				693,39	+2y6	74.8	30.5
				622,36	+2y5	74.8	30.5
		Heavy	603,82	857,5	+2y8	74.8	30.5

				701,41	+2y6	74.8	30.5
				630,37	+2y5	74.8	30.5
5	LLN	Light	1032,99	1178,55	+2y10	106.4	46
				1107,51	+2y9	106.4	46
				742,39	+2y6	106.4	46
		Heavy	1037,99	1188,56	+2y10	106.4	46
				1117,52	+2y9	106.4	46
				752,4	+2y6	106.4	46
6	YTF	Light	458,23	752,39	+2y6	64.5	25.4
				651,35	+2y5	64.5	25.4
				504,28	+2y4	64.5	25.4
		Heavy	463,24	762,4	+2y6	64.5	25.4
				661,35	+2y5	64.5	25.4
				514,29	+2y4	64.5	25.4
6	HLS	Light	418,91	590,33	+3y5	61.7	20.4
				489,28	+3y4	61.7	20.4
				376,19	+3y3	61.7	20.4
		Heavy	422,25	600,33	+3y5	61.7	20.4
				499,29	+3y4	61.7	20.4
				386,2	+3y3	61.7	20.4
6	VLE	Light	400,25	700,42	+2y6	60.3	23.3
				458,3	+2y4	60.3	23.3
		Heavy	404,26	708,44	+2y6	60.3	23.3
				466,31	+2y4	60.3	23.3
7	YVG	Light	584,99	982,62	+3y8	73.8	29.4
				835,55	+3y7	73.8	29.4
				764,51	+3y6	73.8	29.4
		Heavy	588,33	992,63	+3y8	73.8	29.4
				845,56	+3y7	73.8	29.4
				774,52	+3y6	73.8	29.4
7	EQL	Light	871,91	1186,56	+2y9	94.7	40.2
				1039,49	+2y8	94.7	40.2
				876,42	+2y7	94.7	40.2
		Heavy	876,91	1196,56	+2y9	94.7	40.2
				1049,5	+2y8	94.7	40.2
				886,43	+2y7	94.7	40.2
7	SDV	Light	556,77	910,47	+2y7	71.7	28.9

				811,4	+2y6	71.7	28.9
				401,2	+2b4	71.7	28.9
		Heavy	560,77	918,48	+2y7	71.7	28.9
				819,41	+2y6	71.7	28.9
				401,2	+2b4	71.7	28.9
8	EHV	Light	412,24	798,5	+3y6	61.2	20
				661,44	+3y5	61.2	20
				548,36	+3y4	61.2	20
		Heavy	415,58	808,51	+3y6	61.2	20
				671,45	+3y5	61.2	20
				558,36	+3y4	61.2	20
8	EQL	Light	705	1046,51	+3y8	82.5	35.9
				933,43	+3y7	82.5	35.9
				818,4	+3y6	82.5	35.9
		Heavy	708,33	1056,52	+3y8	82.5	35.9
				943,44	+3y7	82.5	35.9
				828,41	+3y6	82.5	35.9
8	SDV	Light	572,75	942,44	+2y7	72.9	29.5
				843,37	+2y6	72.9	29.5
				712,33	+2y5	72.9	29.5
		Heavy	576,76	950,45	+2y7	72.9	29.5
				851,38	+2y6	72.9	29.5
				720,34	+2y5	72.9	29.5
9	EIG	Light	531,3	819,46	+2y7	69.8	28
				633,4	+2y5	69.8	28
				520,31	+2y4	69.8	28
		Heavy	535,3	827,48	+2y7	69.8	28
				641,41	+2y5	69.8	28
				528,33	+2y4	69.8	28
9	AVL	Light	954,48	1134,56	+2y12	100.7	43.2
				1005,52	+2y11	100.7	43.2
				718,41	+2y8	100.7	43.2
		Heavy	958,49	1142,58	+2y12	100.7	43.2
				1013,54	+2y11	100.7	43.2
				726,42	+2y8	100.7	43.2
10	VGY	Light	490,75	824,4	+2y7	66.9	26.5
				661,34	+2y6	66.9	26.5

				562,27	+2y5	66.9	26.5
		Heavy	495,76	834,41	+2y7	66.9	26.5
				671,35	+2y6	66.9	26.5
				572,28	+2y5	66.9	26.5
10	VTS	Light	602,32	916,49	+2y7	75	30.5
				803,4	+2y6	75	30.5
				675,35	+2y5	75	30.5
		Heavy	606,33	924,5	+2y7	75	30.5
				811,42	+2y6	75	30.5
				683,36	+2y5	75	30.5

SD Figure 1: Calibration curves for each heavy standard peptide.
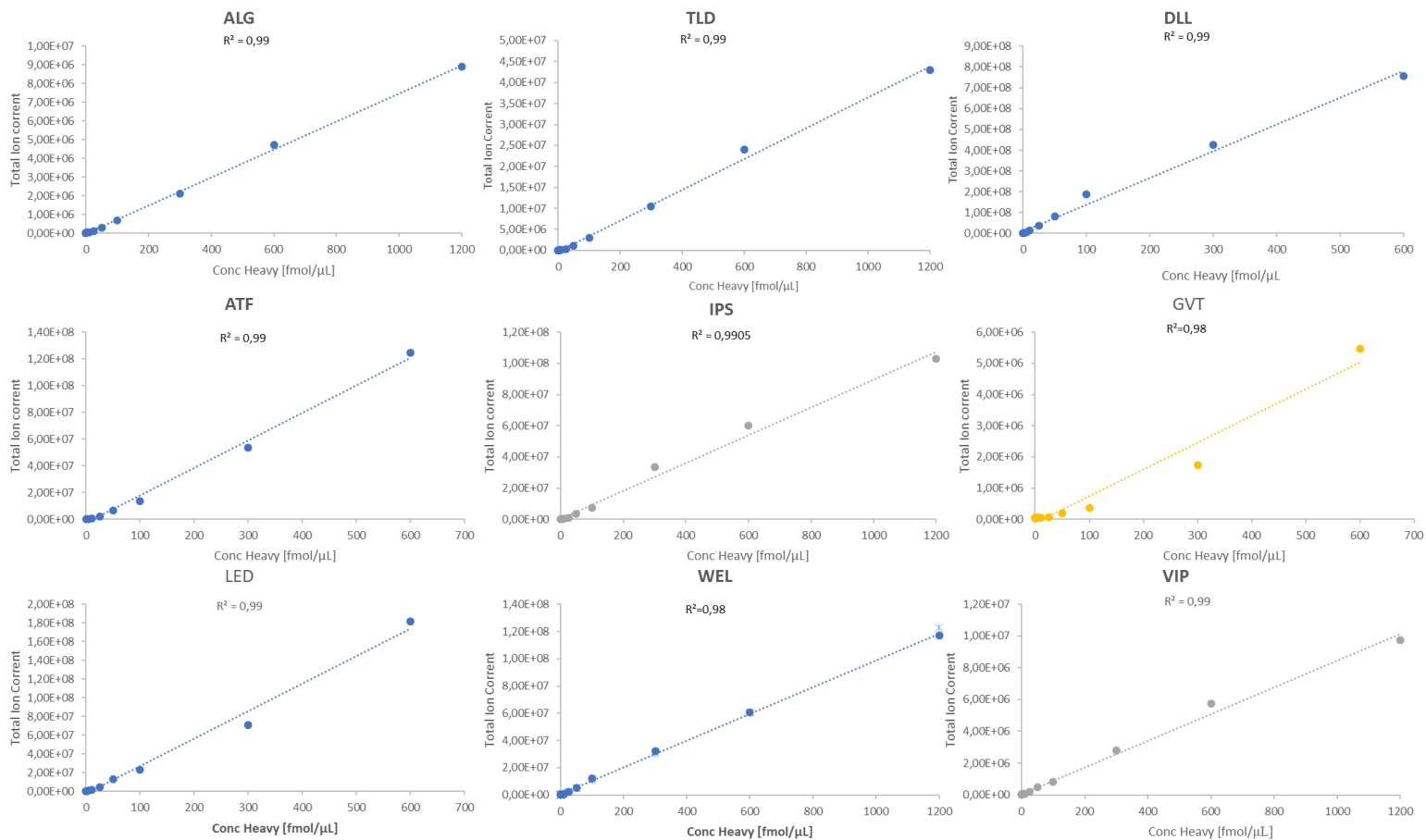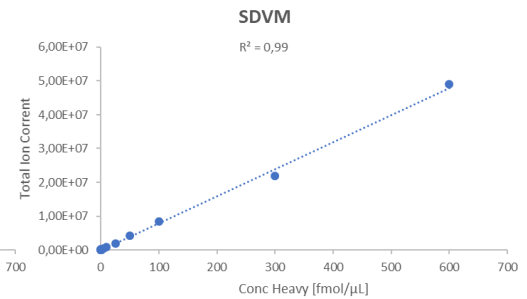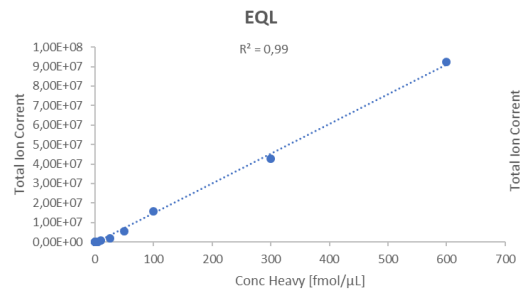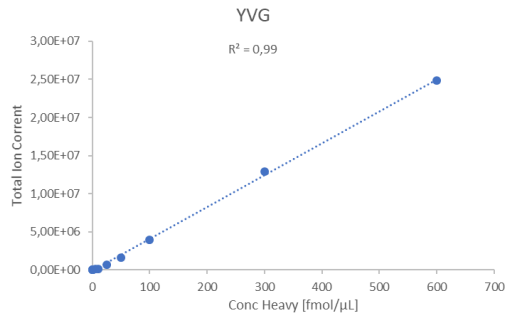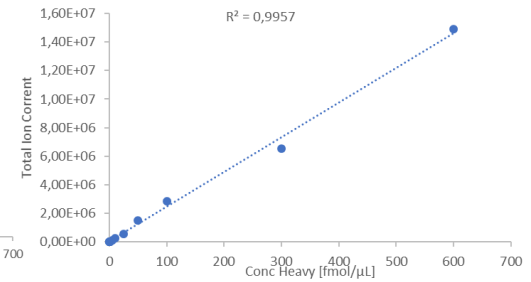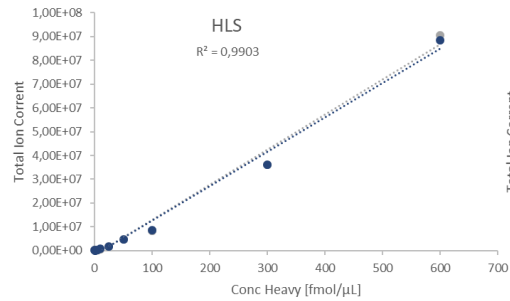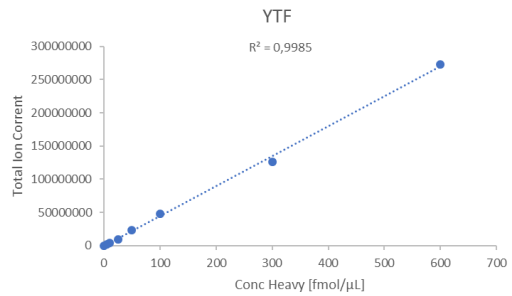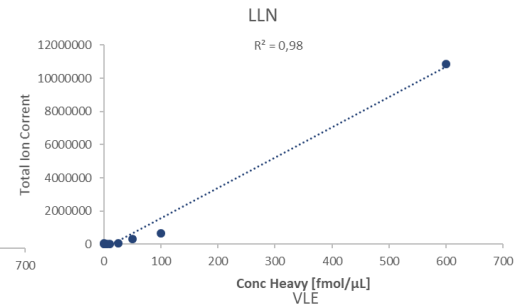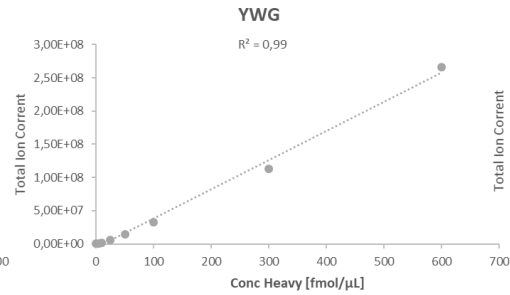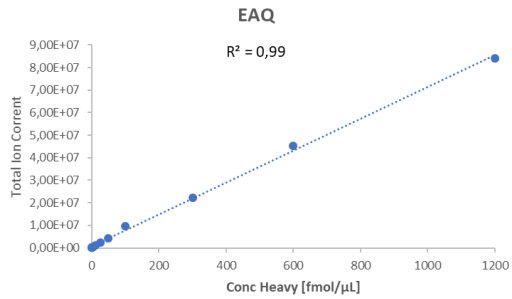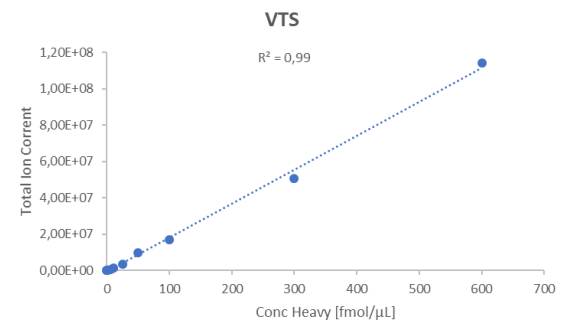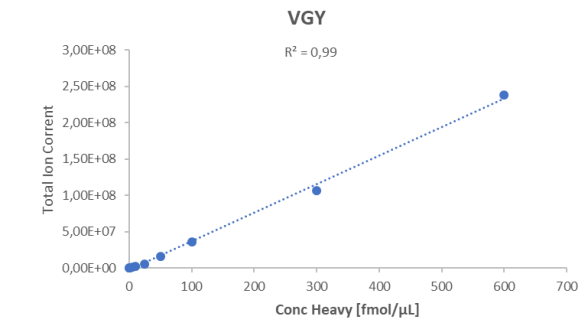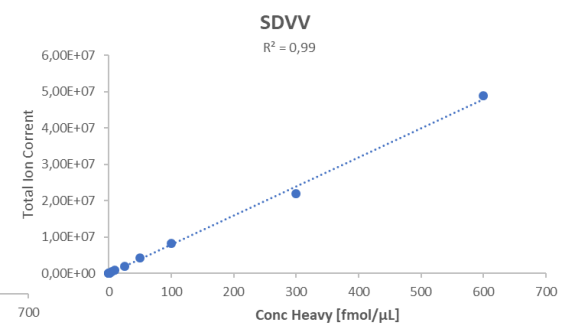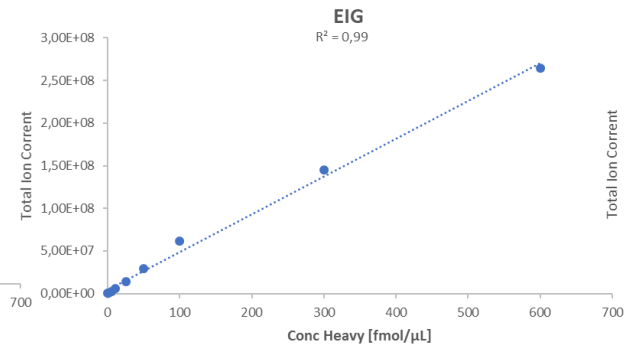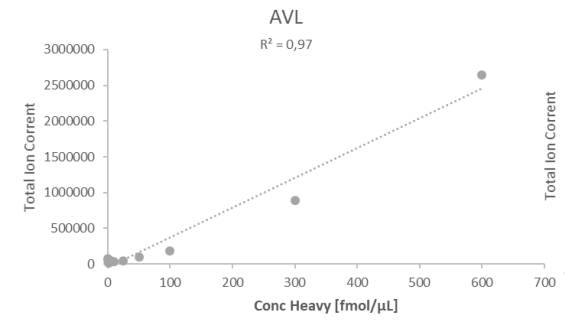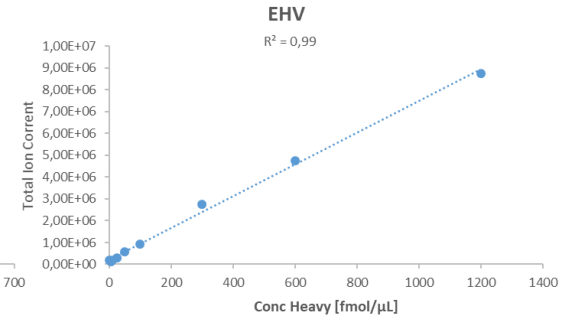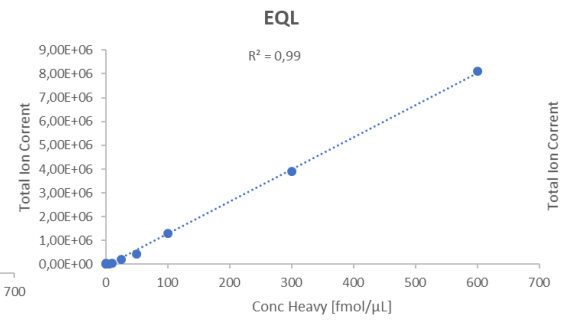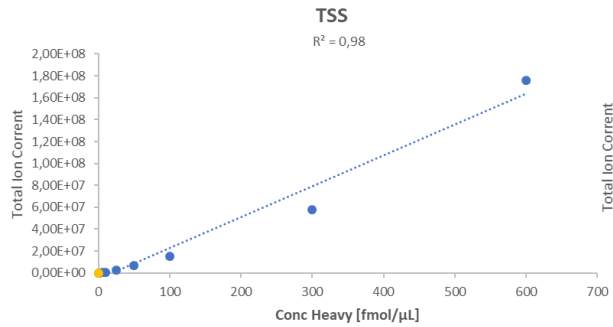
SD Table 2: SIS spiked peptides concentration in sera samples.

Labelled Peptide	Spiked SIS peptides concentration (fmol/μL)
ALG	50
TLD	100
DLL	25
ATF	100
IPS	50
TSS	100
GVT	250
LED	100
WEL	250
VIP	250
EAQ	100
YWG	100
LLN	250
YTF	100
HLS	250
VLE	100
YVG	250
EQL	250
SDVV	250
EHV	250
EQL	250
SDVM	250
EIG	250
AVL	250
VGY	250
VTS	250

Chapter 6

6.1 Introduction

In the last three decades the link between forensic investigations and scientific disciplines, such as Chemistry and Biology, Biology, has been enormously strengthened by the application of new tools for a detailed characterization of the crime scene starting from trace evidence. In particular, DNA testing is fundamental to identify the individuals involved in a crime often leading the investigators to define a unique genetic profile[1].

However, the detection and identification of the type and origin of biofluids at a crime scene are just as important as DNA test to link the nature of the biofluid, the criminal act and the dynamics of the crime. For example, blood stains can indicate some form of physical altercation, whereas detection of semen or vaginal fluid can indicate the involvement of some form of sexual encounter or assault. Moreover, identification of the specific fluid or tissue from which the DNA sample was recovered is fundamental in correctly reconstructing the criminal event and the effective role of the donor.

The most common biofluids found at crime scenes are blood, semen, and saliva, but others such as vaginal fluid, urine, and sweat can also be found. Each of these fluids has one or more presumptive tests that are initially used to give some indication as to the identity of the substance; blood is quickly detected by means of luminol or benzidine[2], whereas traces of semen, saliva, sweat, urine and other biological fluids are detected by ultraviolet light or by other light sources at specific wavelengths[3]. Confirmatory tests, based on biochemical, spectroscopic and microscopy methods, and immunochemical techniques, are necessary to give a legal value to the trace evidence[4-9].

However, these identification methods suffer from several limitations including instability of assayed biomolecules and low selectivity and specificity. As an example of the latter issue, it is not possible to discriminate between alpha-amylase 1 (present in saliva) and 2 (present in semen and vaginal secretion)[10]. Moreover, biochemical assays are specific only for one biological matrix and several cascade tests might then be needed before the biological nature of a

certain stain is uncovered. This is even more challenging when the sample is a mixture of different biological matrices thus greatly increasing the complexity, the cost and the time of the analytical procedures. Finally, some confirmatory tests very often cause sample loss and they are not compatible with downstream individual identification by DNA analysis.

For all these issues, forensic science is looking for a universal confirmatory test for the analysis of unknown stains which will be able to unambiguously identify the type and nature of any biofluids that might be present at a crime scene. The method should be applicable to mixtures, and, more importantly, should preserve the samples for subsequent DNA analysis. Since biofluids have evolved to perform different functions, they contain different proteins, or different combinations of proteins, providing each biological matrix with a unique protein signature that can be used to distinguish among the various biofluids[11-13]. Recently, untargeted proteomics have been introduced for the determination of biological matrices in forensic science based on the identification of the most prominent proteins present in biofluids[11,14]. However, these methods can give poor or uncertain results when high amounts of non-specific proteins occur in the samples, i.e. when low amounts of a particular fluid are present in combination with high amounts of other matrices or when the sample is contaminated by other tissues.

This paper reports a universal method alternative and complementary to "traditional" tests currently used in forensic investigation to identify an unknown stain of biofluids occurring at a crime scene. The method is based on a targeted proteomic approach that makes use of tandem mass spectrometry in multiple reaction monitoring (MRM) mode to selectively monitor a number of specific peptides belonging to proteins biomarkers of individual biological fluids. First, the most prominent and highly specific protein biomarkers for each biological matrix were identified by both conventional LC-MS/MS analyses and comparison with literature data. Specific peptide markers of each protein were then selected on the basis of their mass spectrometric behavior together with their specific precursor ion-product ion transitions as defined by their unique amino acid sequence. Then, a single MRM method was devised to detect the

120

occurrence of the target peptides within the sample leading to the unambiguous discrimination among the different biological matrices in a single analysis due to its high sensitivity, selectivity and accuracy. The optimized method was tested and validated on specimens consisting of four biological matrices (blood, saliva, semen and urine), or mixtures of matrices, spotted on different substrates like cloth, wood, plastic, plaster and paper. Finally, a combined strategy allowing both DNA analysis and biofluids identification on the same sample was developed and demonstrated to be effective in forensic investigation by the analysis of real specimens collected at a crime scene.

6.2 Materials and methods

Control biofluids (blood, saliva, urina, semen) were provided by a private chemical laboratory. Guanidine, Tris(idrossimetil)amminometano cloridrato (TrisHCl), dithiothreitol (DTT), ethylendiaminetetraacetate (EDTA), trypsin, iodoacetamide (IAM), ammonium bicarbonate (AMBIC), trichloroacetic acid (TCA) were purchased from Sigma-Aldrich. Bovine serum albumin (BSA), formic acid (HCOOH), methanol, chloroform, acetonitrile (ACN) are from J.T. Baker. Bradford staining solution was from Bio-Rad. Pipette tips C18 (zip tip) and centrifugal filter units were purchased from Merck Millipore.

6.2.1 In-solution digestion of proteins from biological matrices

Different protocols were performed to improve the number of identified proteins and thus to provide the higher sequence coverage for *blood, urine, semen and saliva*. A protocol of protein precipitation by using chloroform/methanol/water[15] was carried out for each biofluids. After precipitation, protein concentration was determined by Bradford assay using BSA as standard[16]. Samples were submitted to reduction, alkylation and tryptic digestion. Samples were dissolved in denaturant buffer (urea 6 M, Tris 300 mM pH 8.0, EDTA 10 mM) containing DTT (10-fold molar excess on the Cys residues) at 37 °C for 2 h and then iodoacetamide (IAM) was added to perform carboamidomethylation using an excess of alkylating agent (5-fold molar excess on thiol residues). The mixture was then incubated in the dark at room temperature for 30 minutes. The product was purified by chloroform/methanol/water precipitation. Supernatants were removed, and the pellets were dried. Digestion of proteins mixture was carried out in AMBIC 10 mM using trypsin at a 50:1 protein:enzyme mass ratio. The samples were incubated at 37°C for 16 h and after acidification (10% HCOOH) they were dried. To eliminate any impurities the samples were suspended in 200 µL of AMBIC 100 mM, filtrated by centrifugal filter units (0.22 µm), and dried in a speed-vacuum concentrator. Finally, samples were suspended in 20 µL of HCOOH 1% and purified by reverse phase chromatography using ZipTip C18 cartridges (Millipore). Samples were evaporated and suspended in 10 µL of

HCOOH 0.1% and analysed by nanoLC-MS/MS and then by LC-MRM/MS as described below.

In order to improve the number of identified proteins from blood by LC-MS/MS, a step of depletion of abundant proteins was carried out using ProteoPrepR 20 Plasma Immunodepletion KitProteoprep® 20 Plasma Immunodepletion Kit (Sigma Aldrich, Milan, Italy) able to deplete twenty highly abundant proteins from human plasma or serum. Dried extracted proteins were subjected to reduction, carboamidomethylation, chloroform/methanol/water precipitation protocol and tryptic hydrolysis as described above.

Protein precipitation from urine was also performed using cold acetone followed by the same steps of tryptic digestion as described above. Cold acetone at -20°C (400 µL) was added to urine (100 µL) and incubated at -20°C for 2h. After centrifugation at 12,000 rpm for 15 min the pellet was collected and dried under vacuum.

Protein precipitation from semen was also carried out using TCA. Semen was centrifuged (13,000 rpm, 30 min) to remove cellular material, and the proteins were precipitated (4°C, 60 min) using an equal volume of cold TCA at final concentration of 10%. Following further centrifugation (10,000 rpm, 30 min), the supernatant was removed and the protein pellet washed with cold acetone (three times) to remove residual TCA.

Therefore, a unique precipitation/sample preparation procedure based on chloroform/methanol/water was adopted for all the samples because the origin or matrix of casework samples.

6.2.2 Test Specimens

Twenty test forensic samples were prepared by drying different biological fluids (100 µl) or mixture of them on various substrates: cloth, wood, plastic, plaster and paper. Cloths were put directly in a plastic tube while the samples on the other substrates were recovered by a cotton-swab (Agilent Technologies, Palo Alto, CA). Samples were collected in plastic tubes and 2 ml of AMBIC 50 mM were added for 16 h under gentle agitation. Samples were then sonicated for 20 min and submitted to chloroform/methanol/water precipitation. After

precipitation, reduction, alkylation and tryptic digestion were performed as previously described. The resulting peptide mixtures were analysed by LC-MRM/MS as described below.

6.2.3 Analysis of real samples

Real samples consisting of traces of unknown biological fluids on cloths, paper and stubs collected at a crime scene (Sample1, Sample 2, Sample 3) were provided by RIS Laboratory in Rome. Samples were treated as described above (see Test Specimens section) and submitted to tryptic digestion. Then, gDNA was extracted from the samples using the robotic platform Qiagen Biorobot EZ1 Advanced XL using the Qiagen EZ1 DNA Investigator kit[17]. The DNA was then washed and eluted in water. DNA degradation and quantification level were evaluated by multiplex RT-PCR using the Quantifiler™ Trio DNA Quantification Kit (Applied Biosystems, CA, USA) on a 7500 Real-Time PCR System (Applied Biosystems)[18]. Results showed a total DNA amount of 0.081ng/µl for Sample 1, 0.71 ng/µl for Sample 2 and 1.05 ng/µl for Sample 3. Samples were also tested for possible DNA degradation showing a good quality with a degradation index of about 1.0 for all samples.

DNA profiling was obtained by multiple amplification of 17 polymorphic regions comprising the D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, TH01, D16S539, D2S1338, D19S433, ACTBP2 (SE33), D1S1656, D2S441, D10S1248, D12S391, D22S1045 and Amelogenin STR loci using the Investigator ES Splex SE Plus Kit Qiagen (Qiagen, Hilden, Germany)[18]. Amplified DNA was analysed by automated DNA sequencing on an Applied Biosystems 3500xL Genetic Analyzer[19-21]. Data were collected and elaborated using the 3500 Series Data Collection Software v. 2.0and the Gene Mapper ID-X Software v.1.4 (Life Technologies, Carlsbad, CA, USA).

6.2.4 LC-MS/MS analysis

Peptide mixture were analyzed by LC-MS/MS on a 6520 Accurate-Mass Q-TOF LC/MS system (Agilent Technologies) equipped with a 1200 HPLC system and a chip cube (Agilent Technologies). After loading, the peptide

mixture (1 μL) was concentrated and desalted at flow rate of 4 μL/min in a 40 nL enrichment column (Agilent Technologies chip) with 0.1% HCOOH as eluent. The sample was then fractionated on a C18 reverse phase capillary column (75 μm*43 mm in the Agilent Technologies chip) at flow rate of 400 nL/min, with a linear gradient of eluent B (0.1% HCOOH in 95% acetonitrile) in A (0.1% HCOOH in 2% acetonitrile) from 5% to 80% in 50 min. Peptides analysis was performed using data-dependent acquisition of one MS scan (mass range *m/z* 300-2400) followed by MS/MS scan of the five most abundant ions in each MS scan. MS/MS spectra were measured automatically when the MS signal was greater than the threshold of 50000 counts. Doubly- and triply- charged ions were preferably isolated and fragmented over singly charged ions. Data were acquired through Mass Hunter software (Agilent Technologies). The acquired data, containing MS and MS/MS spectra, were transformed in .mgf format and used for protein identification with a licensed version of Mascot Software (www.matrixscience.com).

Mascot search parameters included: NCBInr as database; Trypsin as enzyme, allowed number of missed cleavage 3; *Homo-Sapiens* as taxonomy; carbamidomethyl, C as fixed modifications; oxidation of methionine (oxidation (M));, *Gln pyro-Glu* (N-term Q) as variable modifications; 10 ppm MS tolerance, 0.6 Da MS/MS tolerance and peptide charge, from +2 to +3.

6.2.5 MRM targeted proteomic approach

In order to build up a targeted MRM method, Skyline software (3.7, 64 bit version MacCoss Lab Software, University of Washington, USA) was used for the *in silico* selection of peptides with unique sequence for each selected protein. For each peptide, *m/z* precursor ion, *m/z* product ions and relative collision energy were provided by Skyline. Peptide mixture was analysed by LC-MS/MS analysis using a Xevo TQ-S (Waters) equipped with an IonKey UPLC Microflow Source coupled to an UPLC Acquity System (Waters). For each run, 1 μL peptide mixture was injected and separated on a TS3 1.0 mm × 150 mm analytical RP column (Waters, Milford, MA, USA) at 45°C with flow rate of 3 μl/min using 0.1% HCOOH in water (LC-MS grade) as eluent A and 0.1%

HCOOH in ACN as eluent B. Peptides were eluted (starting 1 min after injection) with a linear gradient of eluent B in A from 7% to 95% in 55 min. The column was re-equilibrated at initial conditions for 4 min. The MRM mass spectrometric analyses were performed in positive ion mode using MRM detection window of 0.5-1.6 min per peptide; the duty cycle was set to automatic and dwell times were minimal 5 ms. Cone voltage was set to 35 V.

6.3 Results

6.3.1 Selection of protein biomarkers of biological matrices

Four biological matrices, blood, saliva, urine and semen, representative of biofluids recovered at crime scenes, were analysed to identify specific protein biomarkers candidates for each matrix. Different procedures for sample preparation were investigated (*Material and Method*s section). As an example, results obtained for blood with the only chloroform/methanol/water precipitation followed by tryptic hydrolysis, are shown in Table 1A where GI number, the name of identified proteins, the number of peptides and the corresponding sequence coverage are reported.

The unsatisfactory number of identified proteins in the original assay severely limited the choice of specific protein biomarkers for blood candidates for the MRM method because the dominant abundance of unselective blood proteins. Thus, a different sample preparation procedure was then explored by a depletion step from the most abundant proteins to improve protein identification (Table 1B). A significantly higher number of peptides were identified for each protein leading to a much larger sequence coverage.

Table 1A. List of the identified proteins following the chloroform/methanol/water precipitation protocol applied to human blood sample. **Table 1B.** List of the identified proteins following the depletion step previous to the chloroform/methanol/water precipitation protocol. The Gi number, the number of peptides, the corresponding Mascot score and sequence coverage are reported for human blood sample.

Table 1A				
GI number	**Identification**	**MascotScore**	**Number of Peptides**	**Sequence Coverage (%)**
Gi|3212456	Chain A, Crystal Structure Of Human Serum Albumin	1186	35	24
Gi|90108664	Chain A, Crystal Structure Of Lipid-Free	602	18	22

Gi number	Identification	MascotScore	Number of Peptides	Sequence Coverage (%)
	Human Apolipoprotein A-I			
Gi|2765421	Immunoglobulin Kappa Heavy Chain	192	4	34
Gi|28637	Alpha-1 Antitrypsin	188	5	5
Gi|386789	Hemopexin Precursor	178	7	10
Gi|28810	Beta-2-Glycoprotein Apolipoprotein H	160	6	1
Gi|4557871	Serotransferrin Precursor	146	6	5
Gi|38026	Zn-Alpha2-Glycoprotein	75	2	5
Gi|112910	Alpha-2-HS-Glycoprotein	71	2	4
Table 1B				
Gi number	**Identification**	**MascotScore**	**Number of Peptides**	**Sequence Coverage (%)**
Gi|1431650	Chain β Of Hemoglobin	156	6	20
Gi|229751	Chain α of Hemoglobin	138	4	15
Gi|3212456	Chain A, Crystal Structure Of Human Serum Albumin	1995	48	34
Gi|194383506	Serum Transferrin	1012	34	29
Gi|90108664	Chain A, Crystal Structure Of Lipid-Free Human Apolipoprotein A-I	562	22	27
Gi|177870	Alpha-2-Macroglobulin	485	17	7
Gi|13787109	Alpha-1-antitrypsin	337	12	14

Improved results were also obtained in saliva, semen, and urine analysis by using the chloroform/methanol/water precipitation protocol. Table S-1 in Supplementary information (SI) reports the lists of the different proteins identified for the other matrices, including the GI number, the number of peptides and the Mascot score.

Proteins identified for each biofluids by LC-MS/MS were then compared with literature data[11] to eventually select the most specific protein biomarkers for each biological matrix. Following this procedure, proteins listed in Table S-2 were chosen as representative for each biological fluid.

6.3.2 Development of MRM methods specific for each biological matrix and a single MRM method for all the matrices

Several unique peptides belonging to the target proteins characteristic of individual biological matrix were selected by *in silico* analysis using the Skyline software that provided the predicted best transitions and collision energy to generate maximal fragment intensities. The *in silico* data of the selected peptides were then compared with the experimental fragmentation spectra obtained by the LC-MS/MS analyses. Peptides defined by Skyline analysis and showing the best signal to noise ratio in the experimental fragmentation spectra were selected for developing MRM methods for each biological fluid. These methods contained all the precursor ion- product ions transitions and collision energy associated with the selected peptides from the defined target proteins constituting the protein signature of each individual biofluids in Table S-2.

As an example, two peptides from α-globin, three from β-globin, and Hemopexin, five from Haptoglobin and six from α-2-macroglobulinfor a total of 25 peptides and 106 transitions were selected to monitor the presence of blood and to build the MRM method for this biological matrix.

The Total Ion Current chromatogram (A) and the MRM TIC analysis of the three peptides 132-143, 82-94 and 30-39 from β-globin (B) are shown in Figure 1. Panel C displays the MRM TIC chromatogram of the three transitions, *m/z* 409.72 to 507.29, *m/z* 409.72 to 604.34 and *m/z* 409.72 to 719.37, used to monitor the α-globin (92-98) peptide. The different transitions perfectly co-eluted at a retention time of 5.24 min thus indicating that they belong to the same precursor ion.

Figure 16: Identification of all the protein of blood fluid by LC-MS/MS analysis and MRM/MS data of the defined target proteins constituting the protein signature of blood. **Panel A**: Total Ion Current (TIC) as a function of time (min.). LCMSMS analysis allowed the identification of protein content on blood tryptic peptide mixture. **Panel B**: BPI Chromatograms for β-globin 132-143, 82-94, 30-39 peptides eluted at 6.18, 11.46 and 16.91

min respectively. **Panel C**: MRM TIC Chromatograms extracted for α-Globin (92-98) peptide. The monitored transitions for this peptide: *m/z* 409.72 to 719.37, *m/z* 409.72 to 604.34 and *m/z* 409.72 to 507.29 are perfectly co-eluted at 5.24 min as an unambiguous identification of blood detection.

Similar MRM methods were developed for the other biological matrices (Table S-2). For urine, as an example, Figure S-1 shows the Total Ion Current chromatogram (A) and the MRM TIC analysis of the two peptides (507-605) and (204-211) from Uromodulin (B). Panel C displays the TIC chromatogram of the three transitions for the Uromodulin (204-211) peptide. The different transitions perfectly co-eluted at a retention time of 2.99 min thus indicating that they all originated from the same precursor ion.

Some of the MRM TIC chromatograms of the selected peptides for the specific detection of proteins from saliva and semen were reported in SI (Figure S-2, FigureS-3).

Once the optimized MRM methods were developed for each individual biological fluid, the next step involved the development of a single MRM method able to detect all the biological fluids in a single analysis. For each selected protein biomarker, all the selected peptide sequences, *m/z* precursor ion, *m/z* product ions and the optimized collision energy were tabulated in Table S-3.

As the mass spectrometer was able to handle a very large number of transitions *per* run, the single MRM method was built up by using a total of 46 peptides and 212 transitions.

6.3.3 Analysis of test specimens

For all the *test specimens* samples, the MRM chromatograms solely showed the mass transitions associated to the target peptides belonging to the specific proteins constituting the unique signature of the defined matrix. No transitions related to peptides from proteins characteristic of other biological fluids were detected for all the analysed samples. As an example, Figure 2A shows the MRM TIC chromatogram of the three transitions, *m/z* 637.86 to 687.35, *m/z* 637.86 to 850.42, *m/z* 637.86 to 949.49, *m/z* 637.86 to 1048.56 and

m/z 637.86 to 1161.64 used to monitor the *β*-globin (30-39) peptide form a blood sample spotted on plaster. Panel B showed the corresponding MRM chromatogram for the (1075-1086) peptide from Mucin-6, a protein specific of semen, illustrating the absence of any non-specific transition of interfering proteins. This result clearly demonstrates that the specific signature of blood could be easily monitored whereas no interference from cross transitions of peptides belonging to other biological matrices was detected.

Figure 2. MRM/MS analysis of test specimens: blood spotted on plaster. **Panel A**: MRM TIC Chromatograms extracted for β-Globin (30-39) peptide. The best five monitored transitions *m/z* 637.87 to 687.36 *m/z*, 637.87 to 850.42, *m/z* 637.87 to 949.49, *m/z* 637.87 to 1048.56, *m/z* 637.87 to 1161.64 are reported in figure eluted at 16.73 min. **Panel B**: MRM TIC chromatogram extracted for (1075-1086) peptide from Mucin-6 a protein specific of semen, illustrating the absence of any non-specific transition (*m/z* 785.38 to 960.42, *m/z* 785.38 to 1057.48, *m/z* 785.38 to 1307.62).

Similar results were obtained on all the other spotted samples and are summarized in Table 2. Each biological matrix could be unambiguously identified by the developed MRM method even on different substrates. No cross transitions were observed whatsoever with the exception of urine where peptides

from PSA and PAP proteins, a specific signature of semen, was also detected. However, the presence of seminal proteins in urine is somehow expected and provides further information on the origin of the sample that was obviously from a male donor.

Table 2. List of proteins detected in each test specimen prepared by drying different biological fluids on various substrates: cloth, wood, plastic, plaster and paper. matrix spotted on substrates.

Matrixes	Substrates	BLOOD				URINE		SALIVA				SEMEN				
		αHemoglobin	βHemoglobin	Hemopexin	Haptoglobin	Uromodulin	Osteopontin	Alpha amylase	Muc 5 Ac	Annexin A1	Custatin S	Prostate Specific Antigen	Prostatic Acid Phosphatase	Muc 6	Semenogelin I	Semenogelin II
Blood	Cloth	X	X	X	X											
	Wood	X	X	X	X											
	Plastic	X	X	X	X											
	Plaster	X	X	X	X											
	Paper	X	X	X	X											
Urine	Wood					X	X					X	X			
	Plastic					X	X					X	X			
Saliva	Cloth							X	X	X	X					
	Wood							X	X	X	X					
	Plastic							X	X	X	X					
	Plaster							X	X	X	X					
Semen	Cloth											X	X	X	X	
	Wood											X	X	X		
	Plastic											X	X	X	X	

X is reported when the unique peptides belonging to the target proteins characteristic of individual biological matrix were identified by MRM analysis.

To test the MRM optimized method capability to unambiguously define the occurrence of individual biofluids when occurring in a 50:50 ratio mixture, blood-semen, blood-urine and semen-urine were spotted onto different substrates. The negative controls, for a total of 5 sample, consisted of the exclusive use of substrates: cloth, wood, plastic, plaster and paper. As shown in Table S-4, each individual component of the binary mixture was unambiguously identified by the specific transitions of the unique peptide signature; whenever urine was present, specific peptides from semen proteins were also detected.

6.3.4 Analysis of real crime scene samples

The developed procedure was then applied to real samples collected at a crime scene by the RIS Laboratory in Rome. Samples were prepared according to the procedure described above and following enzymatic hydrolysis, DNA was extracted from each sample according to the normal procedure performed at the RIS Laboratory. DNA extracts were quantified, the degradation level was evaluated and the DNA profiling obtained as described in Materials and Methods section. The resulting genotypic profile of Sample 1 is shown in Figure S-4 and was identified as a male genotype profile called "Profile A". The quality of the data exceeded the minimum quality requirements established by internal validation procedures for this Lab for comparative procedures to be used for personal identification purposes.

Following DNA analysis, the remainder of the samples containing the mixture of tryptic peptides was submitted to the developed MRM analytical procedure. Figure 3 shows the TIC analysis of the three peptides 30-39, 82-94, and 132-143 monitored for the human β-globin and two peptides 92-98 and 127-138 monitored for the human α-globin (A). Panel B displays the MRM TIC chromatogram corresponding to four transitions, *m/z* 637.87 to 687.35, *m/z* 637.87 to 850.42, *m/z* 637.87 to 949.48 and *m/z* 637.87 to 1048.55, used to monitor the β-globin (30-39) peptide.

Figure 3. MRM/MS analysis of real crime scene samples after DNA extraction. **Panel A**: BPI Chromatograms for α-Globin 92-98 and 127-138 peptides eluted at 5.26 and 14.91 min respectively and BPI Chromatograms for β-Globin 132-143, 82-94, 30-39 peptides eluted at 11.40, 11.50, 16.62 min respectively. **Panel B**: MRM TIC Chromatogram corresponding to four transitions, *m/z* 637.87 to 687.35, *m/z* 637.87 to 850.42, *m/z* 637.87 to 949.48 and *m/z*

637.87 to 1048.55 perfectly co-eluted at 16.59 min, used to monitor the β-globin (30-39) peptide.

As shown in Table 3, the MRM analysis was able to unambiguously define individual protein component occurring in the different samples by their unique peptide signature even after DNA extraction and analysis. Therefore, this combined procedure led to both DNA profiling analysis and the unambiguous identification of blood indicating the biological fluid the DNA was collected from.

Table 3. Identification of the biomarker proteins by MRM analysis in three samples got from RIS laboratories after DNA extraction and analysis.

		Peptide	Sample 1	Sample 2	Sample 3
BLOOD	α Hb	K.VGAHAGEYGAEALER.M [16, 30]	x		
		R.VDPVNFK.L [92, 98]	x		
		K.FLASVSTVLTSK.Y[127, 138]	x		x
	β Hb	R.LLVVYPWTQR.F [30, 39]	x		x
		K.GTFATLSELHCDK.L [82, 94]	x		
		K.VVAGVANALAHK.Y [132, 143]	x		
	Hemopexin	R.ELISER.W [82, 87]	x		
		K.VDGALCMEK.S [401, 409]	x		
	Haptoglobin	K.DIAPTLTLYVGK.K [156, 167]	x		
		R.VMPICLPSK.D [202, 210]	x		
SALIVA	α Amylase	R.WVDIALECER.Y [35, 44]			x
		K.SSDYFGNGR.V [258, 266]			x
	Annexin 1	K.GVDEATIIDILTK.R [58, 70]			x
		R.SEIDMNDIK.A [303, 311]			x

136

	MUC 5AC	K.ILVALCGGN.- [337, 345]			x
		K.GVQLSDWR.D [1335, 1342]			x
		R.AQAQPGVPLR.E [2999, 3008]			x
	Cystatin S	K.QLCSFEIYEVPWEDR.M [115, 129]			
		R.MSLVNSR.C [130, 136]			x
SEMEN	MUC 6	K.VTNEFVSEEGK.F [182, 192]		x	
		R.ETDPCSMSQLNK.V [573, 584]		x	
		R.GVLLWGWR.S [646, 653]		x	
		K.VYHLPYYEACVR.D [1075, 1086]		x	
	SEM II	K.GHYQNVVDVR.E [217, 226]		x	
		K.DIFTTQDELLVYNK.N [251, 264]			
		K.ISYQSSSTEER.H [345, 355]		x	
		K.QDLLSHEQK.G [534, 542]		x	
	PSA	R.IVGGWECEK.H [20, 28]		x	
		K.HSQPWQVLVASR.G [29, 40]		x	
		R.LSEPAELTDAVK.V [121, 132]		x	
		K.FMLCAGR.W [190, 196]		x	
	PAP	R.SPIDTFPTDPIK.E [47, 58]		x	
		K.DFIATLGK.L [185, 192]		x	
		R.ELSELSLLSLYGIHK.Q [236, 250]		x	
	SEM 1	R.LWVHGLSK.E [165, 172]		x	
		K.VQTSLCPAHQDK.L [233, 244]		x	
		K.DVSQSSIYSQTEEK.A [307, 320]			

		K.GESGQSTNR.E [404, 412]		x	
URINE	Osteopontin	K.QNLLAPQTLPSK.S [51, 62]			
		R.GDSVVYGLR.S [145, 153]			
		R.ISHELDSASSEV.- [287, 298]			
	Uromodulin	R.STEYGEGYACDTDLR.G [185, 199]			
		K.VFMYLSDSR.C [356, 364]			
		K.INFACSYPLDMK.V [420, 431]			
		R.VGGTGMFTVR.M [449, 458]			
		R.VLNLGPITR.K [597, 605]			

X is reported when the unique peptides belonging to the target proteins characteristic of individual biological matrix were identified by MRM analysis**.**

6.5 Discussion

Knowledge of the nature of biofluid discovered at a crime scene can influence the outcome of a case. Identification of a suspect's DNA on a victim is quite different if it comes from saliva, suggesting intimate contact, wither consensual or forced, or from blood or seminal fluid, indicating a physical struggle or a sexual assault. However, this is not always an easy task, since many biofluid stains are either invisible to the naked eye or similar in appearance to other fluids or substances. Even when the identity of a stain may seem obvious to a forensic investigator, absolute confirmation is necessary to give a legal value to the trace evidence-to either prove or disprove a fact in a ~~case~~ lawsuit. This is especially important with the possible occurrence of mixtures as single stain could contain multiple biofluids.

Most biochemical and immunologic tests which have been used for presumptive or confirmatory methods suffer from several limitations leading to sample destruction resulting in sample loss for subsequent DNA analysis or to

incompatibility with downstream individual identification assays, i.e. DNA profiling. Moreover, most of the current methods are designed to detect a single biofluid forcing the investigators to decide which test to perform in the presence of a limited amount of sample.

Contrary to DNA, proteins were rarely considered and scarcely used as sources of useful biological traces in crime scene investigations. Proteins tend to be less stable than DNA and are easily degraded making their identification by immunological methods unfeasible. Moreover, the amount of available samples cannot be amplified by "PCR–like" procedures thus requiring the use of analytical techniques with extremely high sensitivity[22].

More recently, proteomics strategies have been applied to the identification of biofluids that overcome most of the previous difficulties. In proteomic approaches, proteins are identified by their peptide fragments making the degradation no longer a problem. Moreover, contrary to DNA, proteins are tissue-specific providing a unique signature to identify biological tissues and fluids. Finally, the tiny amount of trace evidence usually recovered at a crime scene is matched by the extraordinary sensitivity of modern mass spectrometers. Untargeted proteomic approaches demonstrated to be effective in the definition of biofluids in traces from crime scenes by identification of specific proteins belonging to individual biological sample[5]. However, severe limitations in sensitivity, in the analysis of mixed traces and, more important, in preserving the samples for subsequent DNA analysis still exist. A recent study has greatly overcome this drawback by MALD IMS imaging directly on a blood fingermark leaving the ridge detail completely preserved[10]. Even though the proposed method is as destructive as the other approaches, the great compatibility between DNA extraction and MS strategy protocols allows to limit the destruction of the entire trace evidence. Indeed, a combined strategy to provide both DNA analysis profiling by RIS laboratory and identification of the corresponding biological fluid(s) on the same sample collected at real crime scenes has been developed. Samples were deproteinised by treatment with trypsin instead of the commonly used non-specific proteases, i.e. pronase, proteinase K, etc. DNA was then extracted from the digested samples and analysed according to the usual

procedure, yielding optimal DNA profiles. Moreover, the proposed *bottom-up* approach makes this method valid also on old traces that have undergone degradation. Another advantage from the current approach is the unambiguous identification of all target proteins allowing us to univocally discriminate among the different fluids, even in mixture by a single MRM run of less than one hour.

Indeed, four proteins, α-globin, β-globin, Haptoglobin and Hemopexin have been selected for blood signature, four for saliva (α-amylase 1, Mucin-5, Annexin A1 and Cystatin S), five for semen (Semenogelin I and II, PSA, PAP and Mucin-6) and two proteins determined for urine (Uromodulin and Osteopontin). For each protein biomarker, a number of unique tryptic peptides (proteotypics) has been selected with the use of bioinformatics tools and the best precursor ion-fragments transitions employed to set up a single MRM method able to identify any of the four biofluids in unknown stains. The MRM LC-MS/MS analysis recognizes the specific peptides by their unique transitions. The identified proteotypic peptides have been correlated to each set of selected protein biomarkers that in turn unambiguously defined the biological fluid(s) under investigation.

The MRM method optimized on pure samples of the four matrices, blood, urine, semen and saliva, is resulted to have good specificity and selectivity for each biofluid with no cross-contamination observed whatsoever. A single contamination was detected in the urine samples spotted on different surfaces that showed the occurrence of small amount of seminal fluid. However, this presence was somehow expected and contributed to indicate a male donor. Then in a proof of concept, prepared test samples consisting of different fluids spotted on various substrates have been analysed by using the single optimized MRM method.

The current results demonstrated that the developed strategy based on the MRM LC-MS/MS method could be a useful substrate in forensic science due to its capability of providing a universal approach for the identification of unknown stains recovered at crime scenes with high selectivity and specificity. Moreover, this approach can be carried out on the same samples used for DNA profiling revealing the nature of the tissue or fluid the DNA had been recovered from.

6.6 References

(1) Roewer, L. *Investigative genetics* **2013**, *4*, 22.

(2) Tobe, S. S.; Watson, N.; Daeid, N. N. *Journal of forensic sciences* **2007**, *52*, 102-109.

(3) Auvdel, M. J. *Journal of Forensic Science* **1987**, *32*, 326-345.

(4) Webb, J. L.; Creamer, J. I.; Quickenden, T. I. *Luminescence* **2006**, *21*, 214-220.

(5) Virkler, K.; Lednev, I. K. *Forensic Science International* **2009**, *188*, 1-17.

(6) Cocks, J.; du Toit-Prinsloo, L.; Steffens, F.; Saayman, G. *Forensic science international* **2015**, *249*, 225-232.

(7) Horjan, I.; Barbaric, L.; Mrsic, G. *Journal of forensic and legal medicine* **2016**, *38*, 101-105.

(8) Muro, C. K.; Doty, K. C.; de Souza Fernandes, L.; Lednev, I. K. *Forensic Chemistry* **2016**, *1*, 31-38.

(9) Muehlethaler, C.; Leona, M.; Lombardi, J. R. *Analytical chemistry* **2015**, *88*, 152-169.

(10) Pang, B.; Cheung, B. K. *Journal of forensic sciences* **2008**, *53*, 1117-1122.

(11) Van Steendam, K.; De Ceuleneer, M.; Dhaenens, M.; Van Hoofstat, D.; Deforce, D. *International journal of legal medicine* **2013**, *127*, 287-298.

(12) Legg, K. M.; Powell, R.; Reisdorph, N.; Reisdorph, R.; Danielson, P. B. *Electrophoresis* **2017**, *38*, 833-845.

(13) Deininger, L.; Patel, E.; Clench, M. R.; Sears, V.; Sammon, C.; Francese, S. *Proteomics* **2016**, *16*, 1707-1717.

(14) Legg, K. M.; Powell, R.; Reisdorph, N.; Reisdorph, R.; Danielson, P. B. *Electrophoresis* **2014**, *35*, 3069-3078.

(15) Wessel, D.; Flügge, U. *Analytical biochemistry* **1984**, *138*, 141-143.

(16) Bradford, M. M. *Analytical biochemistry* **1976**, *72*, 248-254.

(17) Anslinger, K.; Bayer, B.; Rolf, B.; Keil, W.; Eisenmenger, W. *Legal Medicine* **2005**, *7*, 164-168.

(18) Liu, J. Y. *Forensic Science International: Genetics* **2014**, *13*, 10-19.

(19) Pascali, J. P.; Bortolotti, F.; Tagliaro, F. *Electrophoresis* **2012**, *33*, 117-126.

(20) Moreno, L. I.; McCord, B. *Handbook of Capillary and Microchip Electrophoresis and Associated Microtechniques, third ed. CRC Press, Boca Raton, USA* **2007**, 733-756.

(21) Landers, J. P. *Handbook of capillary and microchip electrophoresis and associated microtechniques*; CRC press, 2007.

(22) Sethi, S.; Brietzke, E. *International Journal of Neuropsychopharmacology* **2016**, *19*.

6.7 Supplementary Data

Table S-1. Identified Proteins By Different Pretreatment of each biological matrix with relative Gi Number, Molecular weight, Number of peptides and Mascot Score. Chloroform/Methanol/Water precipitation protocol showed the best results and it was selected for building up the experimental strategy to apply for real samples treatment.

Urine

Code	Identification	Mascot Score	Numbero of Peptides
Gi|3212456	Chain A, Crystal Structure Of Human Serum Albumin	832	29
Gi|340166	Uromodulin	542	13
Gi|157835127	Chain A, Crystal Structure Of Human Apolipoprotein D (Apod) In Complex With Progesterone	223	8
Gi|4502067	Protein Ambp Preproprotein	142	5
Gi|480312171	Chain H, Crossfab Binding To Human Angiopoietin 2	140	2
Gi|229585	Ig A1 Bur	132	4
Gi|238236	Transmembrane Secretory Component	121	4
Gi|4504893	Kininogen-1 Isoform 2 Precursor	118	5
Gi|29470	Human Basement Membrane Heparan Sulfate Proteoglycan Core Protein	118	3
Gi|178585	Alpha-Amylase	111	2
Gi|4557871	Serotransferrin Precursor	92	3
Gi|3297879	Masp-2	84	2
Gi|256397	Urine Protein 1, Up1=Clara Cell 10 Kda Protein Homolog Ù {N-Terminal} [Human, Urine, Peptide Partial, 53 Aa]	83	3
Gi|4759166	Osteopontin Isoform Opn-B Precursor	81	4
Gi|532598	Ig J-Chain, Partial	78	2
Gi|307151	Mac25	75	2

Gi|15489339	Vasn Protein, Partial	70	2
Gi|5803023	Vesicular Integral-Membrane Protein Vip36 Precursor	69	1
Gi|180550	Plasma Serine Protease Inhibitor Precursor	67	1
Gi|62738629	Chain A, Crystal Structure Of Human Peptidoglycan Recognition Protein (Pgrp-S)	64	2
Gi|261745	1.14 Beta-Trace=23.5 Kda Glycoprotein {N-Terminal} [Human, Cerebrospinal Fluid, Peptide Partial, 19 Aa]	57	1
Gi|31075	E-Cadherin	48	1
Gi|292870	Tyrosine Kinase Receptor	45	1
Gi|6563042	Leukocyte-Associated Ig-Like Receptor 1b	43	1
Gi|4502719	Cadherin-13 Isoform 1 Preproprotein	40	1
Gi|310942621	Chain D, The Crystal Structure Of Pcsk9 In Complex With 1d05 Fab	40	1
Gi|189524	Prostate Specific Antigen Precursor, Partial	40	1
Gi|186651	Kallikrein	36	1
Gi|179617	C1-Inhibitor, Partial	31	1

Saliva

Code	Identification	Mascot Score	Number of Peptides
Gi|131412225	Keratin, Type I Cytoskeletal 13	1020	21
Gi|114644568	Keratin, Type II Cytoskeletal 6A	818	20
Gi|82654947	Keratin, Type II Cytoskeletal 4	779	21
Gi|178585	Alpha-Amylase	458	15
Gi|4502101	Annexin A1	415	10
Gi|375314779	Keratin 1	397	5
Gi|157057089	Mucin 5AC, Oligomeric Mucus/Gel-Forming	247	7
Gi|238236	Transmembrane Secretory Component	201	6
Gi|28336	Mutant Beta-Actin (Beta'-Actin)	158	5
Gi|229585	Ig A1 Bur	132	5
Gi|21669481	Immunoglobulin Kappa Light Chain VLJ Region	126	2
Gi|240102	Cystatin SN	116	2
Gi|609412460	Chain A, Crystal Structure Of Human Lnkh2b-H2a.Z-Anp32e	115	2
Gi|17942573	Chain A, G48a Human Lysozyme	85	2
Gi|7706635	Cornulin	75	3
Gi|662841	Heat Shock Protein 27	68	2
Gi|359513	Cystatin SA	123	4
Gi|4503109	Cystatin-S Precursor	68	1
Gi|38026	Zn-Alpha2-Glycoprotein	64	2
Gi|186833	Lactoferrin	81	5

Semen

Code	Identification	Mascot Score	Numbero of Peptides
Gi|122920512	Chain A, Human Serum Albumin Complexed With Myristate And Aspirin	1581	39
Gi|187122	Lactoferrin	1367	30
Gi|578840955	PREDICTED: Mucin-6 Isoform X1	1069	21
Gi|4506883	Semenogelin-1 Preproprotein	1048	26
Gi|6382064	Prostatic Acid Phosphatase Isoform PAP Precursor	860	23
Gi|31397	Fibronectin Precursor	946	25
Gi|38026	Zn-Alpha2-Glycoprotein	855	18
Gi|62089422	Membrane Alanine Aminopeptidase Precursor Variant	741	18
Gi|189524	Prostate Specific Antigen Precursor, Partial	613	12
Gi|178855	Apolipoprotein J Precursor	613	14
Gi|38648667	Fatty Acid Synthase	546	15
Gi|83699649	Heat Shock 90kda Protein 1, Alpha	494	10
Gi|180550	Plasma Serine Protease Inhibitor Precursor	427	10
Gi|4505821	Prolactin-Inducible Protein Precursor	418	9
Gi|219517850	Membrane Metallo-Endopeptidase	410	9
Gi|180570	Creatine Kinase	376	8
Gi|4503009	Carboxypeptidase E Preproprotein	346	10
Gi|177836	Alpha-1-Antitrypsin Precursor	336	9
Gi|5031863	Galectin-3-Binding Protein Precursor	329	8
Gi|1353350	Prostate-Specific Transglutaminase	313	8
Gi|337760	Cerebroside Sulfate Activator Protein	295	9
Gi|4505763	Phosphoglycerate Kinase 1	294	6
Gi|4507953	14-3-3 Protein Zeta/Delta	287	7
Gi|183129	Gamma-Glutamyltransferase 1 (EC 2.3.2.2)	284	5
Gi|4557305	Fructose-Bisphosphate Aldolase A Isoform 1	281	5
Gi|3641398	NADP-Dependent Isocitrate Dehydrogenase	272	8

Gi|4505591	Peroxiredoxin-1	258	7
Gi|693933	2-Phosphopyruvate-Hydratase Alpha-Enolase	239	5
Gi|386785	Heat Shock Protein	269	7
Gi|1488324	Extracellular Matrix Protein 1	238	7
Gi|4758638	Peroxiredoxin-6	238	6
Gi|226529917	Triosephosphate Isomerase Isoform 2	230	4
Gi|15620780	Glutamate Carboxypeptidase	223	8
Gi|4826870	Nucleobindin-2 Precursor	223	7
Gi|4826643	Annexin A3	215	4
Gi|496078	L-Iditol-2 Dehydrogenase	198	5
Gi|1246085	Acidic Epididymal Glycoprotein Homolog	191	5
Gi|4759000	Ras-Related Protein Rab-3D	189	4
Gi|4557727	Lipoprotein Lipase Precursor	186	5
Gi|4503273	Angiotensin-Converting Enzyme Isoform 1 Precursor	184	4
Gi|5174735	Tubulin Beta-4B Chain	186	6
Gi|190215	Placental Protein 14	177	4
Gi|4502101	Annexin A1	153	4
Gi|181387	Cystatin C	148	3
Gi|31645	Glyceraldehyde-3-Phosphate Dehydrogenase	148	4
Gi|235948	Cystatin SA-III=Potential Precursor Of Acquired Enamel Pellicle	146	4
Gi|189998	M2-Type Pyruvate Kinase	145	3
Gi|1931577	Rab27a	143	4
Gi|662841	Heat Shock Protein 27	138	4
Gi|558379	FALL-39 Peptide Antibiotic	138	4
Gi|9857661	Cell Surface Receptor	134	3
Gi|181250	Cyclophilin, Partial	132	3
Gi|31615558	Chain A, X-Ray Crystal Structure Of Cyclophilin AHIV-1 Ca N- Terminal Domain (1-146) M-Type H87a Complex.	130	4
Gi|37492	Alpha-Tubulin	128	3
Gi|4826898	Profilin-1	121	3
Gi|19923082	Prostate And Testis Expressed Protein 1 Precursor	121	3

Gi|45594240	Limbic System-Associated Membrane Protein Preproprotein	111	4
Gi|5453678	Epididymal Secretory Protein E1 Precursor	106	3
Gi|62089124	Angiotensinogen Precursor Variant	89	3
Gi|291191130	Chain A, Crystal Structure Of Human Seminal Plasma Protein Psp94	79	2
Gi|5031857	L-Lactate Dehydrogenase A Chain Isoform 1	78	2
Gi|36038	Rho GDP Dissociation Inhibitor (GDI)	78	2
Gi|307110	Lysosomal Membrane Glycoprotein-2	75	2
Gi|35830	Ubiquitin Activating Enzyme E1	75	2
Gi|386789	Hemopexin Precursor, Partial	73	3
Gi|3777541	Major Sperm Fibrous Sheath Protein Precursor	72	2
Gi|31075	E-Cadherin	67	2
Gi|5901956	Follistatin-Related Protein 1 Precursor	59	2

Table S-2. List of proteins selected as representative for each biological matrix by matching LC-MS/MS analysis and literature data.

Matrix	Protein Biomarkers
Blood	*Chain A of Hemoglobin*
	Chain B of Hemoglobin
	Hemopexin
	Haptoglobin
	Complement C3
	Alfa-1-Antitrypsin
	Alfa-2-Macroglobulin
	Alpha-2-HS-Glycoprotein
Saliva	*A - Amylase*
	Annexin
	Mucin 5 Ac
	Cystatin S
	Transmembrane Secretory Complement
	Mutant Beta Actin
Urine	*Osteopontin*
	Uromodulin
	AMBP Protein
	Kininogen
Semen	*Epididymal Secretory Protein E1*
	Mucin-6
	Semenogelin I
	Semenogelin II
	Prostate Specific Antigen (PSA)
	Galectin-3-Binding Protein
	Prostatic Acid Phosphatase (PAP)
	Prostate- Specific Transglutaminase
	Prostate And Testis Expressed Protein 1
	Beta-Microseminoprotein/ Crystal Structure Of Human Seminal Plasma Protein Psp 94

Table S-3. MRM method optimized for each biological fluid reporting protein, selected peptide sequence, precursor m/z, product ions m/z and Collision Energy (V).

BLOOD

	Peptide	Precursor m/z	Products m/z	Collision Energy (V)
Chain α of Hemoglobin	R.VDPVNFK.L [92, 98]	409,72	719,37 604,34 507,29	14
	K.FLASVSTVLTSK.Y [127, 138]	626,86	992,56 921,52 834,49 735,42 648,39	22
Chain β of Hemoglobin	R.LLVVYPWTQR.F [30, 39]	637,86	1161,64 1048,55 949,48 850,42 687,35	23
	K.GTFATLSELH**C**DK.L [82, 94]	739,85	1173,55 1102,51 1001,47 888,38 801,35	26
	K.VVAGVANALAHK.Y [132, 143]	575,34	951,53 880,50 823,47 724,41 653,37	20
Alpha 2-glycoprotein	K.EHAVEGD**C**DFQLLK.L [106, 119]	830,88	1323,62 1224,55 1095,51 1038,49 923,46	30

	K.FSVVYAK.C [124, 130]	407,22	666,3821 579,3501 480,2817	14
	K.**C**NLLAEK.Q [218, 224]	424,22	687,4036 573,3606 460,2766	15
Hemopexin	R.ELISER.W [82, 87]	373,70	617,36 504,27 391,19	13
	R.QGHNSVFLIK.G [101, 110]	571,81	1014,57 957,55 820,49 706,44 619,41	20
	K.VDGAL**C**MEK.S [401, 409]	511,73	923,39 808,36 751,34 680,31 567,22	18
Haptoglobin	K.AVGDK.L [77, 81]	245,13	418,22 319,16 262,13	8
	K.DIAPTLTLYVGK.K [156, 167]	645,86	1062,61 991,58 894,52 793,48 680,39	23
	R.VMPI**C**LPSK.D [202, 210]	522,78	945,48 814,44 717,39 604,31	18
	R.NANFK.F [227, 231]	297,15	479,26 408,22	10
	K.YVMLPVADQDQ**C**IR. H [238, 251]	854,41	1314,64 1201,56 1104,51 1005,44 934,40	31
Alpha-	K.QINDYVEK.G	504,75	880,44 767,35	18

	[155, 162]		653,31 538,28	
	R.DTVFALVNYIFFK.G [178, 190]	788,92	1261,69 1114,62 1043,59 930,50 831,44	28
	R.SASLHLPK.L [282, 289]	426,75	765,46 694,42 607,39 494,30	15
Alpha-2-Macroglobulin	K.NEDSLVFVQTDK.S [123, 134]	697,84	1151,59 1036,56 949,53 836,45 737,38	25
	K.FEVQVTVPK.I [228, 236]	523,79	899,51 770,47 671,40 543,35	18
	K.GVPIPNK.V [375, 381]	362,72	667,41 568,34 471,29	12
	K.GHFSISIPVK.S [521, 530]	542,81	1027,59 890,53 743,46 656,43 543,35	19
	K.AIGYLNTGYQR.Q [1003, 1013]	628,32	1184,60 1071,52 1014,50 851,43 738,35	22
	K.MVSGFIPLKPTVK.M [1384, 1396]	708,91	1186,71 1099,68 1042,66 895,59 782,51	25

URINE

	Peptide	Precursor m/z	Products m/z	Collision Energy (V)
Osteopontin	K.QNLLAPQTLPSK.S [51, 62]	655,37	1067,64 954,56 841,47 770,44 673,38	23
	R.GDSVVYGLR.S [145, 153]	483,25	908,48 793,45 706,42 607,35 508,28	17
	R.ISHELDSASSEV.- [287, 298]	637,29	1160,50 1073,47 936,41 807,37 694,28	23
Uromodulin	R.STEYGEGYA**C**DTDLR.G [185, 199]	868,85	1419,58 1256,52 1199,49 1070,45 1013,43	31
	R.FVGQGGAR.M [204, 211]	396,2116	644,34 545,27 488,25	14
	K.VFMYLSDSR.C [356, 364]	559,27	1018,46 871,39 740,35 577,29	20
	K.INFA**C**SYPLDMK.V [420, 431]	729,84	1231,54 1084,48 1013,44 853,41 766,38	26

	R.VGGTGMFTVR.M [449, 458]	512,76	868,43 811,41 710,36 653,34 522,30	18
	R.VLNLGPITR.K [597, 605]	491,80	883,53 770,45 656,40 543,32	17
Protein AMBP Preproprotein/Alpha-1-Microglobulin	K.GV**C**EETSGAYEK.T [88, 99]	665,28	1272,54 1173,47 1013,44 884,40 755,36	24
	R.ETLLQDFR.V [158, 165]	511,26	892,48 791,44 678,35 565,27	18
	R.EY**C**GVPGDGDEELLR.F [334, 348]	854,87	1416,64 1256,61 1199,59 1100,52 1003,46 946,44	31
Kininogen-1 Isoform 2 Precursor	K.AATGE**C**TATVGK.R [101, 112]	583,28	1094,51 1023,47 922,42 865,40 736,36	21
	K.ENFLFLTPD**C**K.S [208, 218]	692,33	1254,61 1140,57 993,50 880,42 733,35 519,22	25
	R.IASFSQN**C**DIYPGK.D [240, 253]	800,37	1415,62 1328,59 1181,52 1094,49 966,43	29

			852,39 301,18	
	K.DFVQPPTK.I [254, 261]	466,24	816,46 669,39 570,32 442,26 345,21	16
	K.ESNEELTES**C**ETK.K [330, 342]	778,32	1426,60 1339,56 1225,52 1096,48 967,44 854,35	28

SALIVA

	Peptide	Precursor m/z	Products m/z	Collision Energy (V)
α- Amylase	R.TSIVHLFEWR.W [25, 34]	644,34	1186,63 1099,60 986,52 887,45 750,39	23
	R.WVDIALE**C**ER.Y [35, 44]	645,81	1104,53 1005,46 890,44 777,35 706,31	23
	K.SSDYFGNGR.V [258, 266]	501,71	915,39 828,36 713,33 550,27	17
	R.ALVFVDNHDNQR.G [306, 317]	714,35	1243,58 1144,51 997,44 898,37 783,34	25
Annexin 1	K.GVDEATIIDILTK.R [58, 70]	694,38	1116,65 987,60 916,57 815,52 702,43	25
	R.SEIDMNDIK.A [303, 311]	532,75	977,46 848,41 735,33 620,30	19
	K.ILVAL**C**GGN.- [337, 345]	458,75	803,41 690,32 591,25 520,21	16
Transme mbrane Secretor	R.TVTIN**C**PFK.T [146, 154]	540,28	978,51 879,43 778,39	19

			665,30 551,26	
	R.ILLNPQDK.D [296, 303]	470,77	827,46 714,37 601,29 487,25	16
	R.LVSLTLNLVTR.A [525, 535]	614,88	1115,67 1016,61 929,57 816,49 715,44	22
Mutant Beta-Actin (Beta'-Actin)	K.AGFAGDDAPR.A [18, 27]	488,72	905,41 848,38 701,32 630,28 573,26	17
	R.GYSFTTTAER.E [196, 205]	566,76	1075,50 912,44 825,41 678,34 577,29	20
	K.EITALAPSTMK.I [315, 325]	581,31	1032,57 919,49 818,44 747,40 634,32	20
Mucin 5 AC	K.GVQLSDWR.D [1335, 1342]	480,74	903,46 804,39 676,34 563,25	17
	K.SEQLGGDVESYDK.I [2212, 2224]	713,82	1082,50 969,41 912,39 855,37 740,34	25
	R.AQAQPGVPLR.E [2999, 3008]	518,79	965,55 837,49 766,45 638,39 541,34	18

157

	K.SMDIVLTVTMVHGK.E [5478, 5491]	765,90	1197,70 1084,61 985,54 872,46 771,41	27
Cystatin S	K.QL**C**SFEIYEVPWEDR. M [115, 129]	985,95	1482,69 1335,62 1206,57 1093,49	35
	R.MSLVNSR.C [130, 136]	403,71	675,37 588,34 475,26	14

SEMEN

	Peptide	Precursor m/z	Products m/z	Collision Energy (V)
Epididymal Secretory Protein E1	K.SGIN**C**PIQK.D [94, 102]	508,76	929,48 872,46 759,38 645,33	18
	K.SEYPSIK.L [116, 122]	412,21	736,38 607,34 444,28	14
Mucin-6 Isoform X1	K.VTNEFVSEEGK.F [182, 192]	619,79	1139,52 1038,47 924,43 795,38 648,31	22
	R.ETDP**C**SMSQLNK.V [573, 584]	705,30	1179,51 1064,48 967,43 807,40 720,37	25
	R.GVLLWGWR.S [646, 653]	493,78	929,53 830,46 717,38 604,29	17
	K.VYHLPYYEA**C**VR.D [1075, 1086]	785,38	1307,62 1170,56 1057,47 960,42 797,36	28
Semenogelin II	K.DIFTTQDELLVYNK.N [251, 264]	849,93	1323,67 1222,63 1121,58 993,52 878,49	30
	K.ISYQSSSTEER.H [345, 355]	643,79	1173,50 1086,46 923,40	22

		795,34 708,31	
	K.QDLLSHEQK.G [534, 542]	549,28	969,50 854,47 741,38 628,30	19
PSA	R.IVGGWE**C**EK.H [20, 28]	539,25	964,41 865,35 808,32 751,30 565,22	19
	R.LSEPAELTDAVK.V [121, 132]	636,83	1072,55 943,51 846,45 775,41 646,37	22
	K.FML**C**AGR.W [190, 196]	427,70	707,33 576,29 463,20	15
Galectin-3-Binding Protein	R.ELSEALGQIFDSQR.G [137, 150]	796,89	1263,63 1134,59 1063,55 950,46 893,44	28
	R.IDITLSSVK.C [216, 224]	488,2897	862,48 747,46 634,37 533,32	17
Prostate And Testis Expressed Protein 1	R.GI**C**TATTEEA**C**MVGR.M [64, 78]	828,3629	1325,5824 1224,5347 1153,49 1052,45 951,40	30
Beta-Microseminoprotein	K.T**C**SVSEWII.- [85, 93]	547,7630	993,47 833,44 746,41 647,34 560,30	19

Semenogelin-1 Preproprotein	R.LWVHGLSK.E [165, 172]	470,27	826,45 640,37 541,31	16
	K.GESGQSTNR.E [404, 412]	468,21	878,39 749,35 662,32 605,30 477,24	16

Table S-4: List of selected protein biomarkers, reporting selected peptide sequence, precursor m/z, product ions m/z and Collision Energy (V).

BLOOD				
	Peptide	**Precursor m/z**	**Products m/z**	**Collision Energy (V)**
Chain α of Hemoglobin	K.VGAHAGEYGAEALER.M [16, 30]	765,37	1302,60 1165,54 1094,51 1037,48 908,44	27
	R.VDPVNFK.L [92, 98]	409,72	719,37 604,34 507,29	14
	K.FLASVSTVLTSK.Y [127, 138]	626,86	992,56 921,52 834,49 735,42 648,39	22
Chain β of Hemoglobin	R.LLVVYPWTQR.F [30, 39]	637,87	1161,64 1048,55 949,48 850,42 687,35	23
	K.GTFATLSELHCDK.L [82, 94]	739,85	1173,55 1102,51 1001,47 888,38 801,35	26
	K.VVAGVANALAHK.Y [132, 143]	575,34	951,53 880,49 823,47 724,41 653,37	20
Hemopexin	R.ELISER.W [82, 87]	373,70	617,36 504,27 391,19	13
	K.VDGALCMEK.S [401, 409]	511,74	923,39 808,36 751,34	18

			680,31 567,22	
Haptoglobin	K.DIAPTLTLYVGK.K [156, 167]	645,87	1062,61 991,58 894,52 793,48 680,39	23
	R.VMPICLPSK.D [202, 210]	522,78	945,48 814,44 717,39 604,31	18
URINE				
Osteopontin	K.QNLLAPQTLPSK.S [51, 62]	655,38	1067,64 954,56 841,47 770,44 673,38	23
	R.GDSVVYGLR.S [145, 153]	483,26	908,48 793,46 706,42 607,36 508,29	17
	R.ISHELDSASSEV.- [287, 298]	637,30	1160,50 1073,47 936,41 807,37 694,28	23
Uromodulin	R.STEYGEGYACDTDLR.G [185, 199]	868,86	1419,58 1256,52 1199,49 1070,45 1013,43	31
	K.VFMYLSDSR.C [356, 364]	559,27	1018,46 871,39 740,35 577,29	20

164

K.INFAC̲SYPLDMK.V [420, 431]	729,84	1231,54 1084,48 1013,44 853,41 766,38	26
R.VGGTGMFTVR.M [449, 458]	512,77	868,43 811,41 710,36 653,34 522,30	18
R.VLNLGPITR.K [597, 605]	491,81	883,53 770,45 656,40 543,32	17
SALIVA			
α Amylase R.WVDIALEC̲ER.Y [35, 44]	645,81	1104,53 1005,46 890,44 777,35 706,31	23
K.SSDYFGNGR.V [258, 266]	501,72	915,39 828,36 713,33 550,27	17
Annexin 1 K.GVDEATIIDILTK.R [58, 70]	694,39	1116,65 987,60 916,57 815,52 702,43	25
R.SEIDMNDIK.A [303, 311]	532,75	977,46 848,41 735,33 620,30	19
K.ILVALC̲GGN.- [337, 345]	458,75	803,40 690,32 591,25 520,21	16

Mucin 5AC	K.GVQLSDWR.D [1335, 1342]	480,75	903,46 804,39 676,34 563,25	17
	R.AQAQPGVPLR.E [2999, 3008]	518,80	965,55 837,49 766,45 638,39 541,34	18
Cystatin S	K.QLCSFEIYEVPWEDR.M [115, 129]	985,95	1482,69 1335,62 1206,57 1093,49	35
	R.MSLVNSR.C [130, 136]	403,71	675,37 588,34 475,26	14
SEMEN				
Mucin 6 Isoform X1	K.VTNEFVSEEGK.F [182, 192]	619,80	1139,52 1038,47 924,43 795,38 648,31	22
	R.ETDPCSMSQLNK.V [573, 584]	705,31	1179,51 1064,48 967,43 807,40 720,37	25
	R.GVLLWGWR.S [646, 653]	493,78	929,53 830,46 717,38 604,29	17
	K.VYHLPYYEACVR.D [1075, 1086]	785,38	1307,62 1170,56 1057,47 960,42 797,36	28

Semenogelin II	K.GHYQNVVDVR.E [217, 226]	593,80	1129,57 992,51 829,45 701,39	21
	K.DIFTTQDELLVYNK.N [251, 264]	849,93	1323,67 1222,63 1121,58 993,52 878,49	30
	K.ISYQSSSTEER.H [345, 355]	643,80	1173,50 1086,46 923,40 795,34 708,31	23
	K.QDLLSHEQK.G [534, 542]	549,28	969,50 854,47 741,38 628,30	19
Prostate Specific Antigen Precursor (PSA)	R.IVGGWECEK.H [20, 28]	539,26	964,41 865,35 808,32 751,30 565,22	19
	K.HSQPWQVLVASR.G [29, 40]	704,38	1270,69 1183,65 1055,59 958,54 772,46	29
	R.LSEPAELTDAVK.V [121, 132]	636,84	1072,55 943,50 846,45 775,41 646,37	22
	K.FMLCAGR.W [190, 196]	427,70	707,33 576,29 463,20	15
PAP	R.SPIDTFPTDPIK.E [47, 58]	665,85	1146,60 1033,52 918,49	24

		817,44 670,37	
K.DFIATLGK.L [185, 192]	432,74	749,45 602,38 489,30	15
R.ELSELSLLSLYGIHK.Q [236, 250]	851,47	1372,78 1243,74 1130,65 1043,62 930,54	31
Semenogelin-1 Preproprotein R.LWVHGLSK.E [165, 172]	470,27	826,45 640,37 541,30	16
K.VQTSLCPAHQDK.L [233, 244]	692,34	1156,54 1055,49 968,46 855,37 695,34	25
K.DVSQSSIYSQTEEK.A [307, 320]	800,87	1299,60 1171,54 1084,51 997,48 884,39	30
K.GESGQSTNR.E [404, 412]	468,21	878,39 749,35 662,32 605,30 477,24	16

Table S-5. Specificity test of MRM method in the identification of biomarker proteins when biological fluids were mixed on cloth.

	BLOOD				URINE		SALIVA				SEMEN				
	αHemoglobin	βHemoglobin	Hemopexin	Haptoglobin	Uromodulin	Osteopontin	Alpha amylase	Muc 5 Ac	Annexin A1	Custatin S	PSA	PAP	Muc 6	Semenogelin I	Semenogelin II
Semen and Blood	X	X	X	X							X	X	X		X
Semen and Urine					X	X					X	X	X	X	X
Blood and Urine	X	X	X	X	X	X					X	X			

X is reported when the unique peptides belonging to the target proteins characteristic of individual biological matrix were identified by MRM analysis.

169

Figure S-1. Total Ion Current chromatogram (A) and the TIC chromatograms analysis of two peptides 507-605 and 204-211 from Uromodulin (B) (396.21 → 644.34 m/z, 396.21 → 545.27 m/z, 396.21 → 488.25 m/z). MRM chromatograms of three transitions for the (204-211) Uromodulin peptide (C). The different transitions perfectly co-eluted at a retention time of 2.99 min thus indicating that they all originated from the same precursor ion
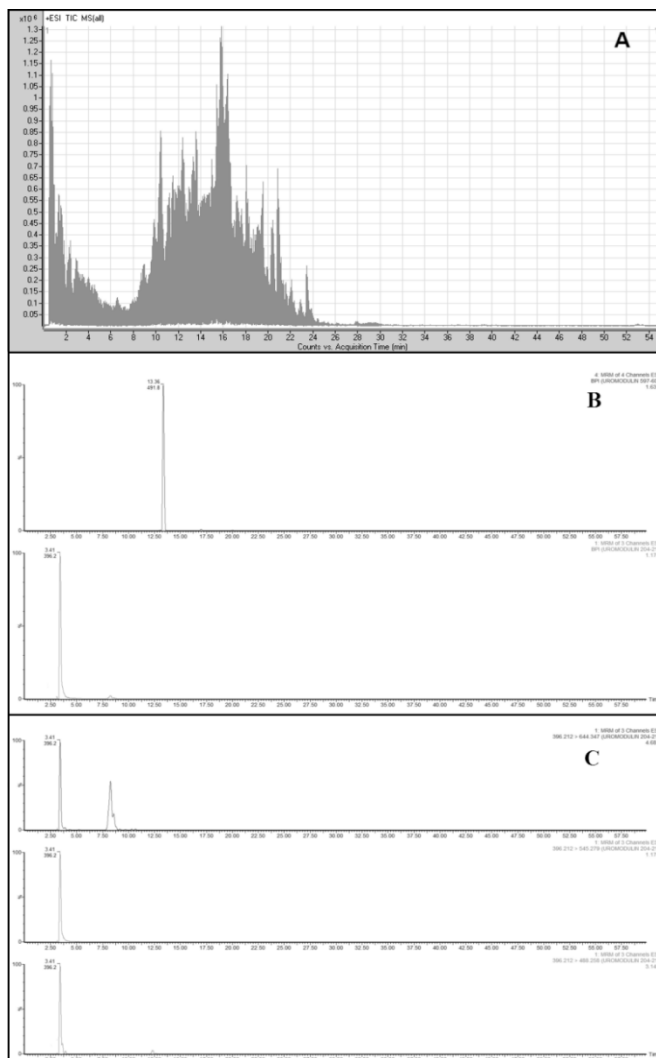
Figure S-2. TIC chromatograms analysis of two peptides 35-44 and 258-266 from α-amylase at retention time of 11.59 min, 3.35 min, respectively (A). MRM TIC chromatogram of three transitions (403.71 → 675.37, 403.71 → 588.34, 403.71 → 475.26) used to monitor the (130-136) Cystatin S peptide (Panel B).
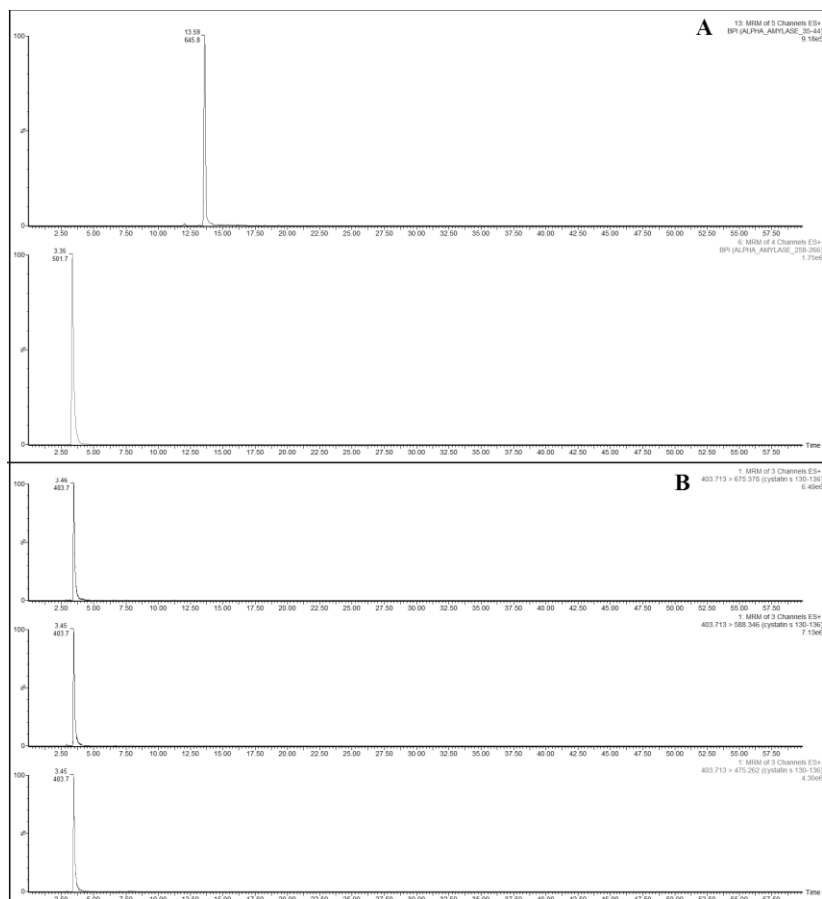
Figure S-3. TIC chromatograms analysis of three peptides 81-91, 165-172, 404-412 from Semenogelin 1 at retention time of 15.55 min, 8.31 min and 6.03 min, respectively (A). MRM TIC chromatograms of three transitions (427.7 →707.33, 427.7→ 576.29, 427.7 →463.2) used to monitor the (190-196) PSA peptide (Panel B).
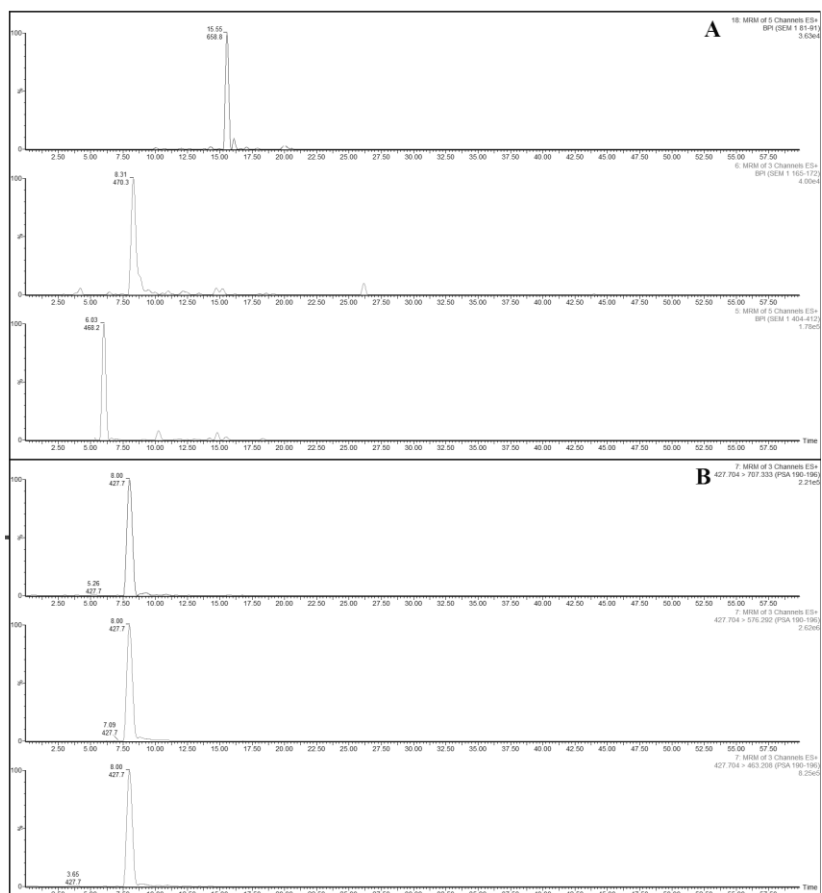
Figure S-4. DNA profiling. DNA profiling obtained by multiple amplification of 17 polimorphic regions comprising the D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, TH01, D16S539, D2S1338, D19S433, ACTBP2 (SE33), D1S1656, D2S441, D10S1248, D12S391, D22S1045 and Amelogenin STR loci using the Investigator ESSplex SE Plus Kit della Qiagen (Qiagen, Hilden, Germany).
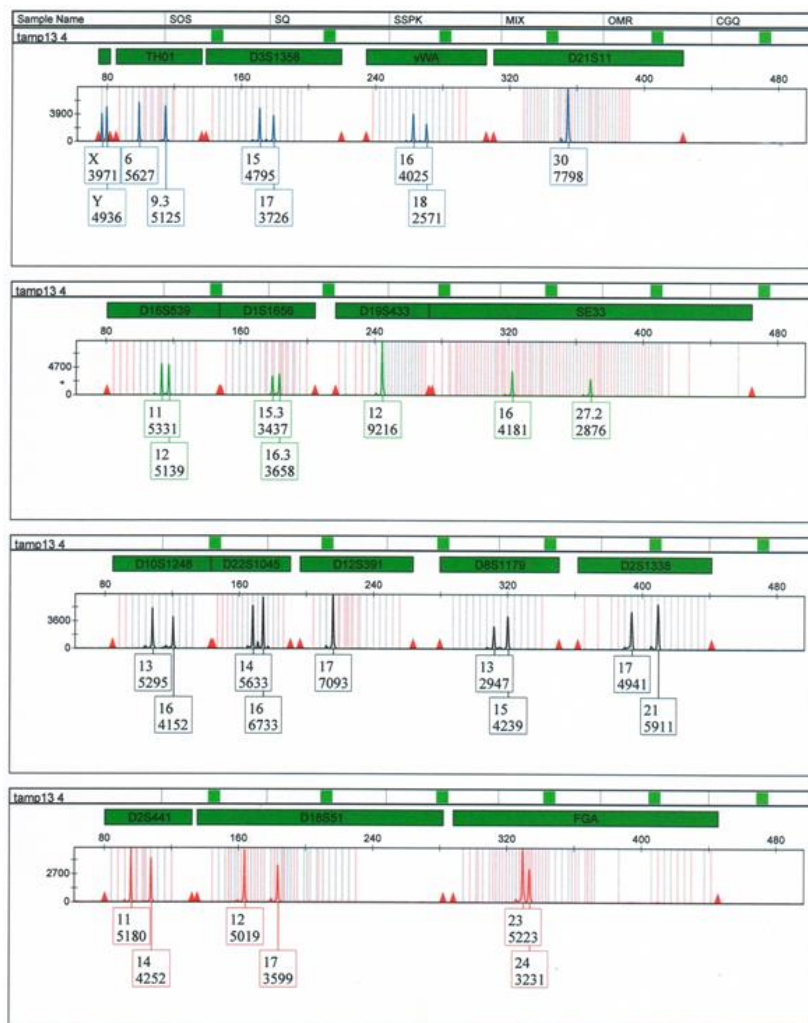
Chapter 7

7.1   Introduction

7.1.1   Human Colorectal Carcinoma

Colorectal cancer (CRC) is the third most common cancer in men and the second most common cancer in women. More than 1.2 million new cases of colorectal cancers are diagnosed globally, with more than 600,000 related deaths in 2008. Both genomic and epigenetic alterations are common in CRC and are the driving forces of tumourigenesis[1].This cancer develops as a result of the pathologic transformation of normal colonic epithelium to an adenomatous polyp and ultimately an invasive cancer. The multistep progression requires years and possibly decades and is accompanied by a number of recently characterized genetic alterations. Mutations in two classes of genes, tumour-suppressor genes and proto-oncogenes, are thought to impart a proliferative advantage to cells and contribute to development of the malignant phenotype[2].

7.1.2 Dna Mutations

Cancer is a result of mutations that inhibit certain gene functions, leading to uncontrollable cell growth. A mutation, which may arise during replication and/or recombination, is a permanent change in the nucleotide sequence of DNA. Damaged DNA can be mutated either by substitution, deletion or insertion of base pairs. Mutations, for the most part, are harmless except when they lead to cell death or tumour formation. Because of the lethal potential of DNA mutations cells have evolved mechanisms for repairing damaged DNA[3].

There are three types of DNA Mutations:

➢   **Base Substitutions**

Single base substitutions are called point mutations, recall the point mutation Glu -> Val which causes sickle-cell disease (*Figure 1*). Point mutations are the most common type of mutation.

➢   **Deletions**

A deletion, resulting in a frameshift, results when one or more base pairs are lost from the DNA. If one or two bases are deleted the translational frame is altered resulting in a garbled message and nonfunctional product. A deletion of three or more bases leave the reading frame intact (*Figure 1*). A deletion of one or more codons results in a protein missing one or more amino acids. This may be deleterious or not.

➢ **Insertions**

The insertion of additional base pairs may lead to frameshifts depending on whether or not multiples of three base pairs are inserted (*Figure 1*).
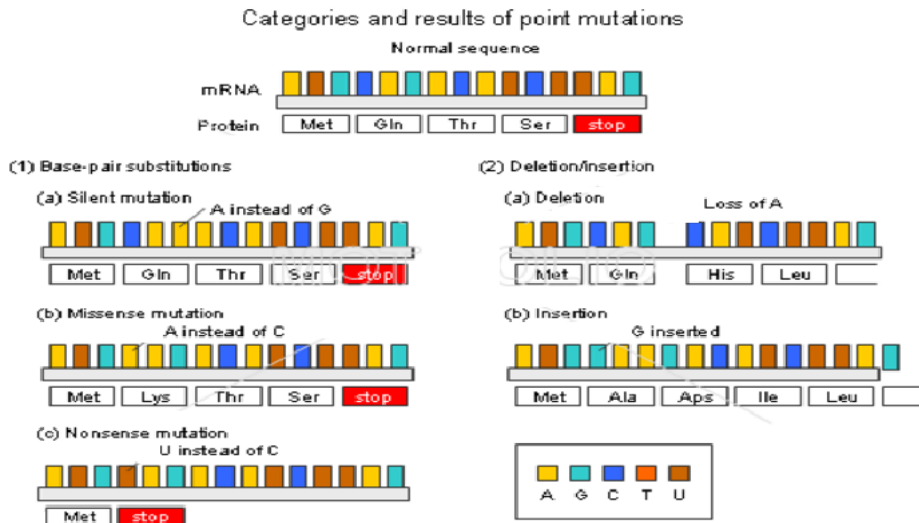
Figure 1: Three types of DNA mutations, base substitutions, deletions and insertions.

7.1.3 HCT116 Tumor Line

Genetic studies carried out on the HCT116 tumor line (studies conducted on mRNA in collaboration with Professor A. Nicosia from the University of Naples Federico II) have demonstrated that mutations of DNA were transcribed in mRNAs determining the modification of some proteins involved in tumorigenesis, which are reported in Table 1.These mRNAs encodes the following hypothetically mutated proteins that shows a mutated amino acid sequence at C-ter (showed in T*able 1*) :

Protein	MUTATED SEQUENCE *
1	13 AA
2	18 AA
3	31 AA
4	44 AA
5	34 AA
6	10 AA
7	187 AA
8	32 AA
9	38 AA
10	22 AA
11	66 AA
12	62 AA
13	43 AA
14	62 AA

Table 1: Sequences mutated at the C-terminal of proteins encoded by mRNA. Protein target are protected by secrecy agreement until publication

Studies conducted on these proteins show how their mutation is related to the onset of colon cancer, for example Protein 1 is involved in innate immune response by recognizing cytosolic double-stranded DNA and inducing caspase-1-activating inflammasome formation in macrophages. Defects in Protein 1 may be a cause of microsatellite unstable colon cancers[4]. Protein 3 mutation may enhance genomic instability in some colorectal cancers[5].

Starting from genetic studies that identified mRNAs, the aim of this study was to understand if:

- mRNAs are translated into proteins or degraded;
- in the case they are not degraded, the mutated proteins are released by the cell or are immediately degraded.

From the analytical point of view, the mutated C-terminal sequence of these proteins may be useful in detecting and analyzing them in the protein extract of the HCT116 cell line. The analytical strategy for the identification of these proteins is based on the development and optimization of a method based on Mass Spectrometry in Multiple Reaction Monitoring (MRM) mode which appears to be more specific, selective and faster than traditional methods such as the ELISA test.

The main part of the method development was to identify for each amino acid sequence peptides that are unique and characterizing the proteins in order to obtain a high specificity and selectivity for MRM analysis. Methods 'optimization was carried out by using a triple quadrupole instruments in MRM ion mode. The first analyzer only selects the values of m/z relative to the precursor ions, the second acts as a collision cell (CID) fragmenting the precursor and the third identifies the specific fragment. Thus, the optimized LC-MRM-MS method was *unique* for all proteins to be analyzed and for the peptide sequences selected, *fast* because in a single analysis the simultaneous identification of all the selected protein is allowed and *highly selective and specific* because all transitions monitored are derived from target peptides. The validation step was carried out by testing the MRM method on other two tumor cell lines LS174T and HT 29. The identification of these mutated proteins in colorectal carcinoma can be considered a preclinical exploratory study required to optimize the search of cancer biomarkers and to make it possible targeted drug therapy.

7.2 Materials and Methods

8 samples from HCT 116, LS174T and HT 29 cell lines respectively were purchased from Prof. Nicosia Laboratory.

7.2.1 Sample Treatment

Proteins extracts obtained from the HCT116 tumour line were submitted to in solution hydrolysis protocol that provides the following steps:

- Reduction (RED) and carboamidomethylation (CAM);
- Precipitation of proteins with $CHCl_3$, MeOH and H_2O;
- Tryptic digestion.

Dried proteins were dissolved in denaturation buffer: 6M Guanidine, 0.3M Tris HCl, 10mM EDTA at pH 8.0. Reduction was carried out by using a 10:1 DTT: cysteine molar ratio. After incubation at 37°C for 2 h, iodoacetamide was added to perform carboamidomethylation using an excess of alkylating agent of 5:1 respect to the moles of thiolic groups. The mixture was then incubated in the dark at room temperature for 30 minutes. The alkylation reaction was stopped with 10 µL of 10% HCOOH, in order to achieve an acid pH.

Alkylated proteins were purified by a precipitation step conducted by adding 400 µL of MeOH, 100 µL of $CHCl_3$ and 300 µL of H_2O. The resulting solution was centrifuged for 5 minutes at 12000 rpm. The proteins will be precipitated at the interface organic-inorganic phases. The supernatant was removed and was added 300 µL of MeOH. The solution was centrifuged for 5 minutes at 12000 rpm and it was clearly visible a protein pellet. The supernatant was removed, and samples was dried under vacuum.

Protein digestion was carried out in 10 mM AMBIC using trypsin at a 50:1 protein:enzyme mass ratio. The sample was incubated at 37°C for 16 h. Then sample was dried under vacuum and stored at -20°C until the LC-MRM/MS analysis.

7.2.2 Development of LC-MRM-MS method

For the LC-MRM-MS method optimization was necessary to perform a sequence alignment to identify target peptides, related to the mutated sequence at the C-terminal, that uniquely identifies the mutated protein.

Sequence alignment between mutated protein sequences (*Table 1*) and corresponding non-mutated proteins and all human proteins was performed using the *Align* function of the UniProt database[6]. This step was necessary to select targeted peptides unique in terms of sequence. From sequence alignment it has been shown that not all proteins provided by Prof Nicosia have been identified as unique peptides; therefore, for the MRM method, only the following mutated proteins listed Table 2 were selected. After the alignment of the sequence, for all mutated proteins Skyline showed a list of target peptides, which uniquely identify the mutated protein, with its precursor-product ion transitions and relative collision energies (*Table 2*).

Table 2: MRM Method contains for selected mutated HCT116 proteins, the specific transitions precursor-product ion selected by Skyline, for all target peptides corresponding to mutated sequence at C terminal. Protein target are protected by secrecy agreement until publication

Protein	Peptide	Precursor ion m/z	Product ion m/z	Collision Energy (V)
1	HRE	334.6956++	531.3249+ 375.2238+	11
	EVK	266.1661+	402.2823+ 303.2139+	9
3	LCL	430.7548++	747.4182+ 587.3875+ 474.3035+	15
	TLL	387.2420++	672.4291+ 559.3450+ 446.2609+	13
4	LTW	417.7182++	721.3450+ 620.2973+ 434.2180+	14
6	KPG	640.8188++	1152.5354+ 1055.4826+ 998.4612+ 897.4135+ 810.3815+ 697.2974+	23
14	QIL	365.2163++	601.3668+ 488.2827+ 375.1987+	12
	EYF	350.1892++	570.3286+ 407.2653+	12
	AMT	453.2417++	834.4390+ 703.3985+ 602.3508+ 503.2824+	16
	AVS	527.3244++	982.6044+ 883.5360+ 796.5039+ 683.4199+ 555.3613+ 442.2772+	18

7.3 Results and Discussion

Regulation of eukaryotic gene expression is crucial in pathophysiological responses to extracellular and intracellular signals in the context of homeostasis maintenance and cellular differentiation or in stress response and cell survival. Damaged DNA can be mutated either by substitution, deletion or insertion of base pairs. Mutations, for the most part, are harmless except when they lead to cell death or tumour formation. Post-transcriptional control of mRNA trafficking and metabolism plays a critical role in the actualization and fine tuning of the genetic program of cells, both in development and in differentiated tissues. We started from prof. Nicosia data related to the presence, in HCT 116 cells, of mutated mRNA belonging to the proteins presented in *Table 1*. The project's focus point was to verify, with an MRM mass spectrometry method, the translation of these mRNAs into the corresponding proteins with the C-ter mutation. In this case was exploited, on one hand, the possibility of conducting multiplex analyses with high specificity and sensitivity deriving from a targeted analysis and, on the other hand, the considerable time and cost saving of the method optimization compared to the immunoenzymatic assays or fluorescence techniques. To set up the MRM method, the main step involved the selection of the peptides to be monitored. Sequence alignments of the mutated protein with the WT protein and all the proteins contained in the UniProt Database belonging to Human was carried out and the selection of peptides with a unique amino acid sequence, the so-called proteotypic peptides, was realized. From sequence alignment not all proteins provided by Prof Nicosia have been identified as unique peptides; therefore, for the MRM method, only the mutated proteins listed in Table 2 Next step was defining the best precursor ion-product ion transitions and other instrumental parameters (dwell time and collision energy), subsequently, samples of the HCT 116 cell line were treated. The eight samples were subjected to a classical protocol of hydrolysis with trypsin, as reported in Materials and Methods. To evaluate the robustness of the method, peptides' mixture were analysed in triplicate and as an example, *Figure 2* shows the MRM TICs Chromatogram for TLL and LCL of protein 2 are reported.
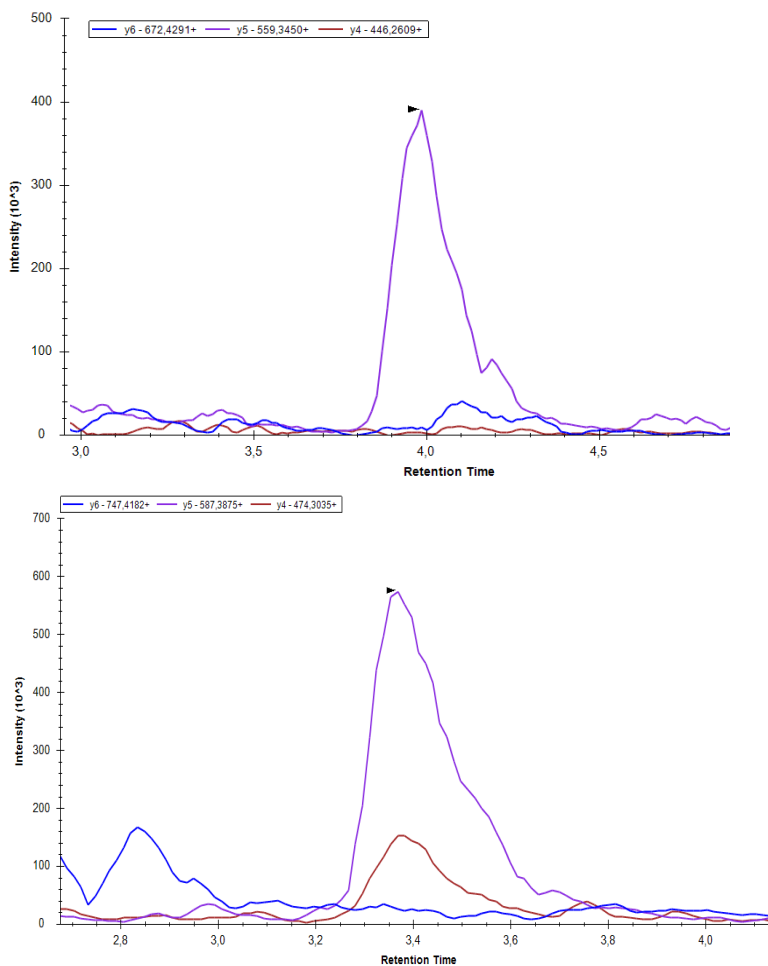
Figure 2: Enlargment of MRM TIC Chromatogram for TLL and LCL respectively. All the monitored transitions are coeluted at 4 and 3,4 min respectively.

Thanks to MRM analysis it was possible to verify the production of the target protein. This kind of analysis is qualitative, the perfect coelution of the monitored transitions for each target peptide allowed the identification, but it also gave a trend of mutated protein expression in the sample. *Figure 3* shows, in the form of histograms, the average of the total ionic currents recorded for each target peptide for each analysed sample, underline the behaviour of each protein. From the strictly analytical point of view, the collected data showed a good reproducibility with coefficients of variation lower than 20
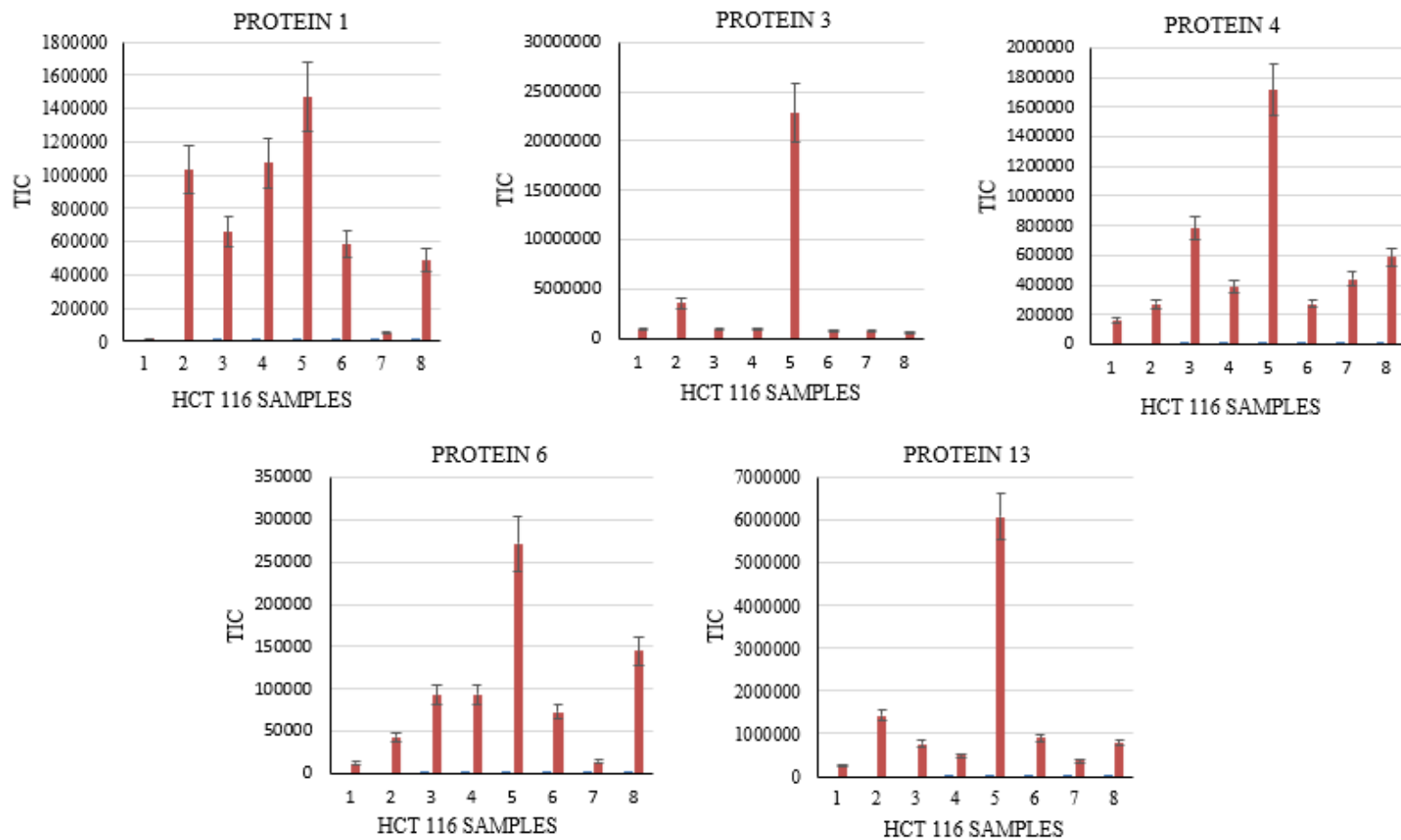
Figure 3: Mutated proteins trend in HCT 116 samples. The histograms showed the average of the total ion current recorded for each transition of each target peptide belonging to the mutated C-ter.

Figure 3 clearly shown that sample 5 has a higher amount of mutated protein in comparison to the others.

From a biological point of view, it was possible for the first time to demonstrate the translation of the target proteins with the mutation of the C-ter due to alternative splicing processes. The presence of proteins that are involved in DNA mismatch repair mechanisms it is possible to hypothesize that the mutated protein, if produced and if functioning, could negatively influence the regulation of mRNA translation which may contain defects and therefore lead to malfunction of the whole cell. Damaged DNA can be repaired by several different mechanisms like Mismatch Repair mechanism that is a system for recognizing and repairing erroneous insertion, deletion, and mis-incorporation of bases that can arise during DNA replication and recombination, as well as repairing some forms of DNA damage. Mismatch repair is strand-specific. During DNA synthesis the newly synthesized (daughter) strand will commonly include errors. In order to begin repair, the mismatch repair machinery distinguishes the newly synthesized strand from the template (parental). However, in other prokaryotes and eukaryotes, the exact mechanism is not clear. Subsequent functional studies will support the idea that this mutated protein can not fully perform its function or that due to the mutation is sent to the proteasome or migrate into other cellular compartments compromising the proper functioning of the cell. The last phase of the work involved the validation of the results obtained for the HCT 116 cell line with two other cell lines: HT 29 and LS174T. These two cell lines are commonly used for the study of adenocarcinomas or colon tumours7. Therefore, eight samples were provided for each cell line that were treated as previously reported for HCT 116 and analysed in triplicate using the same LC-MRM / MS method. Similar results, as reported for HCT 116, have been obtained and are shown in Figures 4 and 5.
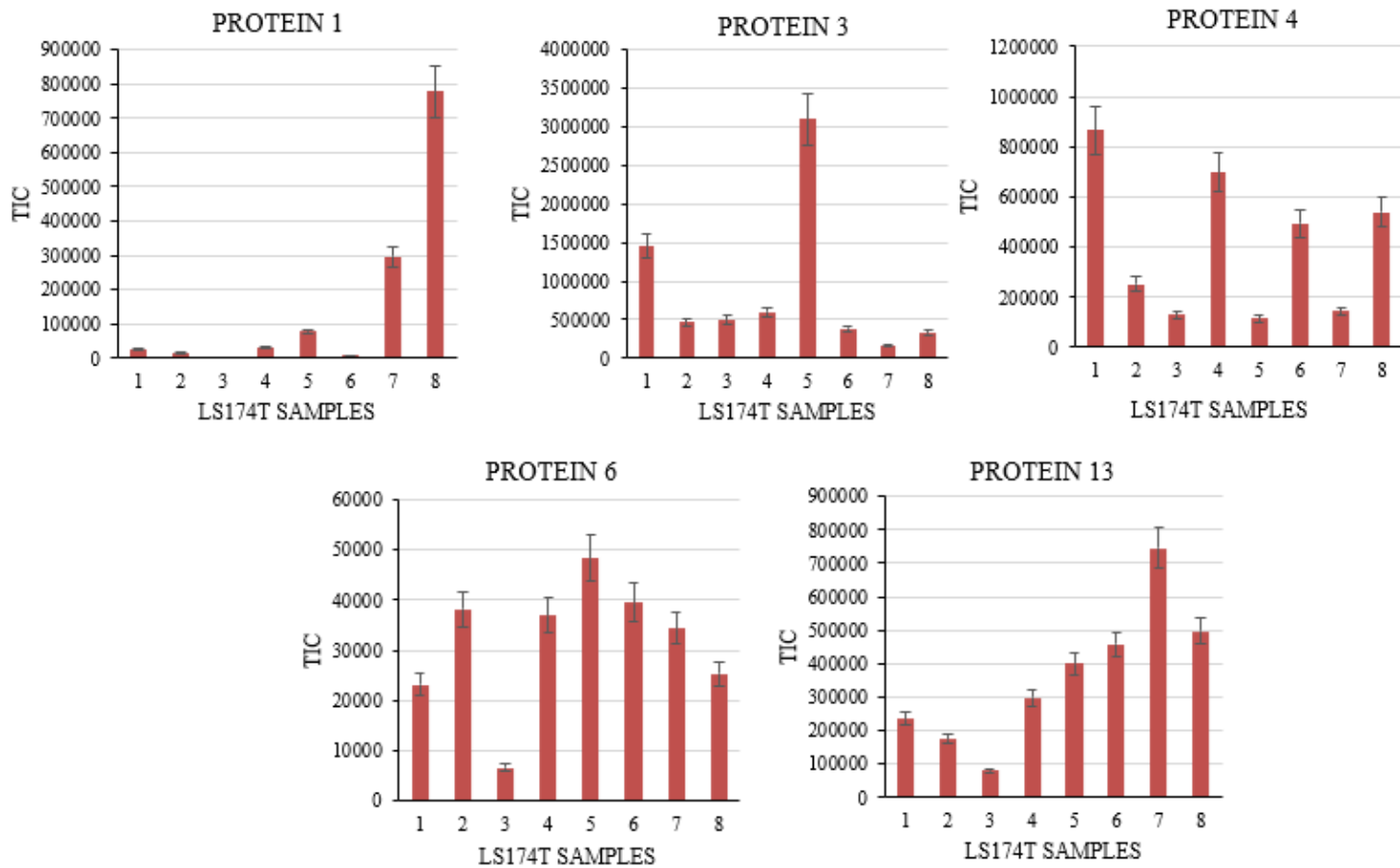
Figure 4: Mutated proteins trend in LS 147T samples. The average of the total ion current recorded for each transition of each target peptide is reported in histogram form.
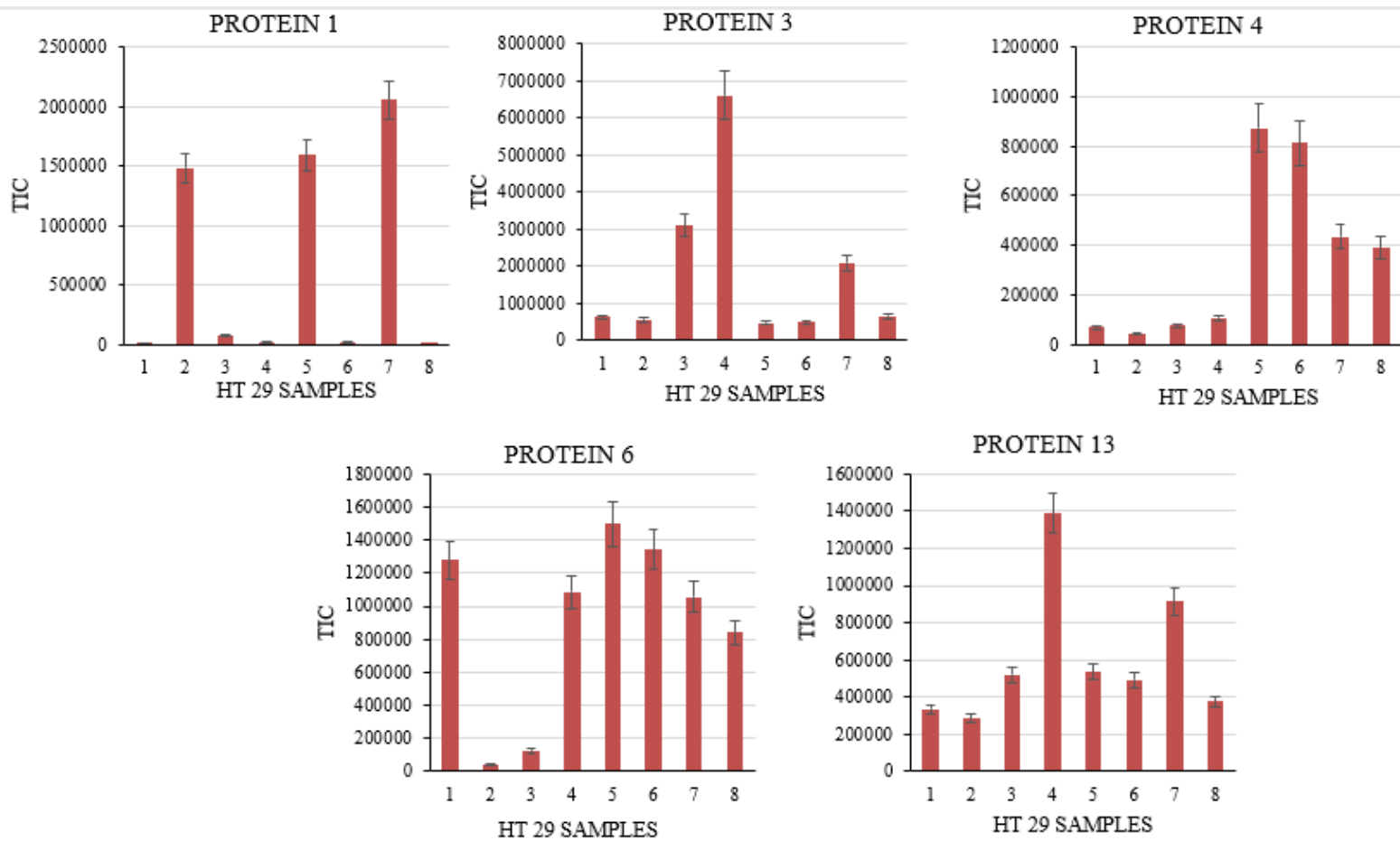
Figure 5: Mutated proteins trend in HT29 samples. The histograms showed the average of the total ion current recorded for each transition of each target peptide belonging to the mutated C-ter.

Figure 4 shows the trend of the mutated proteins in the LS174T tumour line. We can see an increase from sample 1 to sample 8, especially for proteins 13 and 6. For the HT 29 (Figure 5) we can note similar trend for Protein 6 and Protein 4.

The identification of these mutated proteins in these colon carcinoma lines can be considered a preclinical exploratory study required to optimize the search of cancer biomarkers and to make it possible drug therapy targeted. The analyzes carried out on the three HCT 116, HT29 and LS174T cell lines are qualitative analysis. They are to be considered as demonstrations of the existence of proteins with the C-ter mutation and of the expression trend of these proteins within the eight samples. Future developments include the implementation of the MRM method developed with internal standards, isotopically labeled peptides, which confirm the presence of target peptides and allow to perform a quantitative analysis of the same.

7.4 Conclusion

Genetic studies carried out on the HCT116 tumour line (in collaboration with Prof. A.Nicosia of the University of Naples Federico II) have demonstrated that mutations of DNA were transcribed in mRNAs. Such mutations involve hypothetical translation of truncated proteins with respect to WT proteins and having a mutated amino acid sequence at C-ter. This thesis was based on the development and optimization of a MRM-based method of MS ion mode to identify within a cell line extract of the HCT 116 line a series of mutated proteins produced by gene mutations with a mutated sequence at C-ter. The main part of the method development is to identify for each amino acid sequence peptides that are unique and characterizing the protein under study in order to obtain a high specificity and selectivity for MRM analysis. These peptides were subjected in silico analysis by Skyline providing specific output precursor -product ions transitions and instrumental parameters (collision energy) needed for optimizing the method. Methods 'optimization was carried out by using a triple quadrupole instruments in MRM ion mode. Thus, the optimized LC-MRM-MS method was unique for all proteins to be analysed and for the peptide sequences selected, fast because in a single analysis the simultaneous identification of all the selected protein is allowed and highly selective and specific because all transitions monitored are derived from target peptides. This project showed that the translation phase of some of hypothetical mutated proteins takes place and that are not immediately degraded are produced. Mutated C-terminal sequence of these proteins was useful in detecting and analysing them in the protein extract of the HCT116 cell line. From qualitative analysis of LC-MRM-MS results, mutated proteins in HCT116 samples are 1, 3, 4, 6, 14 and they increase in sample 5. The developed LC-MRM-MS method has also been tested on other two tumour lines LS174T and HT29. The identification of these mutated proteins in colorectal carcinoma can be considered a preclinical exploratory study required to optimize the search of cancer biomarkers and to make it possible drug therapy targeted.

7.5 References

(1)   Armaghany, T.; Wilson, J. D.; Chu, Q.; Mills, G. Genetic Alterations in Colorectal Cancer. *Gastrointestinal Cancer Research : GCR* **2012**, *5*, 19-27.

(2)   Gryfe, R.; Bapat, B.; Gallinger, S.; Swallow, C.; Redston, M.; Couture, J. Molecular biology of colorectal cancer. *Current problems in cancer* **1997**, *21*, 233-299.

(3)   Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J. Mutations: types and causes. *Molecular Cell Biology* **2000**, *4*.

(4)   Yamamoto, H.; Adachi, Y.; Taniguchi, H.; Kunimoto, H.; Nosho, K.; Suzuki, H.; Shinomura, Y. Interrelationship between microsatellite instability and microRNA in gastrointestinal cancer. *World journal of gastroenterology: WJG* **2012**, *18*, 2745.

(5)   Woerner, S. M.; Kloor, M.; Schwitalle, Y.; Youmans, H.; Doeberitz, M. v. K.; Gebert, J.; Dihlmann, S. The putative tumor suppressor AIM2 is frequently affected by different genetic alterations in microsatellite unstable colon cancers. *Genes, Chromosomes and Cancer* **2007**, *46*, 1080-1089.

(6)   Orimo, H.; Nakajima, E.; Ikejima, M.; Emi, M.; Shimada, T. Frameshift mutations and a length polymorphism in the hMSH3 gene and the spectrum of microsatellite instability in sporadic colon cancer. *Japanese journal of cancer research* **1999**, *90*, 1310-1315.

(7)   Hatayama, H.; Iwashita, J.; Kuwajima, A.; Abe, T. The short chain fatty acid, butyrate, stimulates MUC2 mucin production in the human colon cancer cell line, LS174T. *Biochemical and biophysical research communications* **2007**, *356*, 599-603.

8. Conclusion

8.1 SRM compared to other quantification techniques

Targeted metabolomics and proteomics by SRM have shown the potential of becoming the technique of choice for the identification and quantification of specific compounds in biological matrices. Several methods can produce metabolic signatures of biomaterials, including nuclear magnetic resonance (NMR), gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), capillary electrophoresis-mass spectrometry (CE-MS), and so on. GC coupled to MS has been extensively used in metabolome analysis because of its high separation efficiency and the easy interfacing of GC with MS. It can be used to analyse a wide range of volatile compounds and semi-volatile compounds through chemical derivatization. LC coupled to MS is useful tool for metabolomics studies. They can resolve and quantify multiple components in crude biological extracts in the nanomole to picomole and even femtomole range from as little as sub microliter volumes. LC-UV-MS is routinely used for metabolites quantification based on the peak area associated with UV signals. Currently LC-MS/MS is considered the workhorse of metabolites detection because the derivatization of polar/on-volatile metabolites for GC-MS analysis is both time consuming and require fewer material compared to LC. Analytical methods based on affinity reagents, such as antibodies, have been routinely used in biological and biomedical research to measure the presence and/or quantity of target proteins by exploiting the antibody specificity and the direct or indirect detectability of a tag such as radioisotopes, enzymes and fluorophores. The methods with the lowest limit of detection have the potential to detect, in principle, low-abundance compounds, down to zeptomole detection limits. However, the development of reagents of suitable specificity and affinity to support the unambiguous measurement of the target protein in a complex background remains challenging, expensive and arduous. A step toward increased multiplexing of affinity reagent–based assays was the development of protein arrays in which microarrays, chips or beads allow for the parallel, multiplexed screening of large numbers of proteins, in an automated manner, and with minimal consumption of sample and reagents.

Another used technique is the direct tagging of the target protein with a detectable and quantifiable label such as a fluorescent protein and its

quantification by flow cytometry or microscopy techniques. The multiplexing capabilities are normally limited, and the introduced tag can potentially affect the localization, function, stability and expression of the target protein. These fusion proteins offer the possibility to quantify the amount of the target protein in single cells, thus revealing the cell-to-cell variability in expression and to provide information about the subcellular localization of proteins. Other techniques to quantify specific proteins in single cells are immunofluorescence, which can be applied to cells or tissue sections, and mass cytometry, which expands the multiplexing capabilities of single-cell probes using multiatom elemental tags. However, both techniques rely on antibodies and thus have limitations associated with their use.

8.2 Advantage of SRM

In recent years there has been a rapid increase of targeted proteomics and metabolomics using SRM methods. As demonstrated in this thesis project, SRM offers several advantages, such as specificity, sensitivity, linearity and it ideally suits for multiple quantitative analysis. The selective nature of SRM reduces the likelihood of interference from matrix and other target compounds, allowing the rapid analysis of more compounds in a single run. In addition, the ability of monitoring more transitions simultaneously allows greater specificity in the identification of specific target compound, reducing false positive. LC-SRM measurements showed that the range of analytes concentration accessible by direct LC-SRM is 4–5 orders of magnitude and that limits of detection at the low attomole level can be routinely reached, even in complex samples at high reproducibility and throughput.

Another important advantage is that the SRM analysis does not require a laborious step of sample pre-treatment or a detailed chromatographic separation offering the possibility of reducing the time analysis, including the extraction and data processing.

SRM analysis are not limited to a single application area but it is a versatile analytical approach. Laboratories working in food safety, environmental monitoring and bioanalysis can all benefit from the advantages of a SRM analysis. Therefore, costly sample preparation procedures can be optimized to deliver an efficient lab oratory workflow.

Moreover, one can measure 50–100 analytes in parallel in the same LC-SRM run if a reasonable duty cycle for accurate quantification and a sufficient dwell time to maintain a low limit of detection are applied. Upon

identity confirmation, the primary transitions are used to quantify the analyte. This approach has the potential to dramatically increase the multiplexing capabilities of SRM.
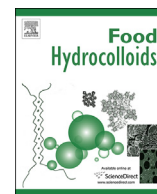
Bitter vetch (*Vicia ervilia*) seed protein concentrate as possible source for production of bilayered films and biodegradable containers

Akram Arabestani [a], Mahdi Kadivar [a], Angela Amoresano [b], Anna Illiano [b],
Prospero Di Pierro [b], Raffaele Porta [b, *]

[a] Department of Food Science, College of Agriculture, Isfahan University of Technology, Isfahan, 84156, Iran
[b] Department of Chemical Sciences, University of Naples "Federico II", Complesso Universitario di Monte Sant'Angelo, Via Cintia 26, 80126, Napoli, Italy

A R T I C L E   I N F O

A B S T R A C T

The manufacture of biodegradable and renewable containers from bitter vetch seed protein-based "bioplastics" was investigated. A bitter vetch seed protein concentrate was prepared and analyzed for proteins, carbohydrates, phenols, other several organic compounds and multi-elements. Protein film forming solutions were cast in the presence of two different glycerol concentrations and the film containing higher plasticizer amount was laminated with an additional corn zein layer. Both lamination process and lower glycerol concentration led to reduce film moisture content, total soluble matter and elongation at break, while both film tensile strength and water vapour barrier properties resulted enhanced. The obtained bioplastics were then processed by a new laboratory plastic moulding equipment specifically designed and fabricated to convert protein-based films to shaped containers. The use of either lower plasticizer concentration or corn zein lamination gave rise to potentially satisfactory vacuum thermoformed containers with acceptable resistance and stability.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The dramatic increase in production and lack of biodegradability of commercial polymers, particularly commodity plastics used in packaging industry and agriculture, focused public attention on a potentially huge environmental accumulation and pollution problem that could persist for centuries. In fact, petroleum-based synthetic polymers occur in the ecosystem as industrial waste products generating multiple serious problems, e.g. visual pollution, blockage of drains, livestock deaths and threat to aquatic life (Shimao, 2001). There is, thus, an increasing interest toward bio-based polymers to produce eco-friendly materials. In fact, a future extensive use of renewable and biodegradable plastics would allow to avoid the petroleum dependence, as well as the incineration and landfill pitfalls, making these "bioplastics" strategic for a sustainable development. There are three main ways to produce bioplastics: i) by starting from bio-based monomers obtained by fermentation or conventional chemistry and polymerizing them in a second step (e.g. polylactic acid), ii) by producing bio-based polymers directly in microorganisms or in genetically modified crops (e.g. polyhydroxyalkanoates), iii) by using different natural polymers (e.g. polysaccharides and proteins). Natural polymers are often modified mechanically, physically or chemically, as well as combined with plasticizers or further polymeric additives, to overcome several problems of different origin. Therefore, most of the raw material converted to bioplastics may come from the nature, green plants and microorganisms, or even from eatable vegetables. Two examples are sugarcane and wheat straw, annually renewable resources, the by-products of which (bagasse and straw waste) can be turned into several articles normally made from plastic or paper. Further widely used biomaterials are polylactic acid, polybutylenes succinate, poly3-hydroxybutyrate, polycaprolactone, starch- and cellulose-based biopolymers (Smith, 2005).

Among the possible natural biomacromolecules, proteins offer a valid alternative to the mentioned biopolymers (Arcan & Yemenicioglu, 2011; Bambdad, Goli & Kadivar, 2006; Borreani & Tabacco, 2015; Giosafatto et al., 2014; Peelman et al., 2013; Razavi, Amini, & Zahedi, 2015; Tang, Xiao, Chen, & Yang, 2011). Numerous potential applications for protein-based bioplastics are imaginable, from biodegradable containers to packaging materials, from agricultural mulching films to food service utensils. However,

---

* Corresponding author.
  E-mail address: raffaele.porta@unina.it (R. Porta).

to date only few commercial products have been made by using proteins as raw material, even though different vegetable and animal proteins have been suggested for obtaining edible coatings and biodegradable packages (Anderson & Lamsal, 2011; Cho, Lee, & Rhee, 2010; Cuq, Gontard & Guilbert, 1998; Peelman et al., 2013). Among these, whey and soy proteins represent a biodegradable and reproducible resource and, exhibiting an acceptable biocompatibility and processability, show an interesting potential mostly in both food and agricultural biotechnology (Di Pierro et al., 2013 and Mariniello et al., 2003). Some applications of whey and soy protein-based materials, especially for food preservation and packaging technology, have been reported (Di Pierro, Sorrentino, Mariniello, Giosafatto, & Porta, 2011; Grewell, Schrader, & Srinivasan, 2014; Rossi Marquez, Di Pierro, Esposito, Mariniello & Porta, 2014; Song, Tang, Wang & Wang, 2011). Also zein, a protein extracted from corn (*Zeamays* L.) during the wet-milling process (Shukla & Cheryan, 2001) has many properties that make it an appealing bioplastic feedstock, being an ethanol hydrophobic by-product and, thus, quite water resistant (Anderson & Lamsal, 2011; Arcan & Yemenicioglu, 2011; Ghanbarzadeh et al., 2006; Helgeson, Graves, Grewell, & Srinivasan, 2009). Some species of leguminous family are cheap protein sources, generally used only for animal feeding (Saki et al., 2008). One of these, bitter vetch (*Vicia ervilia*; BV), is widely cultivated throughout the temperate areas of Europe, western and central Asia, north Africa and Americas for the utilization of its seeds and hay. BV seeds contain up to 25% of protein (Larbi, El-Moneim, Nakkoul, Jammal, & Hassan, 2011; Pastor-Cavada, Juan, Pastor, Alaiz, & Vioque, 2011; Reisi, Zamani, Vatankhah, & Rahimiyan, 2011; Sadeghi, 2011) and, therefore, were recently analyzed as sustainable alternative source to produce biodegradable films, edible coatings and, potentially, properly shaped biodegradable containers (Arabestani, Kadivar, Shahedi, Goli, & Porta, 2013; Porta et al., 2015). Therefore, the specific objective of the present work was to develop biodegradable containers made of BV seed proteins using a simple laboratory equipment newly designed and fabricated. For this purpose a BV protein concentrate (BVPC) was prepared and its content in proteins, carbohydrates, as well as in multi-element, phenolic and other organic compounds, preliminarily analyzed.

2. Materials and methods

2.1. Materials

BV seeds were obtained from a local market in Isfahan, Iran. Zein powder (containing 90% protein dry weight) was purchased from Suvchem (India). Ethanol (96%) was supplied from Sepahan Teb Company (Isfahan, Iran) and polyethylene glycol 400 from Serva Electrophoresis GmbH (Heidelberg, Germany). Sodium hydroxide, hydrochloric acid (37%) and glycerol (about 87%) were purchased from the Merck Chemical Company (Darmstadt, Germany). All reagents for multi-element, organic compound, carbohydrate and protein analyses were from Sigma Chemical Company (St. Louis, MO).

2.2. Preparation of BV seed proteins and BVPC

Proteins were extracted from BV seeds according to Monsoor and Yusuf (2002) by milling the seeds to a fine powder (40 mesh) and soaking the latter in distilled water (1:10, w/v) brought at pH 10 by 0.1 N NaOH. After stirring (IKA® RH basic 2, Germany) at medium speed for 1 h at 25 °C, the mixture was centrifuged at 3200 g for 10 min and the precipitate discarded. The pH of the supernatant was then adjusted to pH 5.4 by 0.1 N HCl addition and the resulting precipitate, collected after centrifugation at 3200 g for

10 min, was finally dissolved by adding 0.1 N NaOH until pH 7.0 was reached and the obtained solution dried in a vacuum oven at 40 °C. The obtained dry concentrate of BV proteins was finally minced in a coffee grinder.

Protein content of both BV seeds and of the derived final BVPC was determined by the Kjeldahl's method (AACC, 2003), using a nitrogen conversion factor of 6.25.

2.3. BV protein identification

In gel trypsin digestion and protein identification. Proteins contained in BVPC sample were analyzed by 12.5% SDS-PAGE. The gel (1.5 mm) was run at constant 25 mA for 1 h, then stained with colloidal Coomassie, and the blue-stained protein bands excised from the gels and analyzed as follows. Gel particles were first separately washed with acetonitrile and with 0.1 M ammonium bicarbonate, then incubated with 10 mM DTT for 45 min at 56 °C and, finally, the cysteine residues alkylated by 5 mM iodoacetamide treatment for 15 min at room temperature in the dark. After washing with ammonium bicarbonate and acetonitrile, an enzymatic digestion was carried out by incubating the sample in the presence of trypsin (10 ng/μL) in 50 mM ammonium bicarbonate buffer, pH 8.5, at 4 °C for 2 h to allow the enzyme to enter the gel. The solution was then removed and a new aliquot of buffer was added and the incubation continued for 18 h at 37 °C. A minimum reaction volume, sufficient for a complete gel rehydration, was used. At the end of the incubation the peptides were extracted by washing the gel particles with 0.1% formic acid in 50% acetonitrile at room temperature and the sample freeze-dried.

Mascot identification. The sample was analyzed by LCMSMS on a 6520 Accurate-Mass Q-TOF LC/MS System (Agilent Technologies, Palo Alto, CA) equipped with a 1200 HPLC system and a chip cube (Agilent Technologies). After loading, the peptide mixture was first concentrated and washed in 40 nL enrichment column (Agilent Technologies chip), with 0.1% formic acid in 2% acetonitrile as eluent. The sample was then fractionated on a C18 reverse-phase capillary column (Agilent Technologies chip) at a flow rate of 400 nL min^{-1}, with a linear gradient of eluent B (0.1% formic acid in 95% acetonitrile) in eluent A (0.1% formic acid in 2% acetonitrile) from 5 to 80% in 50 min. Peptide analysis was performed using data-dependent acquisition of one MS scan (mass range from 300 to 1800 m/z) followed by MS/MS scans of the five most abundant ions in each MS scan. MS/MS spectra were measured automatically when the MS signal surpassed the threshold of 50,000 counts. Double and triple charged ions were preferably isolated and fragmented. The acquired MS/MS spectra were transformed in Mascot generic format (.mgf) and used for protein identification in the unreviewed set of protein entries that are present in the NCBInr database for all bacteria, with a licenced version of Mascot software (http://www.matrixscience.com) version 2.4.0. Additional Mascot search parameters were: peptide mass tolerance 10 ppm, fragment mass tolerance 0.6 Da, taxonomy Viridiplantae (Green Plants, 2,907,072 sequences) allowed trypsin missed cleavages up to 3, carbamidomethylation of cysteines as fixed modification, oxidation of methionine, pyro-Glu N-term Q, as variable modifications. Only doubly and triply charge ions were considered. Ions score was −10 log(P), where P is the probability that the observed match is a random event. Individual ion scores > indicated identity or extensive homology (P < 0.05).

2.4. Analysis of low MW carbohydrates and other organic compounds

1 mg of BVPC was dissolved in 500 μL of 1 M methanolic-HCl at 80 °C for 16 h. The re-N-acetylation of the monosaccharide mixture

was performed by adding 500 μL of methanol, 10 μL of pyridine and 50 μL of acetic anhydride at room temperature for 15 min. Sugars were finally trimethylsilylated in 200 μL of N, O-bis(trimethylsilyl) acetamide at 70 °C for 45 min. The sample was dried down under nitrogen, dissolved in 50 μL of hexane and centrifuged to remove the excess of solid reagents. The hexane supernatant (1 μL) was used for the GCMS analysis (Agilent 6890 Series GC, coupled to a detector MS 5973, Agilent Technologies) by using a fused silica capillary column (30 m, 0.5 mm ID, 0.25 μm ft) from Supelco. The injection temperature was 250 °C. For sugar analyses the oven temperature was increased from 25 °C to 90 °C in 1 min and held at 90 °C for 1 min before increasing to 140 °C at 25 °C/min, to 200 °C at 5 °C/min and finally to 300 °C at 10 °C/min. Conversely, the temperature gradient used for fatty acid analysis was as follows: from 60 °C to 150 °C at 15 °C/min or 20 °C/min and then to 280 °C at 3 °C/min or 7 °C/min. Electron ionisation mass spectra were recorded by continuous quadrupole scanning at 70eV ionisation energy.

For the analysis of other organic compounds BVPC sample was subjected to extraction step by using equal volume of dichloromethane and hexane and directly analyzed by GCMS with the injection system in splitless mode. The DB-5ms capillary column was 30 m long, 0.25 mm I.D., and carried a film of stationary phase 0.25 μm thick. The stationary phase used consisted of polydimethylsiloxane containing 5% of a phenyl derivative. Helium was used as carrier gas at a flow-rate of 1.0 mL/min. Volatiles were desorbed in the GC injector at 230 °C for 3 min. The gradient program for the GC analysis was as follows: 45 °C for 3 min, 12 °C/min to 150 °C, 18 °C/min to 230 °C, 19 °C/min to 250 °C. The GC detector employed was maintained at a temperature of 250 °C. The ionization energy was set at 70 eV and the resulting fragment ions were analyzed in the mass range m/z. 30-450.

2.5. Isolation and derivatization of oligosaccharide chains

BVPC tryptic digestion was carried out in 0.4% ammonium bicarbonate, pH 8.5, at 37 °C overnight using an enzyme to substrate ratio of 1:50 [w/w]. N-linked glycans were released from the peptide mixture by treatment with 0.1 enzyme unit of peptide-N-(4)-(N-acetyl-beta-glucosaminyl)asparagine amidase A (PNGase A) in 50 mM ammonium citrate, pH 6.5, overnight at 37 °C. Total N-linked oligosaccharides released by PNGase A were separated from peptides by reverse phase chromatography on prepacked Sep-pak cartridges (Waters) equilibrated with 5% acetic acid. Peptides were retained on the cartridge and eluted with 5% acetic acid in 40% propanol, whereas the oligosaccharides were eluted in the aqueous phase, collected, freeze-dried and, finally, permethylated. Briefly, 0.5–1 mL of dimethyl sulfoxide/NaOH slurry (prepared by grinding five pellets of sodium hydroxide with approximately 3 mL of dimethyl sulfoxide) was added to the oligosaccharide sample followed by 0.5 mL of methyl iodide. The mixture was shaken for 10 min at room temperature and then quenched with 1 mL of distilled water; the products were extracted into chloroform, which was washed several times with distilled water. The chloroform layer was finally dried under a stream of nitrogen and the residue was dissolved in methanol for MALDIMS analysis.

2.6. Phenol determination

10 g of BVPC were suspended in 100 mL methanol and the mixture, stirred for 24 h at room temperature, was finally centrifuged at 10,000 × g for 10 min at 4 °C. The obtained pellet was washed three times with additional 20 mL methanol. The obtained supernatants (160 mL) were pooled and analyzed for both quantitative (unbound phenol content) and qualitative phenol determination. The final pellet was redissolved in distilled water at pH 12 and analyzed for bound phenol content. Both bound and unbound phenol BVPC content was determined using the Folin-Ciocalteau method described by Rocha and Morais (2002). Gallic acid was used to prepare a standard curve to determine the amount of phenols (2–12 mg). Phenols were expressed as mg of gallic acid equivalents/g BVPC.

Unbound phenol identification was carried out by drying 1 mL BVPC methanol extract and resuspending the obtained powder in 0.5 mL of 5% acetic acid solution. The sample was chromatographed through a Sep-Pak C18 classic cartridge (360 mg sorbent per cartridge, 55–105 μm particle size, 50/pk) equilibrated with 5% acetic acid/40% isopropanol and the elution was then carried out with 3 mL of 5% acetic acid/40% isopropanol and 3 mL of isopropanol. The two fractions were dried and the derived powders then resuspended in 0.1% formic acid for MALDIMS analysis. Positive reflectron MALDI spectra were recorded using a MALDI-TOF Voyager-DE STR (Applied Biosystems, Framingham, MA, USA). The MALDI matrices were prepared by dissolving 20 mg of 2,5-dihydroxybenzoic acid (DHB) in 1 mL acetonitrile/water (90:10 v/v). Typically, 1 μL of matrix was applied to the metallic sample plate and 1 μL of analyte was added. Thus, 1 μL of extract was directly mixed with the matrix on the metallic sample plate and MALDI analysis was performed, avoiding any fractionation or enrichment step. The analyses were performed in the positive ion mode, setting the instruments in the reflector mode in the mass range m/z 100–700. A laser pulse voltage power of 3500 V was applied to improve the detection of high molecular weight compounds in the MS spectra. A total of 1500 laser shots at 337 nm were summed for each spectrum. Raw data were analyzed by using the Data Explorer Software (version 4.9, build 115; Applied Biosystems, Framingham, MA, USA) and reported as monoisotopic masses. In all cases, MALDI mass spectra are dominated by cations M+.

2.7. Multi-element determination

Minute sub-samples of BVPC were analyzed by ICPMS. BVPC (1 mg) was suspended in 1 mL of Milli-Q water in a teflon vessel and digested in a microwave oven (Milestone Ethos 900-Mega II) by adding a mixture of 4 mL of 67% HNO_3, 2 mL of 40% HF and 1 mL of 37% HCl. HNO_3 and HCl were Super Purity Solvent grade (Romil, Cambridge, UK). Mineralization was achieved with the following microwave oven program: 20 min to reach 220 °C at 1400 W; 15 min at 220 °C and 1400 W; ventilation for 30 min. The solution was then quantitatively transferred into polystyrene liners and stored at 4 °C until the ICPMS analysis was performed. The analyses were carried out in triplicate on an Agilent 7700 ICP-MS, equipped with a frequency-matching RF generator and 3rd generation Octopole Reaction System (ORS3), operating with helium as cell gas on diluted samples (1:10 v/v Milli-Q water). The parameters were set as follows: radiofrequency power 1550 W, plasma gas flow 14 L min^{-1}; carrier gas flow 0.99 L min^{-1}; He gas flow 4.3 mL min^{-1}. The Octopole Reaction System was activated to improve the metal quantification because of the interferences by polyatomic species produced by a combination of isotopes coming from plasma, reagents and matrix. 103Rh was used as an internal standard (50 mg mL^{-1} final concentration). Multi-element calibration standards were prepared in 5% HNO_3 at four different concentrations (1, 10, 50, and 100 mg L^{-1}). The standard addition approach for calibration on four concentration levels was used in order to keep the matrix-induced variations to a minimum. At least three replicates of each calibration standard were run. Intra-day repeatability was determined by the measurements of sample three times on the same day. Inter-day repeatability was determined by the analysis of a similar digested solution on three different days over a period of

one week. Moreover, in order to correct possible instrumental drifts, 103Rh was added as internal standard to all the samples and calibration standards. Relative SD was calculated for both series of analyses. Reagent blanks run together with matrices.

2.8. Film preparation

BVPC powder (5 g) was dispersed under constant stirring in 100 mL distilled water and, then, glycerol was added to obtain either 30% or 50% (w/w) concentrations. The pH was adjusted to a pH value of 10 and the film forming solutions were heated for 30 min at 80 °C in a water bath under constant stirring. The film forming solutions were then cooled at room temperature, degassed, and finally cast on teflon–coated glass plates (30 × 30 cm) at 30 °C for 24 h. In addition, some BVPC films obtained in the presence of 50% glycerol were laminated with an additional corn zein layer to obtain bilayer films. To this aim, corn zein (2 g) and polyethylene glycol 400 (0.5 g) were mixed in 30 mL ethanol. The mixture was stirred for 10 min at medium speed and then heated at 80 °C until zein was completely dissolved. The solution was then cooled and poured onto a dried BVPC film and allowed to stand for 30 min at 25 °C before drying for 30 min at 85 °C. Finally, the obtained bilayer film was stored overnight at 25 °C before using.

All the films were manually peeled off and, prior to be tested, were conditioned at 50% RH and 25 °C by placing them in a desiccator over a saturated solution of Mg(NO3)26H2O for 48 h. Film thickness was measured with a micrometer (Electronic digital micrometer, DC-516, sensitivity 0.001 mm) at different positions of each film specimen. At least five thickness measurements were performed for each film and the mean values were considered in the different tests.

2.9. Moisture and total soluble matter determination

Three samples (2 × 2 cm) were cut from each film, weighed (±0.0001 g) into aluminum dishes and dried in an oven at 105 °C for 24 h. The moisture content values were determined as the percentage of initial film weight lost during drying. Conversely, to determine total soluble matter (TSM) content, same amounts of film samples were directly immersed in distilled water at 25 °C. After 24 h the film samples were removed, rinsed with distilled water and dried at 105 °C for 24 h. Film initial dry matter values were inferred from moisture content measurements of the same film. TSM percentage of each film was calculated by using the following equation (Tang, Jiang, Wen, & Yang, 2005):

$$\%TSM = \frac{Initial\ dry\ wt. - Final\ dry\ wt.}{Initial\ dry\ wt.} \times 100$$

2.10. Mechanical properties

Following the ASTM Standard Test Method D 882–91 (ASTM, 1996) each film strip (2.5 × 10 cm) was mounted between the grips of the texture analyzer (Zwick 1446–60, Germany) and tested with initial grip separation of 5 cm and cross-head speed of 1 mm/s. Tensile strength (TS) was calculated by dividing the maximum force by the initial cross-sectional area of the specimens. Percentage of elongation at break (EB) was calculated as the percentage of change in the length of the specimen compared to the original length between the grips.

2.11. Water vapor permeability

Film water vapor permeability (WVP) was determined gravimetrically (Bamdad et al., 2006). Circular steel cups with a diameter of 5.1 cm and a depth of 5.4 cm were filled with 3 g of anhydrous calcium chloride and then covered with protein films. Cups were weighed with their contents, then placed in a desiccator containing 1 L of pure water at 22 °C, and finally weighed every 24 h for a week. WVP was calculated according to the following equation (Kaya & Kaya, 2000):

$$WVP = (\Delta m / \Delta t)\ xL / S\Delta P$$

Where $\Delta m / \Delta t$ is the weight of moisture gain per unit of time (g/h), L is the average film thickness (mm), S is the cup test mouth area (m^2) and ΔP is the water vapor pressure difference between the two sides of the film (KPa).

2.12. Scanning electron microscopy

Film cross section was examined by a Phillips XL30 scanning electron microscope (The Netherlands). At first, the film samples were fractured in liquid nitrogen. Then, for cross-sectional analysis, fractured films were rested vertically on the sides of a rectangular aluminum piece and fixed on stubs using double sided adhesive tape, coated with gold (300 s, 20 mA), and finally observed at a magnification of 1000 × . An acceleration potential of 20 kV was used during micrograph.

2.13. Container production

About 0.12–0.18 mm thick-walled containers of the same shape and size (1.5 cm height, 6.0 cm small diameter and 9.0 cm large diameter) were made starting from 5 g of BVPC. The prepared films were moulded around the outside of a round teflon laboratory device, having an outer diameter of 18.0 cm and a height of 17.5 cm, by IR heating and under vacuum suction. The obtained containers were then removed from moulds and stored at room temperature for 18 months. The teflon laboratory device, designed for manufacturing protein-based containers, is illustrated in Fig. 1. The device consists of a heat source, constituted by a IR lamp along with lamp holder connected to power supply 1, a circular film support 2, a suction chamber provided with both a hole 3 through which vacuum is applied and a bowl 4 presenting nine little holes (1 mm) to give concave plate shape to the sucked films (Panel A). The section images of the device was showed in the Panel B of Fig. 1. Protein film samples (30 × 30 cm) were mounted between the grip of the two rings (18 cm diameter) of the device and then fixed by three screws to the circular support. After that heat generated by the IR lamp conferred an appropriate flexibility to the film, vacuum was applied through the suction chamber hole (3.5 cm diameter) by a 2000 W domestic cleaner until the desired shape of the container was obtained. Finally, redundant edges around the container were cut using scissors and the final product was stored at 25 °C.

2.14. Statistical analysis

All determinations were carried out in triplicate and mean values with standard deviations were calculated. The results were analyzed using the SAS package (version 9.1). The least significant difference test was used to describe means at 5% significance level. 3.
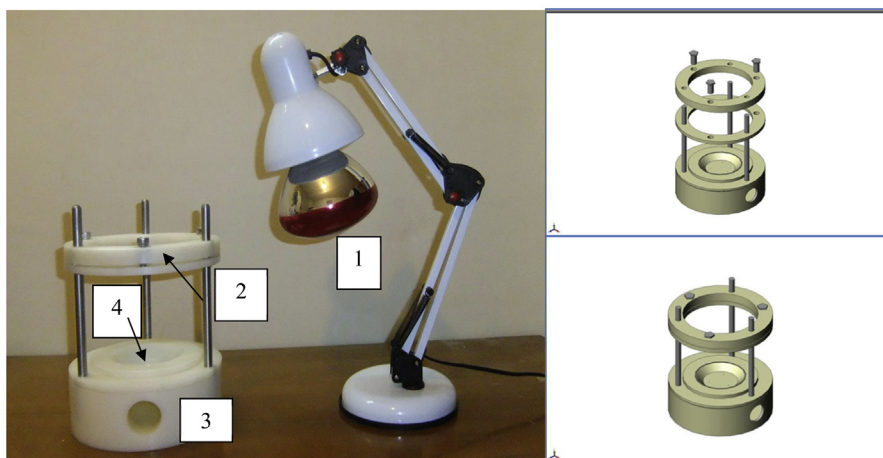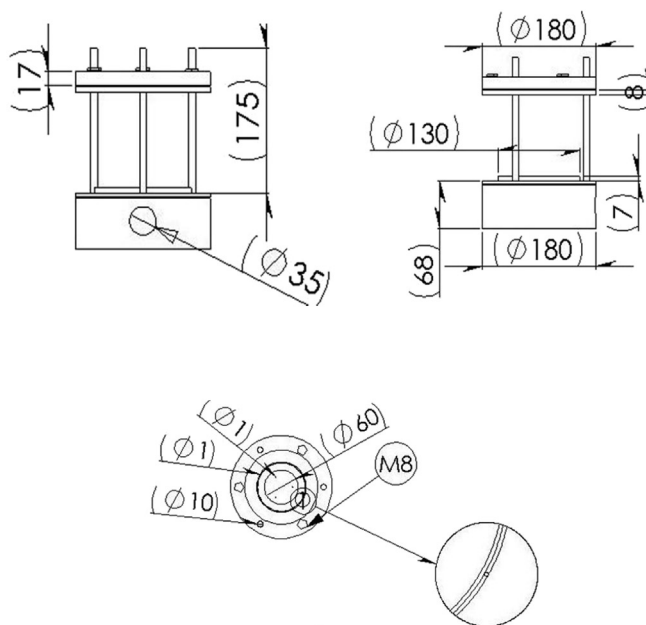
Fig. 1. . Laboratory device for producing protein-derived containers. Panel A: IR lamp (1), circular film support (2), suction chamber hole for vacuum application (3), concave plate (4). Panel B: device sectional images.

3. Results

3.1. BVPC protein identification

The detected protein content values from ground BV seeds (flour), ranging from 27 to 28% (db), were found similar to the amounts previously reported (Haddad, 2006; Pastor-Cavada et al., 2011; Sadeghi, Pourreza, Samei, & Rahmani, 2009). The obtained BVPC, containing 86.45% (db) proteins, was analyzed for protein identification by a preliminary SDS-PAGE separation and staining with colloidal Coomassie Blue. Most intense protein bands were excised from the gels, digested with trypsin and the resulting peptide mixtures fractionated by nano HPLC and sequenced by tandem mass spectrometry generating sequence information on

individual peptides. LCMSMS spectra were used to search for a non-redundant sequence against UniprotSprot databases using the in-house Mascot software, taking advantage of the specificity of trypsin and of the taxonomic category of the samples. The number of measured masses matching within the given mass accuracy was recorded and the proteins showing the highest number of peptide matches were examined leading to the identification of the protein components. As further selection criteria, only the proteins identified by Mascot search with at least three peptides, and found exclusively in the replicates, were selected. The list of proteins identified by this experimental approach (Table 1) showed the occurrence of the same six proteins (convicilin, legumin A2, legumin B, alcohol dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase and allergen Len c 1.0101) in the ten different bands

analyzed (MW between 36 and 63 kDa), probably due to a proteolytic processing of the different proteins.

3.2. BVPC phenol, multi-element, carbohydrate and other organic compound content

It has been recently reported that the total phenol content of BV seed aqueous extract corresponds to about 60 mg gallic acid equivalents/g (Mirzaei, Satarifar, & Zadeh, 2014). Our quantitative determination of the phenol content in BVPC methanol extract (unbound phenols) indicated a value of 1.2 mg gallic acid equivalents/g, whereas the large majority of phenol compounds (20.3 mg gallic acid equivalents/g) was detected in the methanol unsoluble BVPC fraction being, thus, probably protein-bound.

Although MALDIMS is a well known powerful method for the identification of several different molecules, only recently it has been used for the analysis of condensed tannins or simple flavonols. In fact, it has been demonstrated that it is possible to simultaneously detect numerous molecules of this class of compounds in the mass spectra where M+ cations dominate (Carpentieri, Marino, & Amoresano, 2007; Nunes-Miranda et al., 2012). Therefore, BVPC methanol extract was also analyzed by MALDIMS in order to identify the different unbound phenols and Table 2 summarizes the fingerprinting obtained by the analysis carried out on the BVPC methanol soluble fraction. Each peak of the MALDIMS spectrum was assigned to the corresponding species on the basis of literature data and fragmentation spectra, the most abundant compounds resulting 4-methyl-epicatechin-3'-glycoside (m/z, 496.35), malvidin-3-O-glycoside (m/z, 515,23), petunidin-3-(6'-acetyl)-glycoside (m/z, 520.36), spinatoside-4-O-glucuronide (m/z, 522.38) and cyanidin-3-O-(6-O-p-coumaroyl)-pentoside (m/z, 565.18).

Furthermore, ICPMS analysis of BVPC let us to identify and quantify up to fifty-three different metal ions in a single run, even though only thirteen resulted to have a value higher than the ICP limit of detection. Among these, only one element occurred at a value higher than 5 ppm (Mg, 7.17), two elements at values between 1 and 2 ppm (Fe, 1.85; Zn, 1.14) and four at values between 0.1 and 0.5 ppm (Al, 0.50; Cu, 0.24; Mn, 0.18; B, 0.15), whereas only traces of Cr, Cu, Sr, Mo, Ba and Pb were detected.

Monosaccharides and other low MW carbohydrates were instead determined by mild hydrolysis of BVPC sample followed by GCMS analysis of the obtained trimethylsilyl derivatives. The results reported in Table 3 show a high percentage of the monosaccharides glucose (22.1%) and mannose (8.3%), as well as of the disaccharides saccarose (25.3%) and maltose (19.3%). In addition, several other organic compounds like squalene and α −ketobutirric acid (11.2%), as well as tocopherol and oleic alcohol acetate (8%), were identified in BVPC by using a liquid phase extraction procedure coupled to GCMS (Table 3). Additional informations about the possible protein-bound glycosidic structures were obtained by digesting the BVPC sample with trypsin and by releasing the intact N-linked oligosaccharides from the obtained peptides with PNGase A treatment. The glycans were separated from the peptides by a Sep-pack reverse phase chromatographic step, then permethylated and finally analyzed by MALDIMS. The obtained spectra displayied a high heterogeneous glycoform mixture. In particular, two different classes of N-linked glycans were detected: high-mannose structures, representing the most abundant components, and fucosylated complex structures. The m/z of permethylated oligosaccharides released by BVPC and the hypothesized molecular

Table 1
BVPC protein identification by LCMSMS analysis of the main SDS-PAGE protein bands.

	ID NCBI	Protein	MW (kDa)	Number of peptides
BAND 1	gi|164512566	Convicilin	62.7	12
	gi|479104	Legumin B	54.6	15
	gi|126161	Legumin A2	59.3	5
BAND 2	gi|164512566	Convicilin	62.7	45
	gi|479104	Legumin B	54.6	7
	gi|113361	Alcohol dehydrogenase 1	41.1	6
	gi|29539109	allergen Len c 1.0101	47.7	7
	gi|126161	Legumin A2	59.3	5
BAND 3	gi|164512566	Convicilin	62.7	12
	gi|479104	Legumin B	54.6	13
BAND 4	gi|29539109	allergen Len c 1.0101	47.7	15
	gi|164512566	Convicilin	62.7	13
	gi|479104	Legumin B	54.6	5
BAND 5	gi|29539109	allergen Len c 1.0101	47.7	15
	gi|164512566	Convicilin	62.7	14
	gi|479104	Legumin B	54.6	4
	gi|126161	Legumin A2	59.3	3
BAND 6	gi|479104	Legumin B	54.6	10
	gi|126161	Legumin A2	59.3	9
	gi|462138	Glyceraldehyde-3-phosphate dehydrogenase	36.6	8
BAND 7	gi|126161	Legumin A2	59.3	12
	gi|164512566	Convicilin		10
	gi|462138	Glyceraldehyde-3-phosphate dehydrogenase	36.6	8
	gi|479104	Legumin B	54.6	7
	gi|29539109	allergen Len c 1.0101	47.7	8
	gi|113361	Alcohol dehydrogenase 1	41.1	7
BAND 8	gi|29539109	allergen Len c 1.0101	47.7	13
	gi|164512566	Convicilin	62.7	9
	gi|479104	Legumin B	54.6	4
	gi|126161	Legumin A2	59.3	4
BAND 9	gi|29539109	allergen Len c 1.0101	47.7	7
	gi|164512566	Convicilin	62.7	5
	gi|479104	Legumin B	54.6	4
BAND 10	gi|479104	Legumin B	54.6	8
	gi|126161	Legumin A2	59.3	5

Table 2
MALDIMS fingerprinting of phenols occurring in methanol soluble fraction of BVPC[a].

Compound	m/z
Cis-Resveratrol-3-O-glycoside + Na	411.16*
Cis-Resveratrol-3-O-glycoside + K	427.13*
Pelargonidin-3-O-glycoside	433.12*§
Epicatechin-3-O-gallate	441.00*
Cyanidin-3-O-monoglicoside/Kaempferol-3-glycoside	449.13*
Taxifolin-deoxy-hexoside B/A	450.10§
Peonidin-3-O-monoglycoside	463.00*
Delphinidin-3-O-monoglicoside/Quercetin-3-glycoside	465.16*
Isorhamnetin-3-O-glicoside/Quercetin-3-O-glycoside	478.36§
Myricetin-3-glycoside	481.12*
4-Methyl-epicatechin-3'-glycoside	**496.35*§**
Acetone derivate of Peonidin-3-glycoside	501.01*
Delphinidin-3-O-(6-O-acetyl)-monoglycoside	507.30§
Malvidin-3-O-glycoside + Na	**515.23§**
Petunidin-3-(6'-acetyl)-glycoside	**520.36*§**
Spinatoside-4-O-glucuronide	**522.38*§**
Peonidin-3-(6″-malonyl)-glicoside	549.22*
Cyanidin-3-O-(6-O-p-coumaroyl)-pentoside	**565.18***
Crisoeriol-7-O-(6'-malonyl-glycoside)+ Na	570.36§
Delphinidin-3-glicoside-4-vinylphenol adduct	581.17*
Cyanidin-3-O-(6-O-p-coumaroyl)-pentoside + Na	**587.17***
Cyanidin-3-O-monoglicoside-(6-O-p-coumarate)/Petunidin-3-O-glycoside-4-vinylphenol adduct	595.19*
Malvidin-3-O-glycoside Pyruvate/Vitisin A + K	599.44§
Malvidin-3-(6-O-acetyl-glycoside) piruvate	603.14*
Malvidin-Cyanidin (fragment)	617.18*§
Pinotin A/Peonidin-3-(6-O-caffeoyl)glycoside/Peonidin-3,5-O-diglycoside/Malvidin-3-O-glycoside-4-vinyl catechin	625.47§
Delphinidin-3,5-diglucoside	627.22*
Malvidin-Peonidin (fragment)	631.43§
Malvidin-3-O-(6-O-p-coumaroyl) monoglycoside	639.00§
Acetyl Pigment A/Malvidin-3-acetyl-4-vinylphenol adduct	651.19*
Malvidin-3,5-O-diglycoside/Malvidin-caffeoyl-glucoside	655.44§
Malvidin-3-acetyl-glycoside-4-vinylcatechol adduct	667.18*
Acetyl Pigment A/Malvidin-3-acetyl-4-vinylphenol adduct + Na	673.18*
Malvidin-3-acetyl-glycoside-4-vinylguaiacol adduct	681.20*
Malvidin−Malvidin + K (fragment)	685.00§
Acetyl Pigment A/Malvidin-3-acetyl-4-vinylphenol adduct + K	689.16*
Delphinidin -3-O-(6-O-P-Coumaroyl) glycoside Pyruvic Acid	701.40*
Malvidin-3-(6-O-P-Coumaroylglicoside)Pyruvate/A Type Vitisin Of Malvidin-3-P-Coumaroyl glycosilate	707.22§
Delphinidin-3-p- Coumaroyl-glycoside-4- vinylphenol adduct	727.28*
Epicatechin-epicatechin-3-O-galactoside	738.55§
(epi)catechin-Peonidin-3-O-glycoside	751.33*
Delphinidin-catechin-glycoside	753.48§
Catechin-ethyl-cyanidin-glycoside	765.33*
Catechin-Petunidin-glycoside	767.00*
Malvidin-Cyanidin-glycoside/Peonidin-3-glycoside-ethylcatechin	779.23*§
Delphinidin-3-sporoside-5-glycoside	789.29*
Malvidin-3-(6-O-p-coumaroyl)-5-O-diglycoside	801.47§
Malvidin-3-glycoside-4-vinylcatechin	805.26§
Malvidin-Malvidin-glucoside (fragment)/Catechin-Malvidin-acethylglicoside	823.52*
P3/Procyanidin (T2,T3,T4)	866.67§
Malvidin-3-(6-O-p-coumaroyl-glycoside)-epicatechin	927.68§
Malvidin-3-p-coumaroyl-glycoside-4-vinylcatechin	951.37*
Patuletil-3-o-(2″-feruloylglycosyl)-apiosilglycoside	964.33§
Trimer 1G + Na A series[b]	1039.72§
Trimer 1G + Na B series[b]	1041.73§

[a] Phenols eluted by either 5% acetic acid/40% isopropanol (*) or isopropanol (§); the most abundant phenols are reported in bold.
[b] G, galloyl ester; A and B series refer to inter phalanic bonds.

structures of the latter are reported in Table 4.

3.3. Film preparation and properties

The obtained BVPC resulted suitable for the preparation of flexible brown films, sufficiently strong to be easily handled. With the aim to obtain biodegradable finely shaped containers, we prepared three different BVPC films, two of them containing different glycerol concentrations (30 and 50%) whereas the third, containing 50% glycerol, was laminated with a thin layer of zein, the prolamine fraction of corn proteins widely used in bilayer film production. It is known that the amount of plasticizer occurring in the film forming solution is critical for obtaining films with desired flexibility (Vieira, Silva, Santos, & Beppu, 2011) whereas the lamination with another film layer may produce films with improved mechanical and barrier characteristics (Ghanbarzadeh & Oromiehi, 2008, 2009) as a result of the combined properties of the different layers (Rakotonirainy & Padua, 2001; Gonzales, Strumia & Igarzabal, 2011). In fact, zein has a high hydrophobic amino acid content and produces films with lower water vapor permeability and higher tensile strength in comparison with other protein-based films (Cho, Park, & Rhee, 2002; Cho et al., 2010; Ghanbarzadeh et al., 2006; Ghanbarzadeh & Oromiehi, 2009).

Table 5 shows the characteristics of the three different films

Table 3
Low MW carbohydrates and other organic compounds occurring in BVPC[a].

Compound	%
Arabitol	0.9%
Cellobiose	0.9%
Fruttose	0.7%
Galactose	0.6%
Glycerol	0.6%
Glucose	**22.1%**
Maltose	**19.3%**
Mannobiose	0.2%
Mannose	**8.3%**
Myo-inositol	0.6%
NAcetylGalactosamine	0.1%
Pinitol	0.3%
Ribitol	0.3%
Saccarose	**25.3%**
Sorbitol	0.3%
Talose	0.7%
Turanose	0.4%
Dimethylheptane	1.3%
Diossane	3.0%
Isopropyllaurate	4.5%
Metan-3-one semicarbazone	1.5%
Methyl palmitate	0.4%
Methylottane	0.8%
Oleic alchool acetate	**8.6%**
Squalene	**11.1%**
Ter-buthylphenol	0.6%
Tocopherol	**8.0%**
Toluene	1.7%
Trimethyldodecane	0.2%
α −Ketobutirric acid	**11.2%**

[a] The most abundant compounds are reported in bold.

prepared. Moisture content of the BVPC films containing 30% glycerol, as well as that of the zein-laminated films, markedly decreased with respect to the one of the films containing 50% of plasticizer (33% and 45%, respectively). These results were probably due to the hydroscopic properties of glycerol that, by forming hydrogen bonds, is able to effectively retain water molecules into the film matrix (Vieira et al., 2011) while, conversely, the zein high hydrophobic amino acid content contributes to repulse water from the laminated BVPC film network. As a consequence, also the TSM values resulted reduced (of about 25%) both in films containing a lower concentration of glycerol and in the zein-laminated ones depending on their minor hydrophilic characteristics.

Concerning the film mechanical properties, both BVPC films containing 30% glycerol and zein-laminated films exhibited a higher TS (3.5-fold) and a markedly lower EB (20- and 33-fold, respectively) compared to BVPC films prepared in the presence of 50% glycerol (Table 5). Numerous studies showed that high plasticizer concentrations decrease protein film TS and increase their EB by reducing the intermolecular forces along the polypeptide chains (Vieira et al., 2011). In particular, the increase in mechanical resistance and the decrease in deformability of protein films prepared in the presence of lower glycerol concentrations were previously reported (Choi & Han, 2001; Saremnezhad, Azizi, Barzegar, Abbasi, & Ahmadi, 2011). A similar trend was also observed when soy (Cho et al., 2010) or whey protein (Vieira et al., 2011) films, as well as caseinate/whey protein films (Cho et al., 2002), were laminated with corn zein layers.

Moreover, BVPC film containing 30% glycerol and laminated BVPC/zein bilayer film exhibited also a significantly lower WVP (about 25%) with respect to the monolayer BVPC films containing 50% glycerol. Increase in glycerol concentration is known to enhance permeability to water vapor as a consequence of intermolecular space enlargement inside the film network, whereas the

high content of hydrophobic amino acids occurring in zein may explain the more effective water vapor barrier activity observed in laminated bilayer BVPC films.

Finally, also film morphology observed in scanning electron microscopy images shown in Fig. 2 was markedly affected, the different microstructures being able to explain the changes of properties of the three different films.

3.4. Developed containers and stability studies
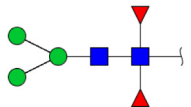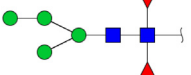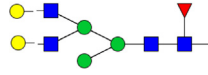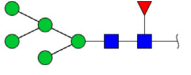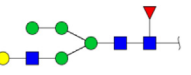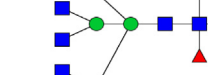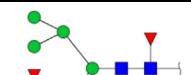
Proteins are considered promising biopolymers to replace petroleum-based commodity plastics, even though they generally exhibit poor mechanical properties regarding processability and end-use application. In fact, protein film brittleness exhibited during thermoformation limits their potential industrial use. To overcome this problem, plasticizers and/or additional film layers could be utilized to provide the needed protein-based film workability. Fig. 3 shows three different representative containers obtained with the three different BVPC films prepared, respectively, at two different glycerol concentrations or after lamination with corn zein of BVPC film containing 50% glycerol, by using a new specific designed device for container manufacture. The containers derived from the BVPC film prepared in the presence of 30% glycerol (Fig. 3, panel B) and the one produced with zein-laminated film (Fig. 3, panel C) were more easily prepared since the originating films showed both resistance and deformability properties more suitable with respect to the BVPC film prepared in the presence of a higher concentration of glycerol (Fig. 3, panel A). Furthermore, the container derived from bilayer film exhibited a stability, shown as structural integrity and mechanical resistance, greater than that of the other two BVPC containers during storage at 25 °C. In fact, no visible changes occurred in the structure and shape of the BVPC/zein bilayer container during 18 months of its storage, while some cracks were evident on the surface of the container prepared with BVPC film obtained in the presence of 30% glycerol following 3 months of storage, and the container derived from 50% glycerol containing film lost its shape only few days after its production because of its excessive flexibility.

4. Discussion

The results of the present study highlight BVPC, easily obtainable from BV seeds, as a promising raw material candidate for the production not only of films but also of biodegradable containers. BV seed proteins were preliminarily identified and BVPC carbohydrate, metal ion, phenolic and other organic compound content also determined. Convicilin, legumins A2 and B, allergen Len c 1.0101 and two oxydoreductase, alcohol dehydrogenase 1 and glyceraldehyde-3-phosphate dehydrogenase, were identified by analysing the ten main protein bands obtained by SDS-PAGE of the BVPC sample. The first four proteins generally occur in the seeds of many leguminous and non-leguminous plants and are known as globulins functioning as protein storage, whereas the two enzymes are certainly related to the embryo development. Moreover, the performed analyses showed the presence in the BVPC of a high content of phenol compounds, the large majority of which was detected in the methanol unsoluble fraction being, thus, probably covalently bound to macromolecules during BVPC preparation. An interesting result was the determination of two different classes of N-linked glycans, i.e. high mannose and fucosylated complex structures, All these findings suggest the occurrence in the BVPC of a complex macromolecular network, probably constituted by polysaccharide/protein complexes with different phenol compounds covalently linked, potentially suitable for the preparation of films with improved mechanical features. The films obtained by

Table 4
Hypothesized structures and m/z of permethylated oligosaccharides released from BVPC by PNGase A treatment.

● Man= Mannose / ◐ Gal = Galactose / ■ HexNAc= N-Acetyl Hexosamine / ▼ Fuc= Fucose

Structure	m/z	Structure	m/z
$Fuc_2Man_3HexNAc_2$	1498.14	$FucMan_3HexNAc_4 Gal_2$	2466.88
$Fuc_2Man_4HexNAc_2$	1702.27	$FucMan_3HexNAc_5 Gal_2$	2520.86
$FucMan_5HexNAc_2$	1731.37	$FucMan_3HexNAc_6 Gal_2$	2712.07
$Man_3HexNAc_5$	1906.42	$Fuc_4Man_3HexNAc_4 Gal_2$	2766.06
$FucMan_4HexNAc_3 Gal$	1976.54	$Fuc_2Man_3HexNAc_7 Gal$	2927.08
$Man_3HexNAc_5 Gal$	2110.54	$FucMan_3HexNAc_7 Gal_2$	2957.18
$Man_3HexNAc_5 Gal_2$	2314.72	$Fuc_2Man_3HexNAc_7 Gal_2$	3131.23
$Fuc_2Man_5HexNAc_3 Gal$	2354.92	$Fuc_2Man_3HexNAc_7 Gal_3$	3335.34

such new low-cost biological source was found to possess, when added with 30% glycerol, as well as with 50% glycerol and then laminated by zein, a high TS and a low EB, becoming capable to be processed to obtain moulded solid containers. In addition, the obtained bilayer BVPC/zein container exhibited a marked stability since no change in its structure and shape and no cracks were observed during 18 months of storage at 25 °C. Moreover, a further result of the present study was the manufacture of a laboratory

Table 5
Properties of different BVPC films used for container preparation*.

Film	Thickness (mm)	Moisture content (%)	WVP (gmm/mhKPa)	TSM (%)	TS (MPa)	EB (%)
BVPC + 50% glycerol	0.10 ± 0.02[a]	27.69 ± 1.18[a]	0.71 ± 0.02[a]	36.57 ± 2.08[a]	5.04 ± 0.81[b]	118.01 ± 18.71[a]
BVPC + 30% glycerol	0.11 ± 0.02[a]	18.60 ± 0.30[b]	0.52 ± 0.06[a]	27.52 ± 0.45[b]	17.82 ± 1.69[a]	5.79 ± 0.64[b]
BVPC/zein + 50% glycerol	0.14 ± 0.04[a]	15.42 ± 0.34[c]	0.55 ± 0.03[a]	27.33 ± 0.96[b]	18.03 ± 0.95[a]	3.54 ± 0.59[b]

*Values superscripted with dissimilar letters (a,b,c) are significantly different (p < 0.05).
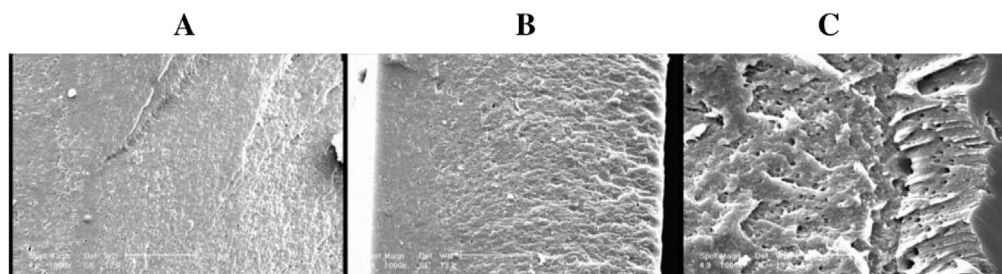
Fig. 2. . Scanning electron microscope images of BVPC films prepared in the presence of 30% glycerol (A), 50% glycerol (B) or 50% glycerol and then zein-laminated (C). Experimental details are given in the text.
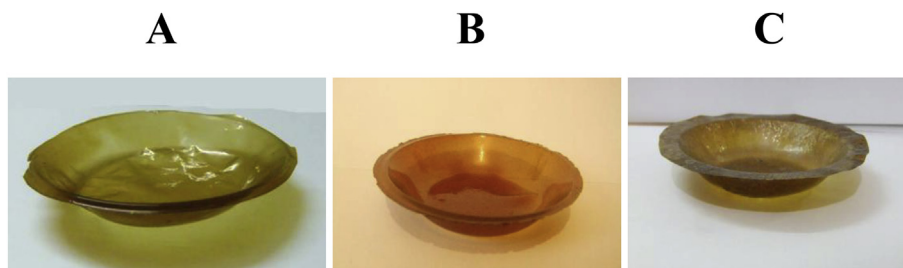
Fig. 3. . Representative BVPC-based bioplastic containers. The containers were obtained by using BVPC films prepared in the presence of 50% (A) or 30% (B) glycerol, or by lamination with corn zein of BVPC films prepared in the presence of 50% glycerol (C). Experimental details are given in the text.

device able to transform protein films in well shaped containers without using the conventional thermoplastic processing techniques generally unsuitable to mould protein-based materials. In fact, it has been demonstrated that BVPC-derived films can be easily transformed into small bowls through a new simple device equipped with only a tunable IR heater lamp and a vacuum unit. However, the obtained biodegradable protein-derived containers were made by hand and, therefore, had sidewalls that were not as uniform as would be expected for those of machine-moulded containers obtainable from a commercial manufacturer. Slight variations in thickness of the walls, in fact, may worsen the structural properties and longevity of the obtained containers, while the commercially fabricated ones might have better features simply due to their uniformity. However, further strategies to increase container stability deserve to be explored, as well as the feasibility of producing BVPC-based bioplastic containers by either modifying standard moulding techniques or refining the device described in the present paper.

In conclusion, although the obtained container made of BV seed protein-based sustainable plastics represents only a promising prototype, the reported results should be considered the first insight concerning the possibility of using BV proteins to fabricate biodegradable articles. Future efforts will be mainly addressed to structurally improve the product with the aim to confer to it the characteristics of durability typical of conventional plastic articles suitable for different uses, and there are several possibilities for improving the BVPC container performance. One of these is applying a coating on its inner walls for reducing the rate of its

degradation. In fact, if the container is layered with a protective material, BV protein degradation might be certainly delayed. Moreover, further research should be addressed to examine how the addition of either plasticizers other than glycerol (Vieira et al., 2011), nanoparticles (Kadam et al., 2013) or chemical/enzymatic cross-linking agents (Porta, Di Pierro, et al., 2011; Porta, Mariniello, et al., 2011) may affect the properties of such protein-based bioplastics as well as container longevity and relative costs. Other physical changes could also be determined by the use of fillers to create new composite materials. Corn stover, as well as lignin and cellulosic fibres, are among the numerous possible additives to incorporate into BVPC films to improve the structural properties of the derived containers. Being many of these potential components underutilized by-products of other processes, their recycle might also partially contribute to offset the current high price of the existing protein-based biodegradable containers.

5. Conclusions

BVPC was easily obtained, analyzed for its protein, carbohydrate, multi-element, phenolic and other organic compound content, and then utilized in manufacturing of films and biodegradable containers by using a new designed and realized plastic moulding equipment. The reported results showed that either the addition of 30% glycerol to BVPC film forming solution or corn zein lamination of bitter vetch protein-based films gave rise to potentially satisfactory vacuum thermoformed containers with acceptable resistance and stability. These findings open new perspectives in using

BV proteins as a sustainable alternative to fossil fuel based plastics to produce a variety of properly shaped biodegradable articles.

Acknowledgements

References

AACC, American Association of Cereal Chemists. (2003). *Approved methods of AACC.* St. Paul, Minnesota. USA: The Association.

Anderson, T. J., & Lamsal, B. P. (2011). Zein extraction from corn products and co-products and modifications for various applications: A review. *Cereal Chemistry, 88,* 159–173.

Arabestani, A., Kadivar, M., Shahedi, M., Goli, S. A. H., & Porta, R. (2013). Properties of a new protein film from bitter vetch (Vicia ervilia) and effect of CaCl₂ on its hydrophobicity. *International Journal of Biological Macromolecules, 57,* 118–123.

Arcan, I., & Yemenicioglu, A. (2011). Incorporating phenolic compounds opens a new perspective to use zein films as flexible bioactive packaging materials. *Food Research International, 44,* 550–556.

ASTM, American Society for testing and materials. (1996). *Standard Test Methods for Tensile Properties of Thin Plastic Sheeting, D* (pp. 882–891).

Bamdad, F., Goli, A. H., & Kadivar, M. (2006). Preparation and characterization of proteinous film from lentil (Lens culinaris): edible film from lentil (Lens culinaris). *Food Research International, 39,* 106–111.

Borreani, G., & Tabacco, E. (2015). Bio-based biodegradable film to replace the standard polyethylene cover for silage conservation. *Journal of Dairy Science, 98,* 386–394.

Carpentieri, A., Marino, G., & Amoresano, A. (2007). Rapid fingerprinting of red wines by MALDI mass spectrometry. *Analytical and Bioanalytical Chemistry, 389,* 969–982.

Choi, W. S., & Han, J. H. (2001). Physical and mechanical properties of pea-protein-based edible films. *Journal of Food Science, 66,* 319–322.

Cho, S. Y., Lee, S. Y., & Rhee, C. (2010). Edible oxygen barrier bilayer film pouches from corn zein and soy protein isolate for olive oil packaging. *LWT-Food Science and Technology, 43,* 1234–1239.

Cho, S. Y., Park, J. W., & Rhee, C. (2002). Properties of laminated films from whey powder and sodium caseinate mixtures and zein layers. *LWT-Food Science and Technology, 35,* 135–139.

Cuq, B., Gontard, N., & Guilbert, S. (1998). Proteins as agricultural polymers for packaging production. *Cereal Chemistry, 75,* 1–9.

Di Pierro, P., Rossi Marquez, G., Mariniello, L., Sorrentino, A., Villalonga, R., & Porta, R. (2013). Effect of transglutaminase on the mechanical and barrier properties of whey protein/pectin films prepared at complexation pH. *Journal of Agricultural and Food Chemistry, 61,* 4593–4598.

Di Pierro, P., Sorrentino, A., Mariniello, L., Giosafatto, C. V. L., & Porta, R. (2011). Chitosan/whey protein film as active coating to extend Ricotta cheese shelf-life. *LWT-Food Science and Technology, 44,* 2324–2327.

Ghanbarzadeh, B., & Oromiehi, A. R. (2008). Biodegradable biocomposite films based on whey protein and zein: barrier, mechanical properties and AFM analysis. *International Journal of Biological Macromolecules, 43,* 209–215.

Ghanbarzadeh, B., & Oromiehi, A. R. (2009). Thermal and mechanical behavior of laminated protein film. *Journal of Food Engineering, 90,* 517–524.

Ghanbarzadeh, B., Oromiehi, A. R., Musavi, M., Emam-Jomeh, Z., Razmi Rad, E., & Milani, J. (2006). Effect of plasticizing sugars on rheological and thermal properties of zein resins and mechanical properties of zein films. *Food Research International, 39,* 882–890.

Giosafatto, C. V. L., Di Pierro, P., Gunning, P., Mackie, A., Porta, R., & Mariniello, L. (2014). Characterization of Citrus pectin edible films containing transglutaminase-modified phaseolin. *Carbohydrate Polymers, 106,* 200–208.

González, A., Strumia, M. S., & Igarzabal, C. I. A. (2011). Cross-linked soy protein as material for biodegradable films: synthesis, characterization and biodegradation. *Journal of Food Engineering, 106,* 331–338.

Grewell, D., Schrader, J., & Srinivasan, G. (2014). Soy-based chemicals and materials, in: Developing protein-based plastics. *ACS Symposium Series, 1178,* 357–370.

Haddad, S. G. (2006). Bitter vetch grains as a substitute for soy bean meal for growing lambs. *Livestock Science, 99,* 221–225.

Helgeson, M. S., Graves, W. R., Grewell, D., & Srinivasan, G. (2009). Degradation and nitrogen release of zein-based bioplastic containers. *Journal of Environmental Horticulture, 27,* 123–127.

Kadam, D. M., Thunga, M., Wang, S., Kessler, M. R., Grewell, D., Lamsal, B. P., et al. (2013). Preparation and characterization of whey protein isolate films reinforced with porous silica coated titania nanoparticles. *Journal of Food Engineering, 117,* 133–140.

Kaya, S., & Kaya, A. (2000). Microwave drying effects on properties of whey protein isolate edible films. *Journal of Food Engineering, 43,* 91–96.

Larbi, A. A. M., El-Moneim, A., Nakkoul, H., Jammal, B., & Hassan, S. (2011). Intra-species variations in yield and quality determinants in Vicia species:1. bitter vetch (Vicia ervilia L. *Animal Feed Science and Technology, 165,* 278–287.

Mariniello, L., Di Pierro, P., Esposito, C., Sorrentino, A., Masi, P., & Porta, R. (2003). Preparation and mechanical properties of edible pectin-soy flour films obtained in the absence or presence of transglutaminase. *Journal of Biotechnology, 102,* 191–198.

Mirzaei, A., Satarifar, M., & Zadeh, F. B. (2014). Antioxidant activities of Iranian bitter vetch in different methods. *World Journal of Pharmacy and Pharmaceutical Sciences, 3,* 186–193.

Monsoor, M. A., & Yusuf, H. K. M. (2002). In vitro protein digestibility of lathyrus pea (Lathyrus satirus), lentil (Lens culinaris), and chickpea (Cicera rietinum). *International Journal of Food Science and Technology, 37,* 97–99.

Nunes-Miranda, J. D., Santos, H. M., Reboiro-Jato, M., Fdez-Riverola, F., Igrejas, G., Lodeiro, C., et al. (2012). Direct matrix assisted laser desorption ionization mass spectrometry-based analysis of wine as a powerful tool for classification purposes. *Talanta, 91,* 72–76.

Pastor-Cavada, E., Juan, R., Pastor, J. E., Alaiz, M., & Vioque, J. (2011). Nutritional characteristics of seed proteins in 28 Vicia species (Fabacease) from southern Spain. *Journal of Food Science, 76,* 1118–1123.

Peelman, N., Ragaert, P., De Meulenaer, B., Adons, D., Peeters, R., et al. (2013). Application of bioplastics for food packaging. *Trends in Food Science and Technology, 32,* 128–141.

Porta, R., Di Pierro, P., Rossi-Marquez, G., Mariniello, L., Kadivar, M., & Arabestani, A. (2015). Microstructure and properties of bitter vetch (Vicia ervilia) protein films reinforced by microbial transglutaminase. *Food Hydrocolloids, 50,* 102–107.

Porta, R., Di Pierro, P., Sorrentino, A., & Mariniello, L. (2011). Promising perspectives for transglutaminase in "bioplastics" production. *Journal of Biotechnology and Biomaterials, 1,* 3. http://dx.doi.org/10.4172/2155-952X.1000102e.

Porta, R., Mariniello, L., Di Pierro, P., Sorrentino, A., & Giosafatto, C. V. L. (2011). Transglutaminase crosslinked pectin- and chitosan-based edible films: a review. *Critical Reviews in Food Science and Nutrition, 51,* 223–238.

Rakotonirainy, A. M., & Padua, G. W. (2001). Effects of lamination and coating with drying oils on tensile and barrier properties of zein films. *Journal of Agricultural and Food Chemistry, 49,* 2860–2863.

Razavi, S. M. A., Amini, A. M., & Zahedi, Y. (2015). Characterization of a new biodegradable edible film based on sage seed gum: influence of plasticizer type and concentration. *Food Hydrocolloids, 43,* 290–298.

Reisi, K., Zamani, F., Vatankhah, M., & Rahimiyan, Y. (2011). Effect of row and soaked bitter vetch (Vicia ervilia) seeds as replacement protein source of cotton seed meal on performance and carcass characteristics of Lori-Bakhtiari fattening ram lambs. *Global Veterinaria, 7,* 405–410.

Rocha, A. M. C. N., & Morais, A. M. B. B. (2002). Polyphenoloxidase activity and total phenolic content as related to browning of minimally processed 'Jonagored' apple. *Journal of the Science of Food and Agriculture, 82,* 120–126.

Rossi Marquez, G., Di Pierro, P., Esposito, M., Mariniello, L., & Porta, R. (2014). Application of transglutaminase-crosslinked whey protein/pectin films as water barrier coatings in fried and baked foods. *Food and Bioprocess Technology, 7,* 447–455.

Sadeghi, G. H. (2011). Effect of cooking and methionine supplementation on nutritional value of bitter vetch seed as a feed ingredient for broilers. *Tropical Animal Health and Production, 43,* 259–264.

Sadeghi, G. H., Pourreza, J., Samei, A., & Rahmani, H. (2009). Chemical composition and some anti-nutrient content of raw and processed bitter vetch (Vicia ervilia) seed for use as feeding stuffin poultry diet. Tropical animal health and production. *41,* 85–93.

Saki, A. A., Pourhesabi, G., Yaghobfar, A., Mosavi, M. A., Tabatabai, M. M., & Abbasinezhad, M. (2008). Effect of different levels of the raw and processed vetch seed (Vicia sativa) on broiler performance. *Journal of Biological Sciences, 8,* 663–666.

Saremnezhad, S., Azizi, M. H., Barzegar, M., Abbasi, S., & Ahmadi, E. (2011). Properties of a new edible film made of faba bean protein isolate. *Journal of Agricultural Science and Technology, 13,* 181–192.

Shimao, M. (2001). Biodegradation of plastics. *Current Opinions in Biotechnology, 12,* 242–247.

Shukla, R., & Cheryan, M. (2001). Zein: the industrial protein from corn. *Industrial Crops and Products, 13,* 171–192.

Smith, R. (2005). *Biodegradable polymers for industrial applications* (pp. 1–535). Woodhead Publ. Limited and CRC Press LLC.

Song, F., Tang, D.-L., Wang, X.-L., & Wang, Y.-Z. (2011). Biodegradable soy protein isolate-based materials: A review. *Biomacromolecules, 102,* 3369–3380.

Tang, C. H., Jiang, Y., Wen, Q. B., & Yang, X. Q. (2005). Effect of transglutaminase treatment on the properties of cast films of soy protein isolates. *Journal of Biotechnology, 120,* 296–307.

Tang, C. H., Xiao, M. L., Chen, Z., & Yang, X. Q. (2011). Properties of transglutaminase-treated red bean protein films. *Journal of Applied Polymer Science, 122,* 789–797.

Vieira, M. G. A., Silva, M. A., Santos, L. O., & Beppu, M. M. (2011). Natural-based plasticizers and biopolymer films: A review. *European Polymer Journal, 47,* 254–263.

Formyl peptide receptor 1 suppresses gastric cancer angiogenesis and growth by exploiting inflammation resolution pathways

Nella Prevete, Federica Liotti, Anna Illiano, Angela Amoresano, Piero Pucci, Amato de Paulis & Rosa Marina Melillo

View supplementary material

Submit your article to this journal

View Crossmark data

Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

Formyl peptide receptor 1 suppresses gastric cancer angiogenesis and growth by exploiting inflammation resolution pathways

Nella Prevete [a,b], Federica Liotti[c], Anna Illiano[d], Angela Amoresano[d], Piero Pucci[d], Amato de Paulis[a], and Rosa Marina Melillo [b,c]

aDipartimento di Scienze Mediche Traslazionali, University of Naples "Federico II," Naples, Italy; bIstituto di Endocrinologia ed Oncologia Sperimentale del CNR "G. Salvatore," Naples, Italy; cDipartimento di Medicina Molecolare e Biotecnologie Mediche, University of Naples "Federico II," Naples, Italy; dDipartimento di Scienze Chimiche, University of Naples "Federico II," Naples, Italy

ABSTRACT

Chronic inflammation can result from inadequate engagement of resolution mechanisms, mainly accomplished by specialized pro-resolving mediators (SPMs) arising from the metabolic activity of lipoxygenases (ALOX5/15) on ω-6 or ω-3 essential polyunsaturated fatty acids (PUFA). We previously demonstrated that formyl peptide receptor 1 (FPR1) suppresses gastric cancer (GC) by inhibiting its inflammatory/angiogenic potential. In this study, we asked whether FPR1 exploits inflammation resolution pathways to suppress GC angiogenesis and growth.

Here, we demonstrate that genetic or pharmacologic modulation of FPR1 in GC cells regulated ALOX5/15 expression and production of the SPMs Resolvin D1 (RvD1) and Lipoxin B4 (LXB4). SPM treatment of GC cells abated their angiogenic potential. Genetic deletion of ALOX15 or of the RvD1 receptor GPR32 increased the angiogenic and tumorigenic activity of GC cells thereby mimicking FPR1 loss. Deletion/inhibition of ALOX5/15 or GPR32 blocked FPR1-mediated anti-angiogenic activities, indicating that ALOX5/15 and GPR32 are required for FPR1's pro-resolving action. An ω-3- or ω-6-enriched diet enforced SPM endogenous production in mice and inhibited growth of shFPR1 GC xenografts by suppressing their angiogenic activity. These data implicate that FPR1 and/or pro-resolving pathway components might be used as risk/prognostic markers for GC; ω-6/3-enriched diets, and targeting FPR1 or SPM machinery may be exploited for GC management.

Introduction

Prolonged inflammation underlies the pathogenesis of various diseases including cancer.[1] Recent years have seen a paradigm shift in our understanding of the etiopathogenesis of inflammation whereby chronic inflammation may not result only from persistence of inflammatory mechanisms, but may also arise from an inadequate engagement of the resolution mechanisms.[2] Resolution is an active process brought about by a series of cellular mechanisms that are activated to restore tissue homeostasis.[3] Lipid specialized pro-resolving mediators (SPMs), which are the main effectors of resolution, derive from the effects exerted by lipoxygenases (ALOX5 and 15) on ω-6 arachidonic acid (AA), or on ω-3 eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), which are essential polyunsaturated fatty acids (PUFA)[3] (Fig. 1A). Recent evidence associates a deficit in pro-resolving pathways with the establishment of various inflammatory-related disorders (e.g., asthma, fibrosis, and autoimmune diseases).[4] However, whether and how defects in resolution mechanisms affect cancer initiation and progression, and what signals control pro-resolving pathways in this context is unknown.

Formyl peptide receptors (FPR1, 2, and 3) are pattern recognition receptors (PRR) of the G-protein-coupled (GPCR) family that recognize both exogenous and endogenous "danger" signals, and trigger inflammation and immune responses.[5] FPRs can also trigger inflammation resolution, depending on the environmental context and on the specific ligand.[6] We recently showed that genetic ablation of FPR1 in AGS gastric cancer (GC) cells, which constitutively express high levels of the receptor, increased their angiogenic and tumorigenic potential. Accordingly, enforced expression of FPR1 in MKN45 GC cells, which constitutively express low levels of the receptor, dramatically impaired GC cell xenograft angiogenesis and growth in immunodeficient mice.[7]

Since genetic deletion of FPR1 in GC cells increased angiogenesis and enhanced the response to pro-inflammatory cytokines,[7] which is a phenotype suggestive of unresolved inflammation, we asked whether, at the gastric level, FPR1 might actively sustain pro-resolving pathways to inhibit GC angiogenesis and growth.

Here, we show that, in GC cells, FPR1 expression/activation levels directly correlate with ALOX expression, and with SPM Resolvin D1 (RvD1) and Lipoxin B4 (LXB4) (3) production (Fig. 1A). The increased angiogenic potential of GC cells lacking FPR1, previously reported,[7] was reverted by RvD1 or LXB4 treatment. Genetic ablation of ALOX15 or of GPR32 (an RvD1 receptor)[8] induced a pro-angiogenic phenotype in GC cells similar to that induced by FPR1 deletion. ALOXs and GPR32 are required for FPR1-mediated anti-angiogenic activity in GC cells. Consistently, administration of ω-3 or ω-6 PUFA-enriched diets, which enforces endogenous production of SPMs,[9] inhibited xenograft growth of FPR1-silenced GC cells by ablating their angiogenic activity.
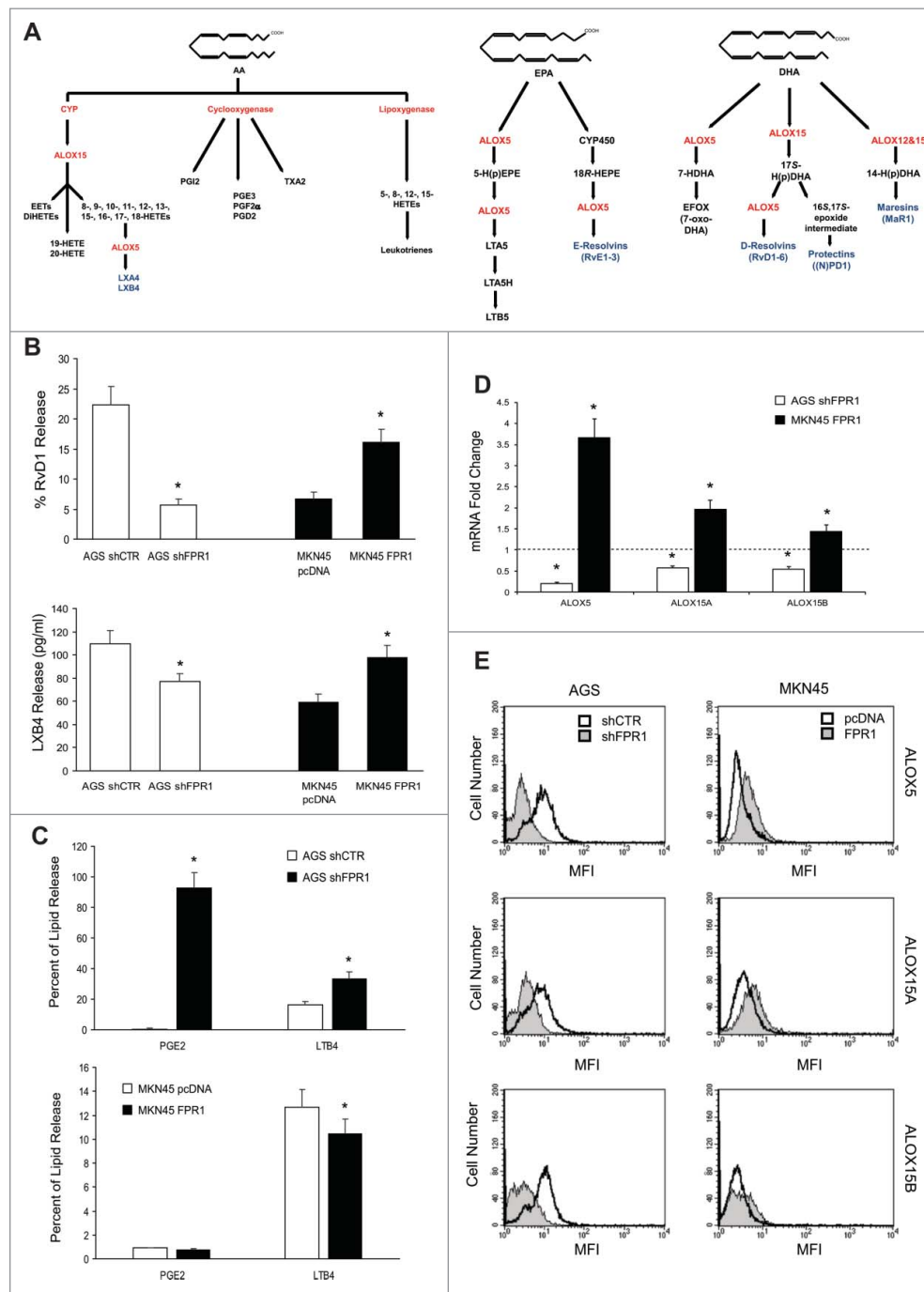
Figure 1. Relation between FPR1 expression and SPM biosynthesis machinery. (A) Schemes of arachidonic acid-AA, docosahexaenoic acid-DHA, and eicosapentaenoic acid-EPA metabolism. (B) AGS shFPR1 cl 15 cells produced significantly lower amounts of RvD1 and LXB4 compared to AGS shCTR cells, while MKN45 cells overexpressing FPR1 (MKN45 FPR1 cl 4) released higher amounts of RvD1 and LXB4 compared to empty vector transfected cells (MKN45 pcDNA), as evaluated by EIA assays. Data are represented as mean $\pm$ SD of five independent experiments. $^{*}p < .05$ compared to the relative control. (C) Increased release of PGE2 and LTB4 from shFPR1 AGS cells and from MKN45 pcDNA cells compared to the relative control cells, assessed by EIA. Data are represented as mean $\pm$ SD of three independent experiments. $^{*}p < .05$ compared to the relative control. (D) AGS shFPR1 cl 15 expressed significantly lower levels and MKN45 FPR1 cl 4 significantly higher levels of ALOX5, ALOX15A, and ALOX15B mRNAs compared to relative controls (dotted line), as assessed by real-time PCR. Data are represented as mean $\pm$ SD of three independent experiments. $^{*}p < .05$ compared to the relative control. (E) ALOX5, ALOX15A, and ALOX15B protein levels were lower in AGS shFPR1 vs. AGS shCTR, and in MKN45 pcDNA vs. MKN45 FPR1 cells, as evaluated by cytofluorimetric analysis. One representative experiment out of three is shown.

Our data indicate that FPR1 signaling activates a pro-resolving program in GC cells that inhibits angiogenesis and growth.

Results

FPR1 controls ALOX5 and ALOX15 expression and the production of SPMs in GC cells

To study whether pro-resolving pathway components are involved in the FPR1-mediated anti-angiogenic and tumor suppressor activity of GC, we used our previously generated GC cell lines, namely, FPR1-silenced AGS (AGS shFPR1) and MKN45 ectopically expressing FPR1 (MKN45 FPR1) or their relative controls (AGS shCTR and MKN45 pcDNA).[7]

RvD1 and LXB4 levels were significantly lower in AGS shFPR1 cells than in AGS shCTR cells. Consistently, MKN45 cells overexpressing FPR1 released higher amounts of RvD1 and LXB4 than did empty vector-transfected cells (Fig. 1B). RvD1 and LXB4 syntheses were reduced in shFPR1 but not in shFPR2 or shFPR3 AGS cells (not shown), which indicates that FPR1 plays a non-redundant role in controlling SPM production, as already observed for
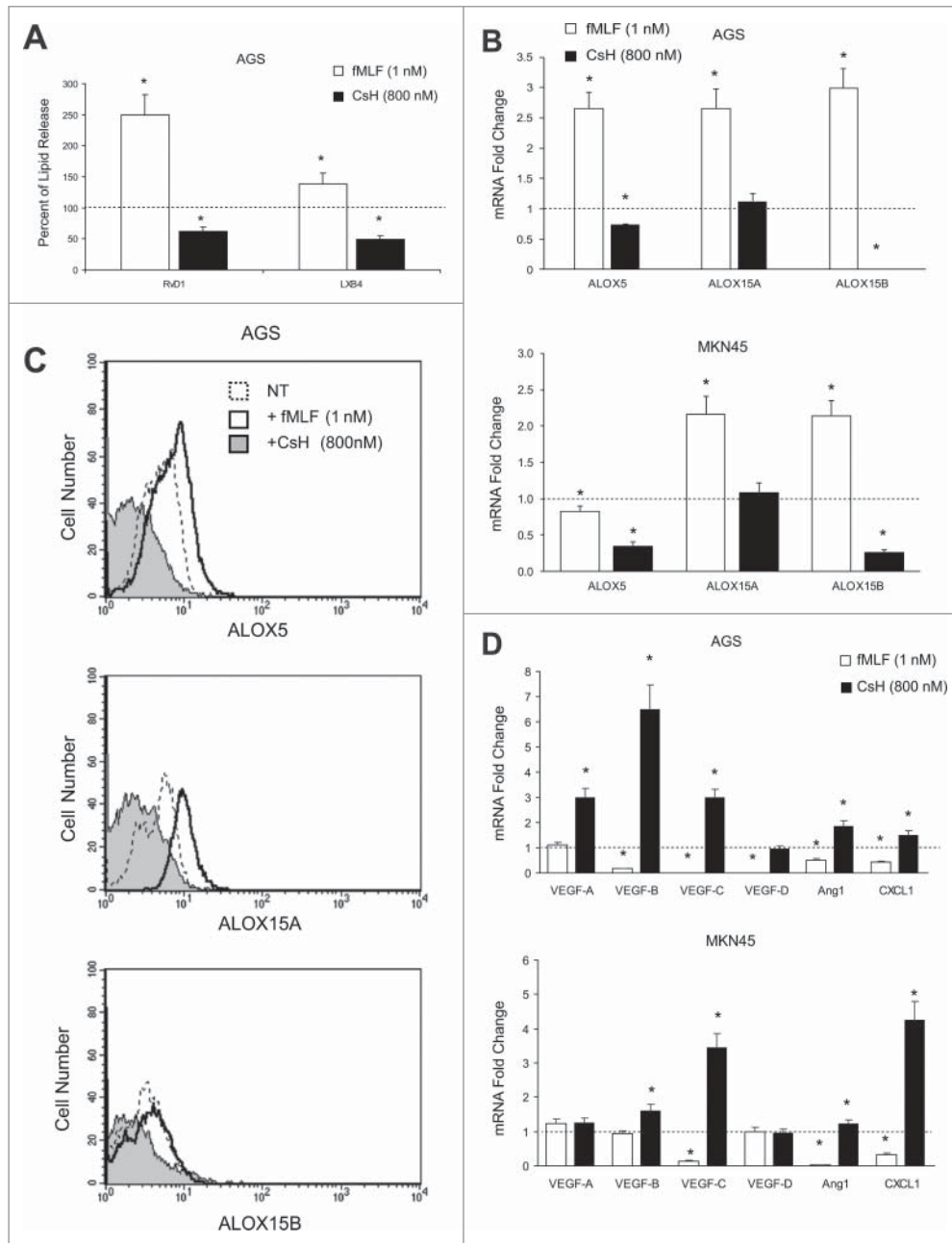
Figure 2. Effects of FPR1 pharmacologic modulation on SPM biosynthesis. (A) EIA assays showing that fMLF (agonist to FPR1-10^{-9} M) treatment of AGS significantly increases RvD1 and LXB4 release compared to untreated cells (dotted line). In contrast, CsH (inverse agonist to FPR1-800 nM) treatment significantly reduced RvD1 and LXB4 release compared to controls (dotted line). Data are represented as mean ± SD of three independent experiments. *p < .05 compared to untreated cells. (B) fMLF induced, whereas CsH inhibited, ALOX5, ALOX15A, and ALOX15B mRNA expression in AGS and MKN45 cells. Data are represented as mean ± SD of three independent experiments. *p < .05 compared to untreated cells (dotted line). (C) fMLF induced, whereas CsH inhibited, ALOX5, ALOX15A, and ALOX15B protein expression in AGS cells, as evaluated by cytofluorimetric analysis. One representative experiment out of three is shown. (D) fMLF inhibited, whereas CsH induced, pro-angiogenic molecule mRNAs expression in AGS and MKN45 cells. Data are represented as mean ± SD of three independent experiments. *p < .05 compared to untreated cells (dotted line).

its tumor suppressor function.[7] We evaluated the lipidomic profiles of controls and FPR1-depleted/overexpressing GC cell supernatants by liquid chromatography tandem-mass spectrometry (LC-MS/MS) (Fig. S1). Compared to controls, AGS shFPR1 cells produced significantly lower levels of SPMs, including 18-HEPE, a resolvin precursor, and LXB4 (Fig. S1A). MKN45 FPR1 cells released significantly higher levels of RvD3 and LXB4 than did control cells (Fig. S1A). LXA4 levels did not differ significantly among GC cells expressing different levels of FPR1 (Fig. S1B). SPM expression parallels the downregulation of pro-inflammatory eicosanoids.[4] Consistently, AGS shFPR1 cells synthesized significantly higher amounts of LTB4 and PGE2 compared to shCTR cells, whereas MKN45 FPR1 cells released lower levels of PGE2 compared to empty vector transfected cells, as assessed by EIA (Fig. 1C).

The levels of the enzymes involved in SPM synthesis (ALOX5, ALOX15A, and ALOX15B)[10,11] were significantly lower in FPR1-silenced GC cells than in shCTR cells, both at mRNA (Fig. 1D) and protein (Fig. 1E) level. The expression of SPM receptors GPR32, ChemR23, and BLT1[8] was lower in AGS shFPR1 cells than in controls (Fig. S1C and D). On the other hand, MKN45 FPR1 cells produced significantly higher levels of pro-resolving enzymes (Fig. 1D and
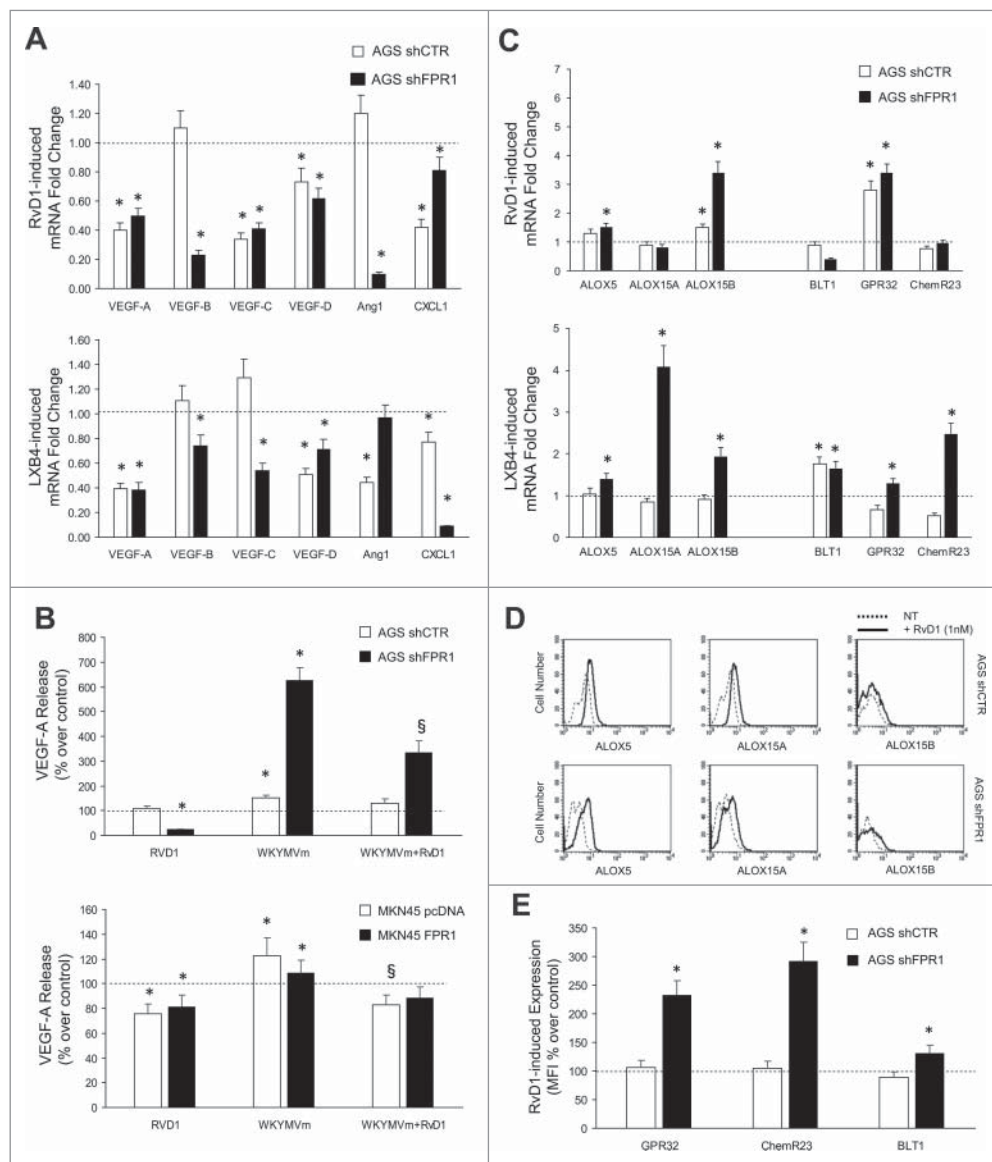
Figure 3. Anti-angiogenic effects of SPMs in GC. (A) RvD1 (1 nM) and LXB4 (1 nM) treatment significantly reduced mRNA expression of pro-angiogenic molecules (VEGF-A, -B, -C, -D, Ang1, and CXCL1) in AGS shCTR and shFPR1 cells. Data are represented as mean ± SD of three independent experiments. $^*p < .05$ compared to the relative untreated control (dotted line). (B) Reduction of spontaneous and WKYMVm-induced VEGF-A release upon RvD1 treatment of the indicated GC cells. Data are represented as mean ± SD of three independent experiments. $^*p < .05$ vs. the relative untreated control (dotted line). (C) RvD1 and LXB4 treatment significantly induced the mRNA expression of the enzymes (ALOX5, ALOX15A, and ALOX15B) and receptors (BLT1, GPR32, and ChemR23) involved in pro-resolving pathways in AGS shCTR and shFPR1 cells. Data are represented as mean ± SD of three independent experiments. $^*p < .05$ vs. the relative untreated control (dotted line). (D) RvD1 treatment significantly induced ALOX5, ALOX15A, and ALOX15B protein expression in AGS shCTR and shFPR1 cells, as evaluated by cytofluorimetric analysis. One representative experiment out of three is shown. (E) RvD1 treatment significantly induced GPR32, ChemR23, BLT1 protein expression in AGS shCTR and shFPR1 cells, as assessed by cytofluorimetric analysis. Data are represented as mean ± SD of three independent experiments. $^*p < .05$ compared to the relative untreated control (dotted line).

E) and of SPM receptors (Fig. S1C and D). These data were reproduced on one more clone and a mass population of both AGS shFPR1 and MKN45 FPR1 GC cells (Fig. S2).

To approach the FPR1-SPMs association from a pharmacologic standpoint, we treated AGS and MKN45 cells with fMLF (10^{-9} M), which is an FPR1 agonist, or with cyclosporine H (CsH 800 nM), which is an FPR1 inverse agonist[12] and verified their ALOX5/15 expression, SPM release, and synthesis of pro-angiogenic mediators (VEGF-A, -B, -C, -D, Ang1, and CXCL1). The release of RvD1 and LXB4 was significantly higher in AGS treated with fMLF (12 h) than in untreated cells. In contrast, RvD1 and LXB4 release was significantly lower in cells treated with CsH (12 h) than in controls (Fig. 2A). fMLF induced, whereas CsH reduced, ALOX5, ALOX15A, and ALOX15B mRNA expression in AGS and MKN45 cells (Fig. 2B). These observations were confirmed at the protein level (Fig. 2C). mRNA levels of pro-angiogenic molecules were significantly reduced by fMLF and induced by CsH treatment (Fig. 2D).

SPMs inhibit the production of pro-angiogenic factors and restore the expression of ALOX5/15 and of SPM receptors in GC cells

To assess whether SPMs (RvD1 and LXB4) block the pro-angiogenic activity of GC cells, we treated or not AGS cells with RvD1 (1 nM) or LXB4 (1 nM), and evaluated the levels of pro-angiogenic factor mRNAs. RvD1 and LXB4 suppressed pro-angiogenic factor mRNA expression in GC cells (Fig. 3A). RvD1 treatment of GC cells significantly reduced VEGF-A release both in basal conditions and upon WKYMVm stimulation (pro-inflammatory FPR2/3 ligand) (Fig. 3B)[7] Similar results were obtained with LXB4 (not shown). This effect was significantly more efficient in GC cells expressing low levels of FPR1 (AGS shFPR1 and MKN45 pcDNA) (Fig. 3A and B). Interestingly, RvD1 treatment restored, and in some cases increased above basal levels, the expression of ALOX5, ALOX15A, ALOX15B, and the SPM receptors (GPR32, ChemR23, and BLT1) at
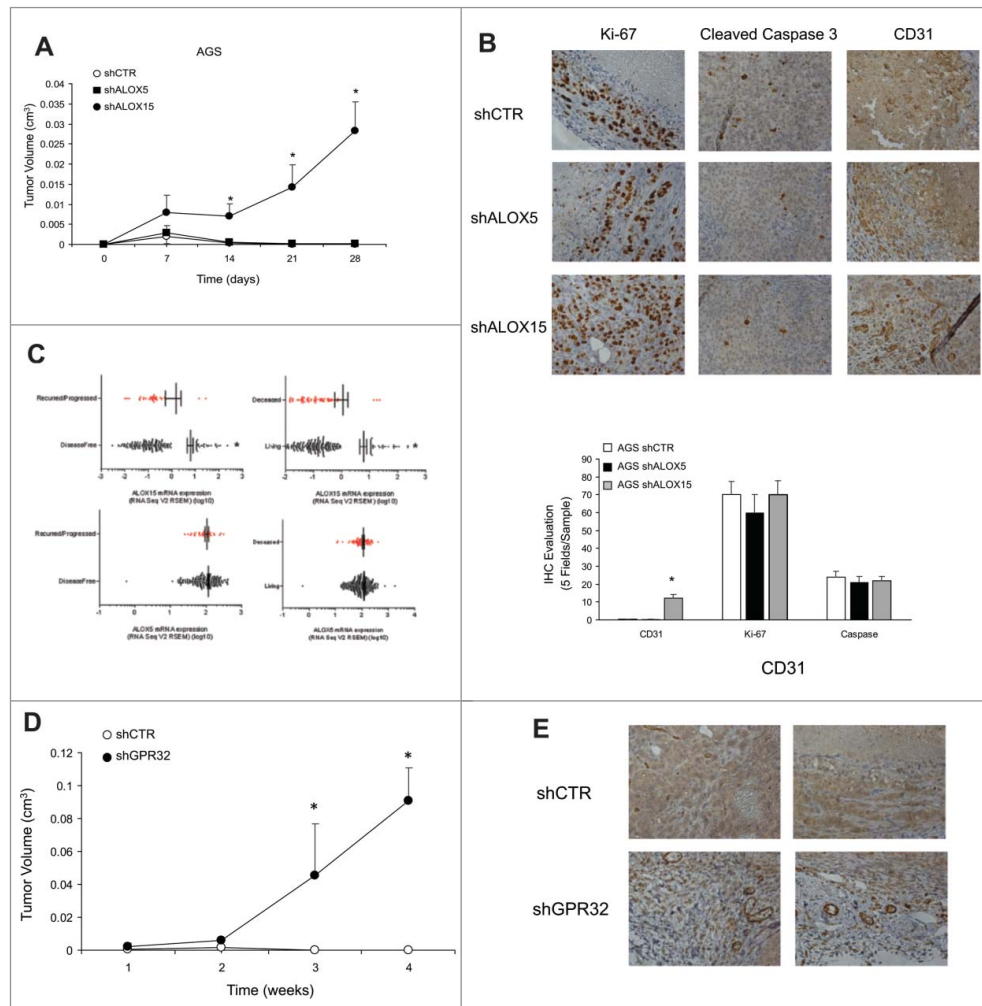
Figure 4. ALOXs and GPR32 involvement in GC angiogenic response. (A) Tumor growth curves of AGS shCTR, shALOX5 (three clones) and shALOX15 (three clones) xenografts in immunodeficient mice. *$p < .05$ compared to shCTR xenografts. (B) Representative images and quantification (five fields/sample) of the proliferation index (Ki-67), vessel density (CD31), and apoptotic rate (Cleaved Caspase 3), assessed by immunohistochemistry, of shCTR, shALOX5, and shALOX15 cell xenografts harvested 28 d post-inoculation. *$p < .05$ compared to shCTR xenografts. (C) ALOX15 and ALOX5 mRNA expression levels of 295 patients affected by gastric adenocarcinoma stratified for disease-free and overall survival status. *$p < .05$ between the two groups. (D) Tumor growth curves of AGS shCTR and shGPR32 xenografts (average of three clones) cells in immunodeficient mice *$p < .05$ compared to shCTR xenografts. (E) Representative images of the vessel density (CD31), assessed by immunohistochemistry, of shCTR and shGPR32 AGS cell xenografts harvested 28 d post-inoculation.

mRNA (Fig. 3C) and protein level (Fig. 3D and E) thereby triggering a positive feed-forward loop. These responses of AGS shCTR to RvD1 were significantly less efficient than that of AGSshFPR1 cells (Fig. 3C). Similar results were obtained with LXB4 (not shown).

ALOX15- and GPR32-silencing increase angiogenesis and tumorigenesis of GC cells thereby mimicking the FPR1-silencing phenotype

To evaluate whether SPM pathway components are necessary to control the angiogenic activity of GC cells, we stably transfected AGS cells with vectors expressing shRNAs targeting ALOX5, ALOX15, GPR32, or control non-targeting shRNAs (shCTR). ALOX5-silencing induced a specific reduction of ALOX5 mRNA levels in AGS cells. However, AGS shALOX5 cells overexpressed ALOX15A and ALOX15B mRNAs (Fig. S3A). Transfection of shRNAs targeting ALOX15 significantly reduced ALOX15A and ALOX15B levels in AGS cells, but did not modify ALOX5 mRNA (Fig. S3A). The effects of shRNAs were confirmed at the protein level (Fig. S3B). shALOX15 and, to a lesser extent shALOX5, impaired AGS cell growth and survival, as did FPR1 silencing[7] (not shown). shALOX5 and shALOX15 AGS cells (three clones/each) constitutively produced lower amounts of RvD1 and increased amounts of PGE2 and LTB4 versus controls, as did shFPR1 cells. LXB4 release was reduced in shALOX15, but not in shALOX5 cells (Fig. S3C). GPR32 and ChemR23 protein expressions were lower in AGS shALOX5 and shALOX15 cells than in controls (not shown). AGS shALOX5 and shALOX15 cells constitutively synthesized higher levels of VEGF-C and VEGF-D mRNAs and released increased amounts of VEGF-A than did controls (Fig. S3D).

To evaluate their tumorigenic potential, we injected AGS shCTR cells or ALOX-silenced clones subcutaneously in athymic mice. AGS shCTR and shALOX5 tumors reached a volume of about 1 mm³ at day 7, and regressed thereafter (Fig. 4A). In contrast, AGS shALOX15 cells formed larger non-regressing tumors, i.e., 2.8 mm³ (Fig. 4A), again replicating the phenotype of FPR1-silenced cells.[7] The tumor growth advantage of shALOX15 cells was due to their higher vessel density compared to controls (Fig. 4B). No differences were detected in Ki-67⁺ and cleaved caspase 3⁺ cells, evaluated far from the necrotic areas, among shCTR, shALOX5, and shALOX15 cell xenografts (Fig. 4B). Our *in vivo* observations are corroborated by data, publicly available in the cBioPortal for Cancer Genomics database (http://www.cbioportal.org),[13,14] reporting the RNAseq analysis of 295 human GC samples. These data show that enhanced mRNA expression of ALOX15, but not of ALOX5, was significantly associated with overall survival status and disease-free status[15] in GC patients (Fig. 4C).

To analyze the effects of GPR32 silencing on the GC cell phenotype, we selected various AGS shGPR32 clones. The low levels of GPR32 were associated with a significant reduction of growth ability and with an increased apoptotic rate of GC cells in culture (not shown). The reduced level of GPR32 in AGS GC cells caused a significant reduction of ALOX5 and ALOX15A/B expression levels

(Fig. S3E) and an increased expression of pro-angiogenic mediators at mRNA (Fig. S3E) and protein (Fig. S3F) level vs. shCTR cells. RvD1 did not reduce angiogenic factor mRNAs in shGPR32 cells (Fig. S3G). When xenotransplanted into nude mice, AGS shGPR32 formed significantly larger tumors than formed by shCTR cells (Fig. 4D). Consistently, the vessel density of shGPR32 xenografts was significantly higher than that of controls (Fig. 4E). No differences were detected in Ki-67⁺ and cleaved caspase 3⁺ cells in shCTR and shGPR32 xenografts (not shown).

FPR1-mediated suppression of angiogenesis in GC cells requires ALOX5, ALOX15, and GPR32

To assess whether FPR1-mediated suppression of angiogenesis requires ALOX5 and/or ALOX15, we stimulated with fMLF (10^{-9} M) GC cells in which ALOX expression/activity was abolished. As shown in Fig. 5A, fMLF significantly reduced the mRNA levels of pro-angiogenic factors (VEGF-B, VEGF-C, and CXCL1) in AGS shCTR, but not in shALOX5 and shALOX15 AGS cells. Similar results were obtained in AGS cells in which ALOXs were pharmacologically inhibited by 10 μM nordihydroguaiaretic acid-(NDGA) (Fig. 5B).[16] These observations suggest that ALOX5 and ALOX15 are required to trigger the anti-angiogenic activity of FPR1 in GC cells.

To verify the requirement of the RvD1 pathway for FPR1-mediated anti-angiogenic activity, we stimulated parental GC cells with fMLF in the presence or absence of a GPR32-neutralizing antibody. fMLF (10^{-9} M) induced ALOX5 and ALOX15A mRNA expression and concomitant VEGF-B and CXCL1 mRNA downregulation in AGS. These effects were reverted by the GPR32-blocking antibody (5 μg/mL) (Fig. 5C). Similar results were obtained in MKN45 cells (Fig. S4A). Consistently, fMLF did not induce ALOXs or suppress pro-angiogenic mediators in AGS shGPR32 cells (Fig. 5D), although shCTR and shGPR32 cells released comparable amounts of RvD1 both in basal conditions and upon fMLF stimulation (not shown). We then assessed the involvement of FPR2, the other RvD1 receptor.[8] fMLF stimulation of AGS shFPR2 cells induced ALOX5 and ALOX15A mRNA expression, and VEGF-B and CXCL1 downregulation, and these effects were reverted by the anti-GPR32 blocking antibody (Fig. 5E). Taken together, these results indicate that RvD1, by binding GPR32, but not FPR2, mediates FPR1's anti-angiogenic effects in GC cells.

We then asked which signaling pathways were involved in the anti-angiogenic response of GC cells to RvD1 and whether differences can be detected in FPR1-lacking cells compared to controls. We found that Akt, MAPK, JNK, and p38 signaling pathways were induced in GC cells treated with RvD1 (1 nM) for different times (15 min–12 h) without significant differences in the activation kinetic and strength between AGS shCTR and shFPR1 cells, or MKN45 pcDNA and FPR1 cells (Fig. 5F). In contrast, RvD1 treatment induced a significant STAT3 and SRC activation only in AGS shFPR1 and MKN45 pcDNA cells expressing lower levels of FPR1 compared to the relative control cells (Fig. 5F). RvD1-mediated reduction of VEGF-

A release was reverted by pre-treatment of AGS shFPR1 cells with the two STAT3 inhibitors FLLL31 (10 μM) and 5–15DPP (15 μM)[17,18] (Fig. 5F). Similar results were not detectable in AGS shCTR cells (Fig. 5F). STAT3 inhibition reverted RvD1 effects on CXCL1 and Ang1 mRNAs levels (Fig. S4B). Instead, when MAPK [U0 126 (25 μM)] or Akt [LY294002 (15 μM)] inhibitors[7] were used, no effects on VEGF-A release were observed (Fig. S4C).

ω-6 and ω-3 diets reduce the tumorigenic potential of shFPR1 GC cells in vivo

Dietary interventions based on differential PUFA ω-3 or ω-6 intake affect AA-, EPA-, and DHA metabolism and the derived SPM endogenous production in mice (Fig. 1A).[9] To assess the effects exerted by ω-6- or ω-3-enriched diets on GC cell growth, we xenotransplanted AGS shCTR or shFPR1 cells in
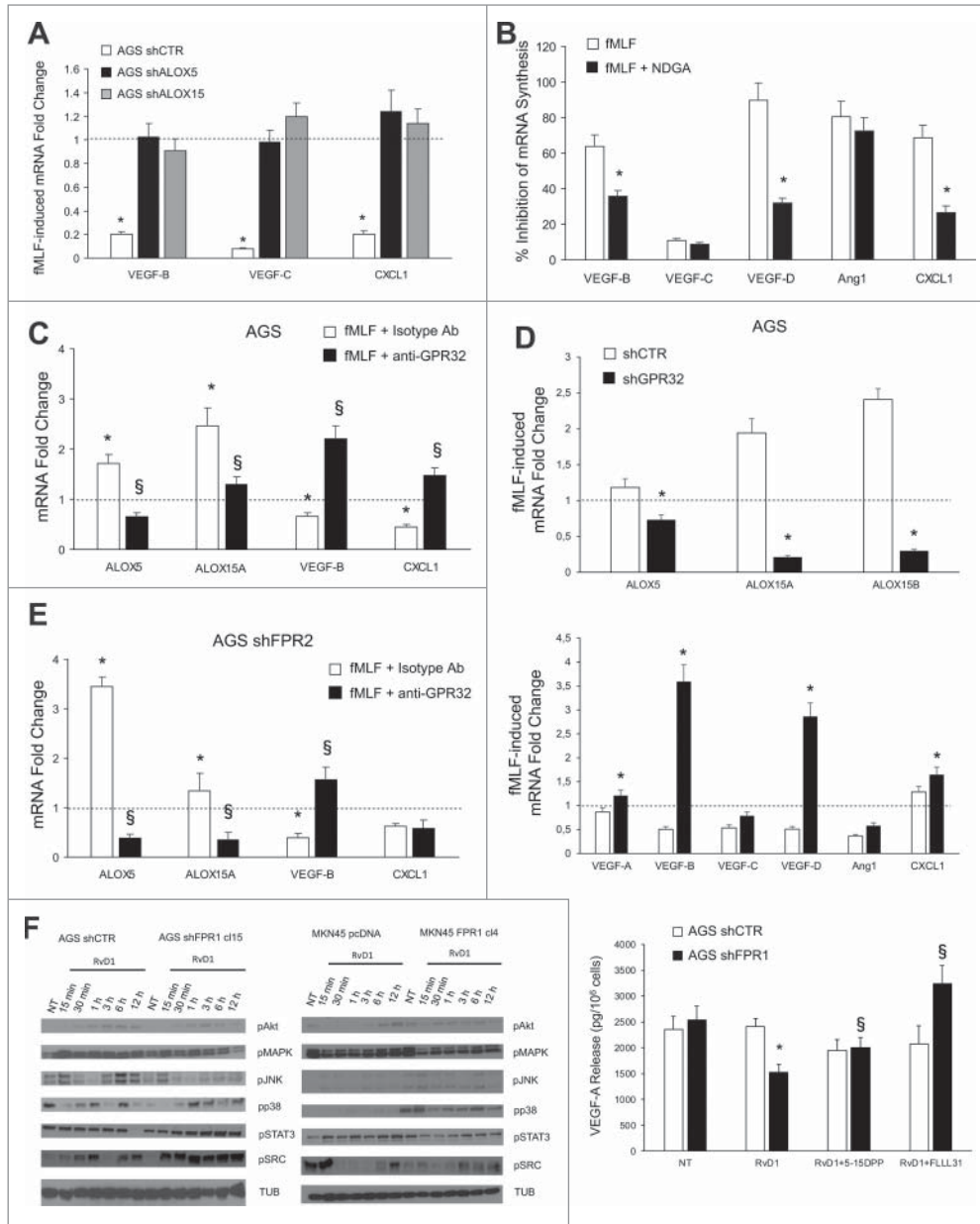
Figure 5. ALOXs and GPR32 are required for FPR1 tumor suppressor role. (A) VEGF-B, -C, and CXCL1 mRNA synthesis was inhibited in shCTR, but not in shALOX5 or shA-LOX15 AGS cells upon fMLF (10^{-9} M) treatment. Data are represented as mean ± SD of three independent experiments. $^*p < .05$ vs. the relative untreated cells (dotted line). (B) Pre-treatment of AGS cells with NDGA (ALOX inhibitor) reverted the ability of fMLF to inhibit the mRNA synthesis of pro-angiogenic molecules. Data are represented as mean ± SD of three independent experiments. $^*p < .05$ compared to fMLF treated cells. (C) In AGS cells, fMLF induced ALOX5 and ALOX15A mRNA overexpression and concomitant VEGF-B and CXCL1 mRNA downregulation. These effects were reverted by a neutralizing GPR32 antibody. An isotype-matched antibody was used as a control. Data are represented as mean ± SD of three independent experiments. $^*p < .05$ compared to untreated cells (dotted line). $^§p < .05$ compared to isotype-matched control. (D) fMLF significantly induced ALOX5, ALOX15A, and ALOX15B mRNA levels and significantly reduced the mRNA levels of pro-angiogenic mediators in AGS shCTR, but not in shGPR32 cells. Data are represented as mean ± SD of three independent experiments. $^*p < .05$ compared to shCTR cells. (E) A neutralizing GPR32 antibody inhibited fMLF-induced ALOX5 and ALOX15A mRNA overexpression and concomitant VEGF-B and CXCL1 mRNA downregulation in AGS shFPR2 cells. An iso-type-matched antibody was used as a control. Data are represented as mean ± SD of three independent experiments. $^*p < .05$ compared to untreated cells (dotted line). $^§p < .05$ compared to isotype-matched control. (F) Activation kinetics of AKT, MAPK, JNK, p38, STAT3, and SRC in AGS shCTR, AGS shFPR1, MKN45 pcDNA, MKN45 FPR1 cells assessed by western blot for their phosphorylated forms. Two STAT3 inhibitors (5-15 DPP and FLLL31) reverted the anti-angiogenic effects of RvD1 in AGS shFPR1 cells. Data are represented as mean ± SD of three independent experiments. $^*p < .05$ compared to untreated cells (NT). $^§p < .05$ compared to RvD1 treated cells.

immunocompromised mice randomly fed a classic, an ω-6/ω-3 balanced (grape seed-colza oils 50/50%, CTRL), an ω-6 enriched (grape seed oil 100%, ω-6), or an ω-3 enriched (colza/fish oils 80/20%, ω-3) diet, 2 weeks before and 4 weeks after xenotransplantation. No differences were detected between animals fed a classic diet and those fed a balanced diet (not shown) demonstrating that a 5% PUFA-enriched diet did not change the tumorigenic potential of GC cells. As shown previously,[7] AGS shCTR cells formed small tumors that regressed in mice fed a balanced, ω-3-, or ω-6-enriched diet, without statistical differences among groups (Fig. 6A). In contrast, AGS shFPR1 cells formed tumors that progressively increased in size (Fig. 6A).[7] Both ω-3 and ω-6-enriched diets significantly suppressed AGS shFPR1 xenograft growth. However, growth inhibition was significant at day 7 post-injection in ω-6-fed mice vs. day 14 in ω-3-fed mice (Fig. 6A). The administration of an ω-3 or ω-6 enriched diet to AGS shFPR1 xenotransplanted mice abolished statistical differences in tumor growth curves between shFPR1 and shCTR cells.

Vessel density was higher in shFPR1 xenografts than in shCTR xenografts (Fig. 6B); ω-6 and ω-3 diets reverted this difference (Fig. 6B). Vessel density was unchanged in shCTR xenografts irrespective of the diet administered (Fig. 6B). Both ω-6 and ω-3 diets significantly reduced Ki-67 staining of shFPR1, and, to a lesser extent, of shCTR excised tumors (Fig. 6C). Diet composition did not affect apoptotic rate as assessed by cleaved caspase-3 evaluation (not shown).

To investigate the effect of dietary intervention on mouse serum lipid composition, we profiled fatty acids by LC-MS/MS. Concentrations of AA were significantly higher in mice fed an ω-6 diet than in those fed a balanced or an ω-3 diet (Table 1). Consistently, EPA ad DHA concentrations were higher in ω-3-fed mice than in the other two groups of mice (Table 1). The concentrations of LXB4 were significantly higher in the sera of ω-6-fed mice than in controls. Concentrations of RvD1, RvD3, and RvE2 were significantly higher in the serum of mice fed an ω-3 diet than in controls. Interestingly, the ω-3 diet significantly reduced PGE2, LTB4, and TXB2 serum concentrations vs. controls (Table 1). Lipid serum concentrations did not differ among mice fed the same diet but xenotransplanted with shCTR or shFPR1 AGS cells (not shown).
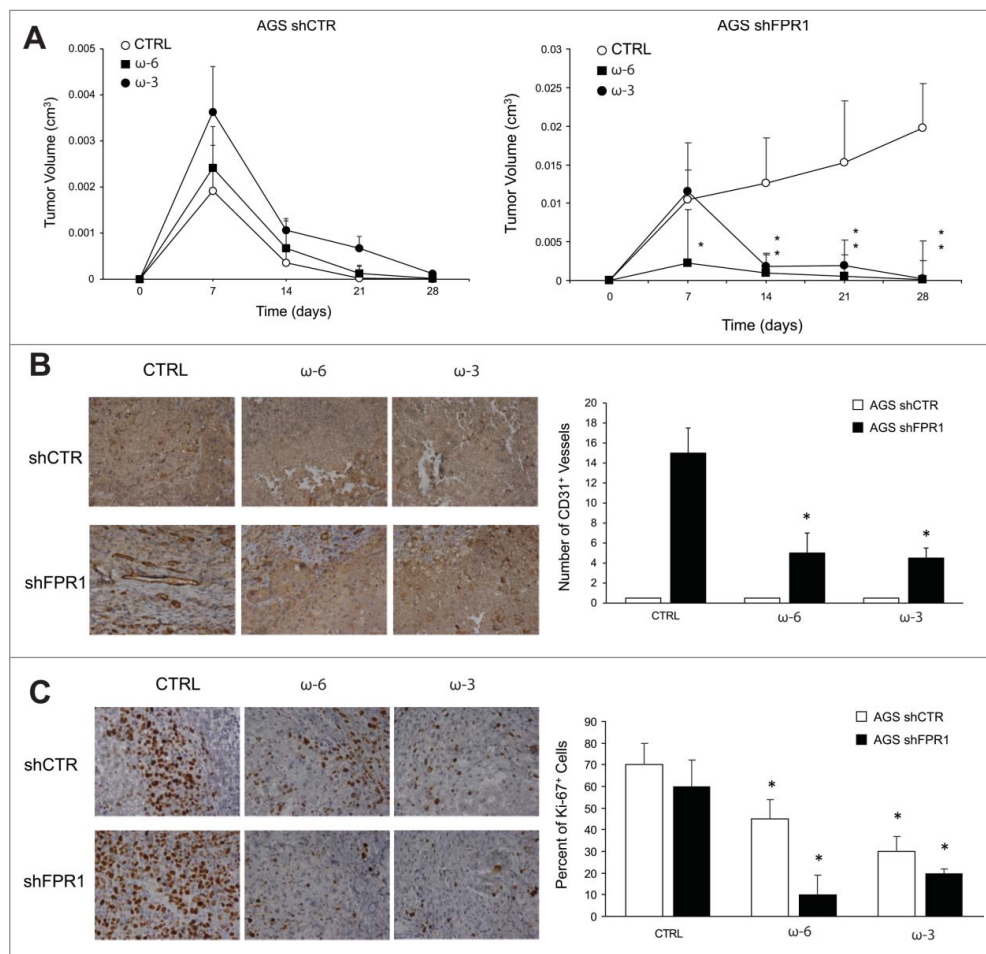
Figure 6. Effects of ω-6 and/or ω-3 increased consumption on GC growth/angiogenesis in a xenotransplantation mouse model. (A) Tumor growth of AGS shCTR and shFPR1 xenografts in CD1 nu/nu mice fed an ω-6/ω-3 balanced (grape seed/colza oils 50/50%, CTRL), an ω-6 enriched (grape seed oil 100%, ω-6), or an ω-3 enriched (colza/fish oils 80/20%, ω-3) diet. *p < .05 vs. CTRL diet. (B) Vessel density (CD31) assessed by immunohistochemistry of shCTR and shFPR1 cell xenografts harvested 28 d post-inoculation from mice fed the three diets as in panel A. Representative images and the relative quantifications (five fields/sample) are shown. *p < .05 vs. the relative CTRL diet. (C) Proliferation index (Ki-67) assessed by immunohistochemistry of shCTR and shFPR1 cell xenografts harvested 28 d post-inoculation from mice fed with the three diets. Representative images and the relative quantifications (five fields/sample) are shown. *p < .05 vs. the relative CTRL diet.

To assess whether the effects exerted by an ω-3 or ω-6 diet on the proliferation rate of GC xenografts was a direct event or whether it was indirectly due to differences in tumor vascularization levels, we studied the effects of RvD1 (1 nM), LXB4 (1 nM), AA (20 μM), DHA (20 μM), and EPA (20 μM) on AGS shCTR and shFPR1 cell proliferation in vitro. We found no difference in BrdU incorporation between treated and untreated cells at the concentrations of lipids above mentioned (Fig. S5). AA, DHA, and EPA significantly reduced GC cell viability when used at higher concentrations (50 μM), as already reported by others.[19-21] However, these concentrations were never reached in vivo in diet-fed mice.[9,22] (Table 1).

Discussion

The role of FPR1 in cancer is complex. In certain cancer types, FPR1 exerts tumor-promoting activities by stimulating the motility and growth of cancer cells[23-26] However, in mouse liver and colon, FPR1- or FPR2-genetic deletion abolishes epithelial homeostasis, increases inflammation after injury and, in colonic mucosae, increases tumorigenesis.[27,28] Recent studies showed that decreased FPR1 concentrations in tumor-infiltrating antigen-presenting cells compromise chemotherapy-induced antitumor immunity.[29,30] In human GC, high FPR1 expression, evaluated by immunohistochemistry, has been correlated with advanced tumor stage and poor overall survival.[31] However, an FPR1 polymorphism linked to reduced FPR1 activity has been associated with an increased risk of stomach cancer in humans.[32,33]

Here, we provide evidence that FPR1, expressed by GC cells, plays a tumor suppressor role by constitutively activating an inflammation resolution program that includes the expression of ALOX5/15, SPMs (RvD1 and LXB4), and SPM receptors (BLT1, ChemR23, GPR32). FPR1 loss/inhibition suppresses this program and induces an inflammatory/angiogenic phenotype in GC cells. Interestingly, the GC-associated FPR1 polymorphism was also associated with periodontitis,[34] a disease whose pathogenesis has been linked to defective pro-resolving pathways and in which SPMs exert protective effects.[35-38]

We demonstrate that the anti-angiogenic potential of SPMs, already reported in other contexts,[39] was more evident in cells expressing low levels of FPR1, despite their lower levels of SPM receptors (BLT1, GPR32, and ChemR23), vs. controls. This effect might be due to constitutively high levels of RvD1 and LXB4 produced by control cells, which could saturate/desensitize their receptors thereby rendering these cells less responsive to exogenous SPM stimulation.

Various studies have recognized that ALOXs exert tumor-promoting or tumor-suppressive activities.[40] Our study introduces a new concept in this field, namely, that another component of pro-resolving pathways, i.e., the Resolvin D1 receptor GPR32,[8] can exert tumor suppressor activity in GC. The LXB4 receptor remains unidentified; consequently, we were unable to investigate its role in GC. We demonstrate that the ALOX enzymes and GPR32 are induced by FPR1 and necessary for FPR1-mediated SPM production and anti-angiogenic activity in GC. In vivo, vessel density and tumor growth were higher in shALOX15 and shGPR32, but not in shALOX5, xenografts than in controls, despite their comparable angiogenic potential in culture. This discrepancy could be explained by the compensatory upregulation of ALOX15 expression in shALOX5 cells that enabled the production of other SPMs, including LXB4. Thus, qualitative or quantitative differences in SPM production between shALOX5 and shALOX15 cells and/or their interaction with tumor stroma might result in a different angiogenic potential in vivo. Be that as it may, in line with our data, an RNAseq analysis showed that higher ALOX15, but not ALOX5, mRNA expression was significantly associated with better overall survival of GC patients (TCGA, http://www.cbioportal.org).[13-15]

Several studies have demonstrated that ω-3 PUFA plays a direct anti-tumor or an adjuvant role in various cancers[41,42] by modulating cell proliferation or apoptosis[43-45] in the high micromolar range.[13,19,20] We believe that, in our system, the effects of an ω-3 or ω-6 diet on xenograft growth were predominantly due to differences in vessel density rather than to a direct effect on cell proliferation. In fact, in the low micromolar range, which corresponds to the concentrations we found in vivo upon dietary modification,[7,22] we found that PUFA suppressed the angiogenic response but did not directly affect GC cell proliferation in culture. In the present study, vessel density in shCTR xenografts did not differ among mice fed different diets, although cell proliferation was slightly, albeit significantly, reduced in mice fed an ω-3 or ω-6 diet. We ascribe this phenomenon to the very low in vivo angiogenic potential and to the massive presence of necrotic areas in these tumors.[7] However, it should be noted that since PUFA administration induces the production of many intermediate metabolites that exert anti-angiogenic and anti-proliferative effects in vivo,[46] we cannot exclude the possibility that in the xenograft microenvironment, these intermediates, together with PUFA, could reach a concentration sufficient to induce xenograft growth inhibition by directly modulating cell proliferation.

Table 1. Lipidomic profile of mice fed ω-6 and/or ω-3 enriched diets.

	DIET		
	CTRL	ω-6	ω-3
AA	6353.6 ± 1444.1	16991 ± 925.8*	5677.2 ± 1295
PGD2	281 ± 43.9	237.3 ± 33.3	327 ± 75.1
PGE2	183.5 ± 45.1	172 ± 66.7	92.7 ± 20*
PGF2α	76 ± 33	57 ± 23	47.5 ± 14.7
TXB2	740.8 ± 64.1	911.4 ± 147.1	466 ± 148.7*
LTB4	740.8 ± 141.2	775 ± 145.9	355.8 ± 109.8*
LXA4	5748.6 ± 258.5	5012.6 ± `185.1	5017 ± 230.5
LXB4	134.3 ± 20.9	318.7 ± 33.4*	183.4 ± 21.5
EPA	1476.2 ± 280.6	2167.8 ± 671.3	5028.1 ± 462.8*
RvE1	108.3 ± 29.9	32.6 ± 17.1	130.8 ± 26.4
RvE2	3251.5 ± 532.2	3797 ± 474.2	8862.7 ± 576.1*
RvE3	213.4 ± 44.6	162.3 ± 42.5	249.7 ± 73.6
DHA	8886.4 ± 1586.2	10220.2 ± 1364.6	17447.5 ± 772*
RvD1	45.8 ± 19.5	52.3 ± 4.7	95.7 ± 12.5*
RvD2	46.6 ± 14.6	47.5 ± 4.9	86 ± 29.4
RvD3	2664.7 ± 365.9	3252 ± 446.5	5101.1 ± 380.3*
RvD5	30.5 ± 5.9	20 ± 8.2	46.7 ± 18.2
RvD6	56.7 ± 15.8	40.7 ± 17.4	98.2 ± 25.6
PD1	56.6 ± 6.3	39.5 ± 6.3	56.5 ± 28

Lipid concentration in serum of mice fed the indicated diets (10 mice/group). Values are reported as nM.
*$p < .05$ vs. the CTRL diet group.

Pattern recognition receptors trigger inflammation,[47] but can also promote its resolution.[12,48,49] This is particularly important for the gastrointestinal tract, which, being continuously exposed to food- and microbioma-derived antigens, is subject to a chronic low-level of mucosal inflammation[50] that must be counteracted to avoid tissue damage. In injured mouse colonic mucosa, FPR1 promotes tissue restitution by recognizing the endogenous N-terminal-derived annexin A1 (Ac2-26) peptides[51,52] or specific commensal microbiota strains.[53] Our data suggest that FPR1 is constitutively active in GC cells, which implies the presence of an endogenous ligand endowed with pro-resolving activity. A possible candidate ligand is peptide Ac2-26 that derives from AnxA1, which, as we previously showed, is constitutively expressed in GC cells.[7]

In conclusion, we describe a novel molecular mechanism in GC, actively controlled by FPR1, that links PUFA metabolism and SPM biosynthesis with angiogenesis. This raises the possibility of new prognostic tools and therapeutic interventions for GC. It is feasible that FPR1 and/or components of pro-resolving pathways may represent novel risk factors or prognostic markers of GC. Administration of SPMs or pharmacologic stimulation of FPR1, ALOX15, or GPR32 could suppress GC angiogenesis and growth. Lastly, as SPM production can be largely determined by diet, increasing ω-3 or ω-6 consumption might be a tool in the management of GC.

Materials and methods

Cell culture

The AGS and MKN45 cell lines derived from poorly differentiated human gastric adenocarcinoma were grown as previously described.[7] To generate AGS cells stably expressing ALOX5, ALOX15, or GPR32 shRNA, we used pools of five constructs (Qiagen, Valencia, CA, USA) containing 21-mer short hairpin RNAs (shRNA) directed to various coding regions of each target gene. Transfectants were selected in medium with 500 ng/mL puromycin.

RNA isolation and real-time PCR

Total RNA was isolated and retrotranscribed according to the manufacturer's instructions (Promega, Madison, WI, USA). Real-time quantitative PCR was performed on the CFX96 system (Bio-Rad, Hercules, CA, USA) using the PE SYBR Green PCR kit (Applied Biosystems, Grand Island, NY, USA). The target-specific primers used for real-time PCR are listed in Table S1. No-reverse transcribed mRNA samples served as a negative control. The expression levels of each target were calculated relative to that of control cells, arbitrarily considered equal to 1. Results were normalized to β-actin mRNA levels.

Flow cytometric analysis

Cells were incubated (30 min at 4°C) with specific or isotype control antibodies (Abs). ALOX5, ALOX15A, and ALOX15B Abs were from Santa Cruz Biotechnology (Dallas, TX, USA), anti-GPR32 was from Acris (Herford, Germany), anti-BLT1

from LSBio (Seattle, WA, USA), and anti-ChemR23 from MyBiosource (San Diego, CA, USA). Cells were analyzed with an FACS Calibur cytofluorimeter using CellQuest software (BD Biosciences, Mississauga, ON, Canada). When necessary, we performed cell membrane permeabilization using the Cytofix/Cytoperm kit (BD Biosciences).

Protein studies

Protein extractions, immunoblotting, and immunoprecipitation were carried out according to standard procedures. Anti-phospho-MAPK, -Akt, -p38, -JNK, -STAT3, -SRC antibodies (Abs) were from Cell Signaling Technology (Danvers, MA, USA). Anti-tubulin was from Sigma-Aldrich (St. Louis, MO, USA), and secondary anti-mouse and anti-rabbit Abs coupled to HRP were from Bio-Rad.

Xenografts in mice

Each group of 10 mice (4-week-old female CD1 nu/nu mice, Charles River, Wilmington, MA) was inoculated subcutaneously with shCTR, shFPR1, shALOX5, shALOX15, shGPR32 AGS cells (1×10^7 cells/mouse). Tumor diameters were measured at regular intervals with a caliper. Tumor volumes (V) were calculated with the formula: $V = A \times B^2/2$ (A = axial diameter; B = rotational diameter). This study was conducted according to Italian regulations for experimentation on animals; the study protocol was approved by the Italian Ministry of Health. Paraffin-embedded tumors were analyzed by immunohistochemistry with anti-Ki-67 antibody from Biocare Medical (Concord, CA), anti-cleaved caspase 3, and anti-CD31 from R&D Systems (Minneapolis, MN).

Fatty composition of diets

The three isocaloric diets containing 5% fat (w/w) were designed as previously described.[9,22] Pellets were prepared by Mucedola (Milan, Italy). We used grape seed and colza oils (50/50%) to prepare the pellets for the control diet (CTRL), grape seed oil (100%) for the ω-6 (ω-6), and colza oil/fish oils (80/20%) for an ω-3 fatty acid-enriched diet (ω-3). The oils were obtained from Sigma-Aldrich (St. Louis, MO). Pellets were stored under vacuum at −20°C. The diets were changed twice a week in each animal cage to avoid oxidative degradation of lipids.

ELISA and EIA assays

VEGF-A contents in culture supernatants were measured in duplicate determinations with a commercially available ELISA (R&D Systems). RvD1, LTB4, PGE2, and LXB4 contents in culture supernatants were measured in triplicate determinations with a commercially available EIA (Cayman Chemical, Ann Arbor, MI).

LC-MS/MS instrumentation and conditions

To extract the lipids, 120 μL samples (mice serum or cell culture supernatants), together with deuterated internal standards,

were loaded onto solid phase extraction Waters Oasis HLB cartridges (60 mg sorbent, 30 μm particle size), and eluted with 2 mL methanol and 2 mL ethyl acetate into polypropylene tubes containing 6 μL of a glycerol solution (30% in methanol). The SPE eluates were dried and residues were then reconstituted in 100 μL methanol. 5 μL of methanol extract were analyzed by using a 4000QTrap mass spectrometer (Applied Biosystems) coupled to a 1,100 nanoHPLC system (Agilent Technologies, Waldbronn, Germany). The lipids were separated by using a micro-C18 column (10 cm x 1.0 mm, 5u). The mobile phase was generated by mixing eluent A (0.1% acetic acid) and eluent B (acetonitrile/isopropanol 50/50) and the flow rate was 30 nL/min. Starting condition was 20% to 95% B in 5 min. Tandem mass spectrometry was performed using a turbo ion spray source operated in a negative mode (curtain gas 20psi, GS 1/2 50/50psi, ion spray voltage $-5,500$ V, DP, -60 V; Dwell 25 ms and temperature of 550°C), and the multiple reaction monitoring mode was used to detect a unique product ion arising from collision-induced fragmentation of the protonated parent compound as reported in Table S2.

Statistical analysis

Values from groups were compared by using the paired Student t test or the Duncan test. A p value < 0.05 was considered statistically significant.

Disclosure of potential conflicts of interest

The authors whose names are listed above inform the Editors-in-Chief, Deputy Editors, Senior Editors, or Scientific Editors that they have NO relationships that they believe could be construed as resulting in an actual, potential, or perceived conflict of interest with regard to this manuscript.

ORCID

Nella Prevete http://orcid.org/0000-0002-0186-5431
Rosa Marina Melillo http://orcid.org/0000-0002-9233-5275

References

1. Balkwill FR, Capasso M, Hagemann T. The tumor microenvironment at a glance. J Cell Sci 2012; 125(Pt 23):5591-6; PMID:23420197; http://dx.doi.org/10.1242/jcs.116392
2. Serhan CN. Pro-resolving lipid mediators are leads for resolution physiology. Nature 2014; 510(7503):92-101; PMID:24899309; http://dx.doi.org/10.1038/nature13479
3. Serhan CN, Chiang N, Van Dyke TE. Resolving inflammation: dual anti-inflammatory and pro-resolution lipid mediators. Nat Rev Immunol 2008; 8(5):349-61; PMID:18437155; http://dx.doi.org/10.1038/nri2294
4. Serhan CN, Chiang N, Dalli J. The resolution code of acute inflammation: novel pro-resolving lipid mediators in resolution. Semin Immunol 2015; 27(3):200-15; PMID:25857211; http://dx.doi.org/10.1016/j.smim.2015.03.004
5. Ye RD, Boulay F, Wang JM, Dahlgren C, Gerard C, Parmentier M, Serhan CN, Murphy PM. International Union of Basic and Clinical Pharmacology. LXXIII. Nomenclature for the formyl peptide receptor (FPR) family. Pharmacol Rev 2009; 61(2):119-61; PMID:19498085; http://dx.doi.org/10.1124/pr.109.001578
6. Dufton N, Perretti M. Therapeutic anti-inflammatory potential of formyl-peptide receptor agonists. Pharmacol Ther 2010; 127(2):175-88; PMID:20546777; http://dx.doi.org/10.1016/j.pharmthera.2010.04.010
7. Prevete N, Liotti F, Visciano C, Marone G, Melillo RM, de Paulis A. The formyl peptide receptor 1 exerts a tumor suppressor function in human gastric cancer by inhibiting angiogenesis. Oncogene 2015; 34(29):3826-38; PMID:25263443; http://dx.doi.org/10.1038/onc.2014.309
8. Cash JL, Norling LV, Perretti M. Resolution of inflammation: targeting GPCRs that interact with lipids and peptides. Drug Discov Today 2014; 19(8):1186-92; PMID:24993159; http://dx.doi.org/10.1016/j.drudis.2014.06.023
9. Gobbetti T, Ducheix S, le Faouder P, Perez T, Riols F, Boue J, Bertrand-Michel J, Dubourdeau M, Guillou H, Perretti M et al. Protective effects of n-6 fatty acids-enriched diet on intestinal ischaemia/reperfusion injury involve lipoxin A4 and its receptor. Br J Pharmacol 2015; 172(3):910-23; PMID:25296998; http://dx.doi.org/10.1111/bph.12957
10. Ivanov I, Kuhn H, Heydeck D. Structural and functional biology of arachidonic acid 15-lipoxygenase-1 (ALOX15). Gene 2015; 573 (1):1-32; PMID:26216303; http://dx.doi.org/10.1016/j.gene.2015.07.073
11. Brash AR, Boeglin WE, Chang MS. Discovery of a second 15S-lipoxygenase in humans. Proc Natl Acad Sci U S A 1997; 94(12):6148-52; PMID:9177185; http://dx.doi.org/10.1073/pnas.94.12.6148
12. Prevete N, Liotti F, Marone G, Melillo RM, de Paulis A. Formyl peptide receptors at the interface of inflammation, angiogenesis and tumor growth. Pharmacol Res 2015; 102:184-91; PMID:26466865; http://dx.doi.org/10.1016/j.phrs.2015.09.017
13. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013; 6(269):pl1; PMID:23550210; http://dx.doi.org/10.1126/scisignal.2004088
14. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012; 2(5):401-4; PMID:22588877; http://dx.doi.org/10.1158/2159-8290.CD-12-0095
15. Cancer Genome Atlas Research N. Comprehensive molecular characterization of gastric adenocarcinoma. Nature 2014; 513(7517):202-9; PMID:25079317; http://dx.doi.org/10.1038/nature13480
16. Lu JM, Nurko J, Weakley SM, Jiang J, Kougias P, Lin PH, Yao Q, Chen C. Molecular mechanisms and clinical applications of nordihydroguaiaretic acid (NDGA) and its derivatives: an update. Med Sci Monit 2010; 16(5):RA93-100; PMID:20424564
17. Lin L, Hutzen B, Zuo M, Ball S, Deangelis S, Foust E, Pandit B, Ihnat MA, Shenoy SS, Kulp S et al. Novel STAT3 phosphorylation inhibitors exhibit potent growth-suppressive activity in pancreatic and breast cancer cells. Cancer Res 2010; 70(6):2445-54; PMID:20215512; http://dx.doi.org/10.1158/0008-5472.CAN-09-2468
18. Uehara Y, Mochizuki M, Matsuno K, Haino T, Asai A. Novel high-throughput screening system for identifying STAT3-SH2 antagonists. Biochem Biophys Res Commun 2009; 380(3):627-31; PMID:19285012; http://dx.doi.org/10.1016/j.bbrc.2009.01.137
19. Finstad HS, Drevon CA, Kulseth MA, Synstad AV, Knudsen E, Kolset SO. Cell proliferation, apoptosis and accumulation of lipid

droplets in U937-1 cells incubated with eicosapentaenoic acid. Biochem J 1998; 336 (Pt 2):451-9; PMID:9820824; http://dx.doi. org/10.1042/bj3360451

20. Sun SN, Jia WD, Chen H, Ma JL, Ge YS, Yu JH, Li JS. Docosa-hexaenoic acid (DHA) induces apoptosis in human hepatocellular carcinoma cells. Int J Clin Exp Pathol 2013; 6(2):281-9; PMID:23330014

21. Peng Y, Zheng Y, Zhang Y, Zhao J, Chang F, Lu T, Zhang R, Li Q, Hu X, Li N. Different effects of omega-3 fatty acids on the cell cycle in C2C12 myoblast proliferation. Mol Cell Biochem 2012; 367(1-2):165-73; PMID:22610825; http://dx.doi.org/10.1007/s11010-012-1329-4

22. Ducheix S, Montagner A, Polizzi A, Lasserre F, Marmugi A, Bertrand-Michel J, Podechard N, Al Saati T, Chetiveaux M, Baron S et al. Essential fatty acids deficiency promotes lipogenic gene expression and hepatic steatosis through the liver X receptor. J Hepatol 2013; 58(5):984-92; PMID:23333450; http://dx.doi.org/10.1016/j.jhep.2013.01.006

23. Khau T, Langenbach SY, Schuliga M, Harris T, Johnstone CN, Anderson RL, Stewart AG. Annexin-1 signals mitogen-stimulated breast tumor cell proliferation by activation of the formyl peptide receptors (FPRs) 1 and 2. FASEB J 2011; 25(2):483-96; PMID:20930115; http://dx.doi.org/10.1096/fj.09-154096

24. Yang Y, Liu Y, Yao X, Ping Y, Jiang T, Liu Q, Xu S, Huang J, Mou H, Gong W et al. Annexin 1 released by necrotic human glioblastoma cells stimulates tumor cell growth through the formyl peptide receptor 1. Am J Pathol 2011; 179(3):1504-12; PMID:21782780; http://dx.doi.org/10.1016/j.ajpath.2011.05.059

25. Belvedere R, Bizzarro V, Popolo A, Dal Piaz F, Vasaturo M, Picardi P, Parente L, Petrella A. Role of intracellular and extracellular annexin A1 in migration and invasion of human pancreatic carcinoma cells. BMC Cancer 2014; 14:961; PMID:25510623; http://dx.doi.org/10.1186/1471-2407-14-961

26. Snapkov I, Oqvist CO, Figenschau Y, Kogner P, Johnsen JI, Sveinbjornsson B. The role of formyl peptide receptor 1 (FPR1) in neuroblastoma tumorigenesis. BMC Cancer 2016; 16:490; PMID:27432059; http://dx.doi.org/10.1186/s12885-016-2545-1

27. Chen K, Liu M, Liu Y, Yoshimura T, Shen W, Le Y, Durum S, Gong W, Wang C, Gao JL et al. Formylpeptide receptor-2 contributes to colonic epithelial homeostasis, inflammation, and tumorigenesis. J Clin Invest 2013; 123(4):1694-704; PMID:23454745; http://dx.doi.org/10.1172/JCI65569

28. Giebeler A, Streetz KL, Soehnlein O, Neumann U, Wang JM, Brandenburg LO. Deficiency of formyl peptide receptor 1 and 2 is associated with increased inflammation and enhanced liver injury after LPS-stimulation. PloS One 2014; 9(6):e100522; PMID:24956481; http://dx.doi.org/10.1371/journal.pone.0100522

29. Baracco EE, Pietrocola F, Buque A, Bloy N, Senovilla L, Zitvogel L, Vacchelli E, Kroemer G. Inhibition of formyl peptide receptor 1 reduces the efficacy of anticancer chemotherapy against carcinogen-induced breast cancer. Oncoimmunology 2016; 5(6):e1139275; PMID:27471610; http://dx.doi.org/10.1080/2162402X.2016.1139275

30. Vacchelli E, Ma Y, Baracco EE, Zitvogel L, Kroemer G. Yet another pattern recognition receptor involved in the chemotherapy-induced anticancer immune response: formyl peptide receptor-1. Oncoimmunology 2016; 5(5):e1118600; PMID:27467929; http://dx.doi.org/10.1080/2162402X.2015.1118600

31. Cheng TY, Wu MS, Lin JT, Lin MT, Shun CT, Hua KT, Kuo ML. Formyl Peptide receptor 1 expression is associated with tumor progression and survival in gastric cancer. Anticancer Res 2014; 34(5):2223-9; PMID:24778024

32. Otani T, Ikeda S, Lwin H, Arai T, Muramatsu M, Sawabe M. Polymorphisms of the formylpeptide receptor gene (FPR1) and susceptibility to stomach cancer in 1531 consecutive autopsy cases. Biochem Biophys Res Commun 2011; 405(2):356-61; PMID:21216225; http://dx.doi.org/10.1016/j.bbrc.2010.12.136

33. Seifert R, Wenzel-Seifert K. Defective Gi protein coupling in two formyl peptide receptor mutants associated with localized juvenile periodontitis. J Biol Chem 2001; 276(45):42043-9; PMID:11559706; http://dx.doi.org/10.1074/jbc.M106621200

34. Maney P, Emecen P, Mills JS, Walters JD. Neutrophil formylpeptide receptor single nucleotide polymorphism 348T>C in aggressive periodontitis. J Periodontol 2009; 80(3):492-8; PMID:19254133; http://dx.doi.org/10.1902/jop.2009.080225

35. Pouliot M, Clish CB, Petasis NA, Van Dyke TE, Serhan CN. Lipoxin A (4) analogues inhibit leukocyte recruitment to Porphyromonas gingivalis: a role for cyclooxygenase-2 and lipoxins in periodontal disease. Biochemistry 2000; 39(16):4761-8; PMID:10769133; http://dx.doi.org/10.1021/bi992551b

36. Van Dyke TE, Hasturk H, Kantarci A, Freire MO, Nguyen D, Dalli J, Serhan CN. Proresolving nanomedicines activate bone regeneration in periodontitis. J Dent Res 2015; 94(1):148-56; PMID:25389003; http://dx.doi.org/10.1177/0022034514557331

37. Hasturk H, Kantarci A, Goguet-Surmenian E, Blackwood A, Andry C, Serhan CN, Van Dyke TE. Resolvin E1 regulates inflammation at the cellular and tissue level and restores tissue homeostasis in vivo. J Immunol 2007; 179(10):7021-9; PMID:17982093; http://dx.doi.org/10.4049/jimmunol.179.10.7021

38. Hasturk H, Kantarci A, Ohira T, Arita M, Ebrahimi N, Chiang N, Petasis NA, Levy BD, Serhan CN, Van Dyke TE. RvE1 protects from local inflammation and osteoclast- mediated bone destruction in periodontitis. FASEB J 2006; 20(2):401-3; PMID:16373400; http://dx.doi.org/10.1096/fj.05-4724fje

39. Jin Y, Arita M, Zhang Q, Saban DR, Chauhan SK, Chiang N, Serhan CN, Dana R. Anti-angiogenesis effect of the novel anti-inflammatory and pro-resolving lipid mediators. Invest Ophthalmol Vis Sci 2009; 50(10):4743-52; PMID:19407006; http://dx.doi.org/10.1167/iovs.08-2462

40. Klil-Drori AJ, Ariel A. 15-Lipoxygenases in cancer: a double-edged sword? Prostaglandins Other Lipid Mediat 2013; 106:16-22; PMID:23933488; http://dx.doi.org/10.1016/j.prostaglandins.2013.07.006

41. Dunbar BS, Bosire RV, Deckelbaum RJ. Omega 3 and omega 6 fatty acids in human and animal health: an African perspective. Mol Cell Endocrinol 2014; 398(1-2):69-77; PMID:25458696; http://dx.doi.org/10.1016/j.mce.2014.10.009

42. Gerber M. Omega-3 fatty acids and cancers: a systematic update review of epidemiological studies. Br J Nutr 2012; 107(Suppl 2):S228-39; PMID:22591896; http://dx.doi.org/10.1017/S0007114512001614

43. Song M, Nishihara R, Cao Y, Chun E, Qian ZR, Mima K, Inamura K, Masugi Y, Nowak JA, Nosho K et al. Marine omega-3 polyunsaturated fatty acid intake and risk of colorectal cancer characterized by tumor-infiltrating T Cells. JAMA Oncol 2016; 2(9):1197-206; PMID:27148825; http://dx.doi.org/10.1001/jamaoncol.2016.0605

44. Black HS, Rhodes LE. Potential benefits of omega-3 fatty acids in non-melanoma skin cancer. J Clin Med 2016; 5(2):pii: E23; PMID:26861407; http://dx.doi.org/10.3390/jcm5020023

45. Eltweri AM, Thomas AL, Metcalfe M, Calder PC, Dennison AR, Bowrey DJ. Potential applications of fish oils rich in omega-3 polyunsaturated fatty acids in the management of gastrointestinal cancer. Clin Nutr 2017; 36(1):65-78; PMID:26833289; http://dx.doi.org/10.1016/j.clnu.2016.01.007

46. Sapieha P, Stahl A, Chen J, Seaward MR, Willett KL, Krah NM, Dennison RJ, Connor KM, Aderman CM, Liclican E et al. 5-Lipoxygenase metabolite 4-HDHA is a mediator of the antiangiogenic effect of omega-3 polyunsaturated fatty acids. Sci Transl Med 2011; 3(69):69ra12; PMID:21307302; http://dx.doi.org/10.1126/scitranslmed.3001571

47. Takeuchi O, Akira S. Pattern recognition receptors and inflammation. Cell 2010; 140(6):805-20; PMID:20303872; http://dx.doi.org/10.1016/j.cell.2010.01.022

48. Parlato M, Yeretssian G. NOD-like receptors in intestinal homeostasis and epithelial tissue repair. Int J Mol Sci 2014; 15(6):9594-627; PMID:24886810; http://dx.doi.org/10.3390/ijms15069594

49. Rakoff-Nahoum S, Medzhitov R. Toll-like receptors and cancer. Nat Rev Cancer 2009; 9(1):57-63; PMID:19052556; http://dx.doi.org/10.1038/nrc2541

50. Fichtner-Feigl S, Kesselring R, Strober W. Chronic inflammation and the development of malignancy in the GI tract. Trends Immunol 2015; 36(8):451-9; PMID:26194796; http://dx.doi.org/10.1016/j.it.2015.06.007

51. Babbin BA, Laukoetter MG, Nava P, Koch S, Lee WY, Capaldo CT, Peatman E, Severson EA, Flower RJ, Perretti M et al. Annexin A1 regulates intestinal mucosal injury, inflammation, and repair. J Immunol 2008; 181(7):5035-44; PMID:18802107; http://dx.doi.org/10.4049/jimmunol.181.7.5035

52. Leoni G, Alam A, Neumann PA, Lambeth JD, Cheng G, McCoy J, Hilgarth RS, Kundu K, Murthy N, Kusters D et al. Annexin A1, formyl peptide receptor, and NOX1 orchestrate epithelial repair. J Clin Invest 2013; 123(1):443-54; PMID:23241962; http://dx.doi.org/10.1172/JCI65831

53. Alam A, Leoni G, Quiros M, Wu H, Desai C, Nishio H, Jones RM, Nusrat A, Neish AS. The microenvironment of injured murine gut elicits a local pro-restitutive microbiota. Nat Microbiol 2016; 1:15021; PMID:27571978; http://dx.doi.org/10.1038/nmicrobiol.2015.21

CrossMark
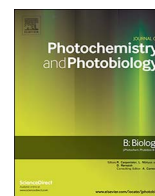
Malvidin and cyanidin derivatives from açai fruit (*Euterpe oleracea Mart.*) counteract UV-A-induced oxidative stress in immortalized fibroblasts

Ganna Petruk[a], Anna Illiano[a], Rita Del Giudice[a], Assunta Raiola[b], Angela Amoresano[a], Maria Manuela Rigano[b], Renata Piccoli[a,c], Daria Maria Monti[a,c,*]

[a] *Department of Chemical Sciences, University of Naples Federico II, Complesso Universitario Monte Sant'Angelo, via Cinthia 4, 80126 Naples, Italy*
[b] *Department of Agricultural Sciences, University of Naples Federico II, Via Università 100, 80055 Portici, Naples, Italy*
[c] *Istituto Nazionale di Biostrutture e Biosistemi (INBB), Rome, Italy*

A B S T R A C T

UV-A radiations are known to induce cellular oxidative stress, leading to premature skin aging. Consumption of açai fruit (*Euterpe oleracea Martius*) is known to have many health benefits due to its high level of antioxidants. Herein, we analyzed the ability of phenolic compounds extracted from this fruit to attenuate UV-A-induced oxidative stress in immortalized fibroblast.

A methanol/water açai extract was fractionated by HPLC and each fraction tested for anti-oxidant stress activity. Immortalized fibroblasts were pre-incubated with açai fractions and then exposed to UV-A radiations. Açai extract was found to be able to strongly protect cells from oxidative stress. In particular, reactive oxygen species (ROS) production, GSH depletion, lipid peroxidation and no increase in the phosphorylation levels of proteins involved in the oxidative stress pathway was observed in cells pre-incubated with the extract and then irradiated by UV-A. Mass spectrometry analyses of HPLC fractionated extract led us to the identification of malvidin and cyanidin derivatives as the most active molecules able to counteract the negative effects induced by UV-A irradiation.

Our results indicate, for the first time, that açai fruit is a valuable natural source for malvidin and cyanidin to be used as anti-stress molecules and represent good candidates for dietary intervention in the prevention of age related skin damage.

1. Background

In the last few years, an increasing attention has focused on age-related diseases, including skin aging. Oxidative stress caused by aging is considered a general initiating factor of neurodegeneration and carcinogenesis [1]. Indeed, it is known that the skin is constantly exposed to oxidative stress induced by reactive oxygen species (ROS), generated by endogenous (i.e. enzyme activities) or exogenous sources [2,3].

UV radiation from sunlight is one of the most important health-related environmental factors because of its hazardous effects, which include generation of skin cancer, suppression of the immune system, and premature skin aging [2]. In particular, UV-A (400–315 nm)

radiations are weakly absorbed by DNA, but rather excite endogenous chromophores, leading to DNA damage. This occurs through the production of reactive singlet oxygen that specifically reacts with guanine within the DNA molecule [4,5]. UV-A radiations may also promote the formation of hydroxyl radicals via the photosensitized production of superoxide anions. Because of their high reactivity and low specificity, hydroxyl radicals likely induce a wide range of DNA damage [6]. Fibroblasts, cells of the dermis, are continuously exposed to UV-A radiations, which are able to penetrate deeply in the skin. The cellular antioxidant defense system is composed of endogenously produced antioxidant molecules, but the intrinsic mechanism for antioxidant defense is gradually impaired with aging, resulting in an inability to deal with ROS generation [7]. A continuous and regular

Table 1

Total phenolic acids, total flavonoids and hydrophilic antioxidant activity in açai extract.

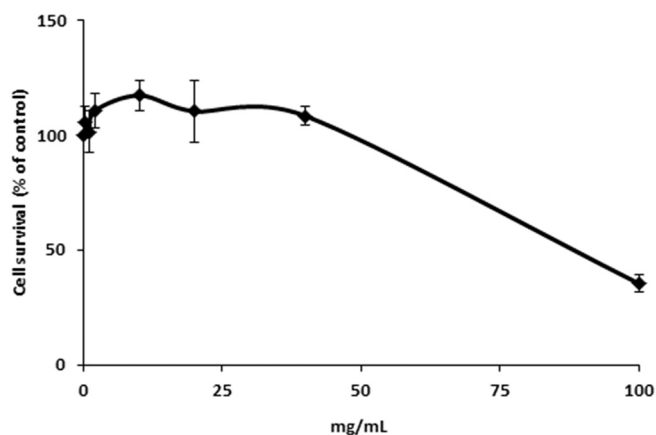| Açai antioxidant determination	Value
Total phenols (mg/100 g DW)	192.41 ± 10.78
Total flavonoid (mg/100 g DW)	159.67 ± 7.65
HAA (mmol TE/100 g DW)	2.1 ± 0.17

Fig. 1. Effect of açai extracts on the viability of BALB/3T3 fibroblasts. Dose-response curve of BALB/3T3 cells after 48 h incubation with increasing concentrations of açai extracts. Cell viability was assessed by the MTT assay and expressed as described in Materials and Methods section. Values are given as means ± S.D. (n ≥ 3).

intake of vitamins, trace metals, polyunsaturated fatty acids, and polyphenols from food sources contributes to counteract oxidative stress [8,9] and in preventing or retarding age-related diseases [10–12]. As an example, we recently report the beneficial effects of tomato extracts in counteracting oxidative stress in different cell lines [13,14]. Therefore, a continuous search for natural extracts highly rich in antioxidant molecules is needed to identify and provide novel natural drugs against aging-associated diseases.

Açai (*Euterpe oleracea Martius*) tree is a large palm found in the Amazon flood plain. The fruit of this palm is a small purple-black berry which reaches about 10 mm in diameter and is usually consumed in all states of Brazil since it is rich in α-tocopherol, fibers, lipids, polyphenols (including anthocyanins), and mineral ions [15,16]. The high polyphenol content, mostly composed of anthocyanins and flavones, is thought to confer to açai fruit several health-promoting effects, including anti-inflammatory, immunomodulatory, antinociceptive, and antioxidant properties [17–25]. In particular, açai extracts were shown to increase plasma antioxidant capacity [21], to decrease oxidative stress in endothelial cells [26], to attenuate tumor growth in mice affected by esophageal cancer [27] and to lower the level of blood cholesterol in animal models for hypercholesterolemia [28].

A positive role of açai in modulating ROS production and activating antioxidant genes expression in rat liver was also reported [29] and, more recently, Peixoto and colleagues demonstrated the beneficial effect of açai extract in counteracting oxidative stress and aging in *C. elegans* [18].

Although several reports on the spectrum of health benefits of açai have been reported, only few studies are available so far on the identification of the antioxidant molecules responsible for the reported beneficial effects on human health.

Here, a methanol/water extract from *Euterpe oleracea* fruits was analyzed for its antioxidant activity on fibroblasts exposed to UV-A-induced insults. A combined approach of bioassays and mass spectrometry analyses led to the identification of açai bioactive compounds.

2. Methods

2.1. Açai Extracts

Methanolic extracts from açai fruit were obtained as reported by Rigano et al. [30], starting from commercially available dried powder (2 g, Tuialimentos, Brasil). The mixture was dried in a rotovapor (R-210, Buchi), and dissolved in 5% dimethyl sulfoxide (DMSO) in PBS (1 mL).

2.2. Antioxidant Compounds Determination and Antioxidant Activity Analysis

Total phenols content was determined in the whole açai extract according to the method of Singleton et al. [31] modified as reported by Rigano et al. [30]. Briefly, an equal volume of Folin-Ciocalteu's phenol reagent and two volumes of ddH$_2$O were added to the hydrophilic extract. After 6 min, Na$_2$CO$_3$ was added (7% final concentration). After 90 min the absorbance was read at 760 nm. A standard curve was obtained by using gallic acid in the range 0–70 μg/mL. The total phenolic content was expressed as mg of gallic acid equivalents (GAE)/100 g dry weight (DW) of açai. Three independent analyses
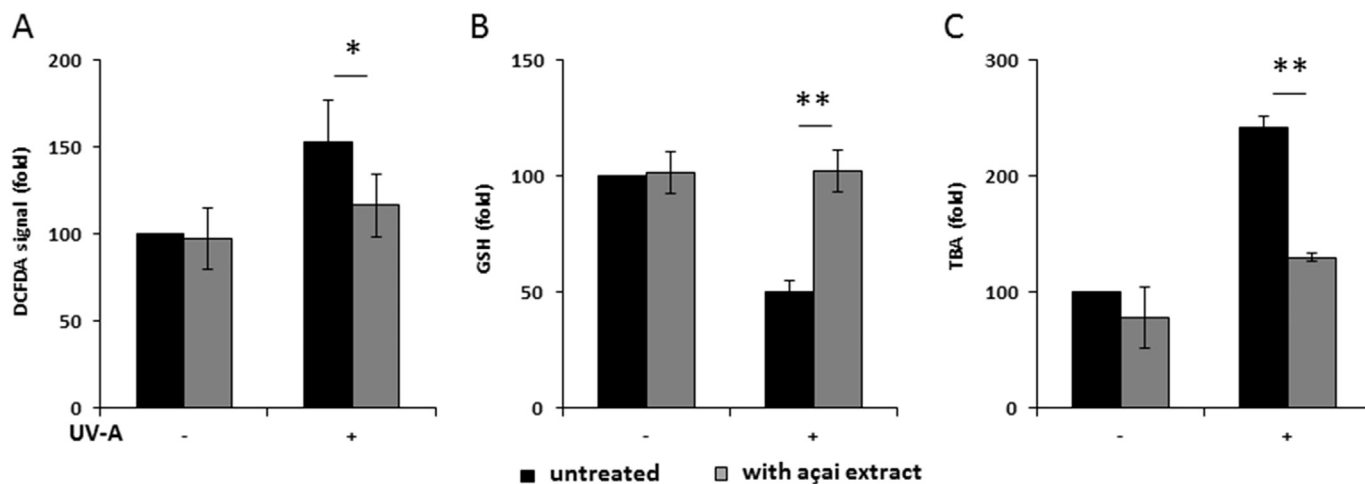

Fig. 2. ROS production, GSH oxidation and lipid peroxidation in BALB/3T3 cells irradiated by UV-A in the presence of açai extracts. Cells were pre-incubated in the presence of 10 mg/mL açai extract (grey bars) for 2 h and then irradiated by UV-A (100 J/cm^2). A, intracellular ROS levels were determined by DCFDA assay; B, intracellular GSH levels determined by DTNB assay; C, lipid peroxidation levels determined by TBARS assay. Values are expressed as fold increase with respect to control (i.e. untreated) cells. Data shown are the means ± S.D. of three independent experiments. * indicates p < 0.01; ** indicates p < 0.001.
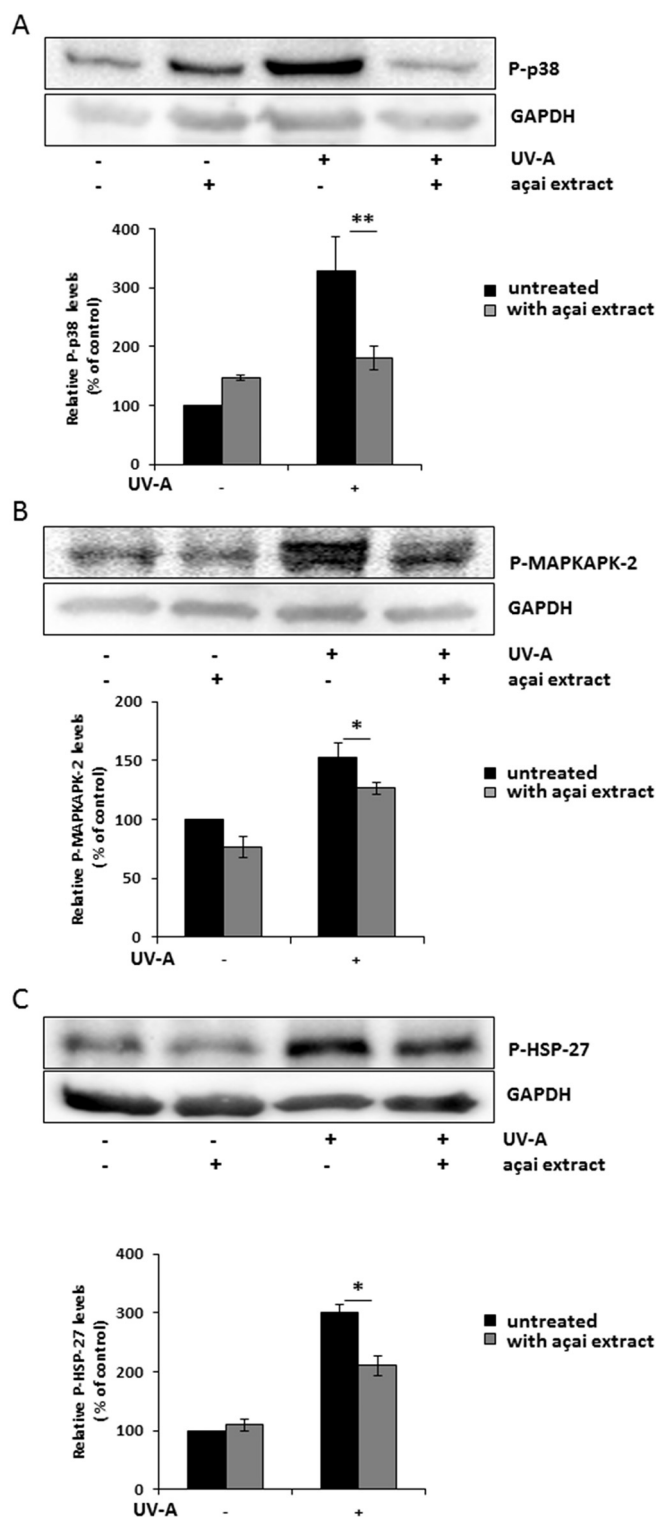
A

B

C

Fig. 3. Effect of açai extracts on UV-A-induced oxidative stress markers in BALB/3T3 fibroblasts. Cells were incubated with 10 mg/mL açaiextract2h prior to UV-A irradiation (100 J/cm^2) and then cells were incubated for 90 min. Western blots show the phosphorylation levels of p38 (A),MAPKAPK-2 (B) and HSP-27 (C), with the relative densitometric analysis in the absence (black bars) or in the presence (grey bars) of açai extract. GAPDH was used as internal standard. Data shown are the means ± S.D. of three independent experiments. * indicates p < 0.01; ** indicates p < 0.001.

were carried out.

Total flavonoids were estimated by the aluminum chloride colorimetric assay reported by Marinova et al. [32] with modifications reported by Raiola et al. [33]. Briefly, NaNO$_2$ (5%) was added to

methanolic extract and, after 5 min incubation, AlCl$_3$ (10% w/v) was added. After 6 min, NaOH (1 M) was added and the absorbance measured at 510 nm. A standard curve was obtained by using quercetin in the range 0–100 μg/mL. Total flavonoids content was expressed as mg quercetin equivalents (QE)/100 g DW. Three independent analyses were carried out.

HAA was evaluated in the water-soluble fraction using the protocol described by Miller et al. [34] with modifications reported by Del Giudice et al. [35]. Briefly, methanolic extracts were allowed to react with an ABTS$^{\bullet+}$ solution for 2.5 min, and then the absorbance was measured at 734 nm using a spectrophotometer. A standard curve, obtained by using Trolox, was linear between 0 and 20 μM Trolox. Values were expressed as mmol TE/kg DW of açai. Three separate analyses were carried out with each sample.

2.3. Cell Culture and MTT Assay

BALB/3T3 fibroblasts (clone A31, from ATCC) were cultured in Dulbecco's Modified Eagle's Medium (Sigma-Aldrich, St Louis, Mo, USA), supplemented with 10% foetal bovine serum (HyClone), 2 mM ʟ-glutamine and antibiotics, all from Sigma-Aldrich, in a 5% CO$_2$ humidified atmosphere at 37 °C.

Cells were seeded in 96-well plates (100 μL/well) at a density of 5×10^3/well. For dose-dependent cytotoxicity assays, 24 h after seeding, increasing volumes of açai methanolic extracts were added to the cells, to reach a final concentration ranging from 0.2 to 100 mg/mL. After 48 h incubation, cell viability was assessed by the MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay, as described by Monti et al. [36]. Cell survival was expressed as the percentage of viable cells in the presence of extract compared to that of control samples. Two groups of cells were used as control, i.e. cells untreated with the extract and cells supplemented with identical volumes of DMSO. Each sample was tested in three independent analyses, each carried out in triplicates.

2.4. Oxidative Stress

To analyze oxidative stress, cells were plated at a density of 4×10^4 cells/cm^2. 24 h after seeding, cells were incubated for 2 h in the presence or absence of 10 mg/mL of açai extract, or equivalent amount of each isolated fraction, and then irradiated for 10 min with UV-A light (100 J/cm^2) (treatment before injury). In a second group of experiments, cells were irradiated for 10 min with UV-A (100 J/cm^2) and then incubated for 2 h in the presence or absence of açai extract or equivalent amount of each isolated fraction (treatment after injury).

2.5. DCFDA Assay

To estimate ROS production, the protocol described in [13] was followed. Briefly, at the end of incubation, cells were incubated with 2′,7′-dichlorodihydrofluorescein diacetate (H$_2$-DCFDA, Sigma-Aldrich). Fluorescence intensity was measured by a Perkin-Elmer LS50 spectrofluorimeter (525 nm emission wavelength, 488 nm excitation wavelength, 300 nm/min scanning speed, 5 slit width for both excitation and emission). ROS production was expressed as percentage of DCF fluorescence intensity of the sample under test, with respect to the untreated sample. Each value was assessed by three independent experiments, each with three determinations.

2.6. DTNB Assay

To estimate intracellular glutathione levels, a procedure previously described was followed [14]. Briefly, at the end of incubation, cells were detached by trypsin, lysed and protein concentration was determined by the Bradford assay. Then, 50 μg of proteins were incubated with 3 mM EDTA, 144 μM 5,5′-dithiobis-2-nitrobenzoic acid (DTNB) in

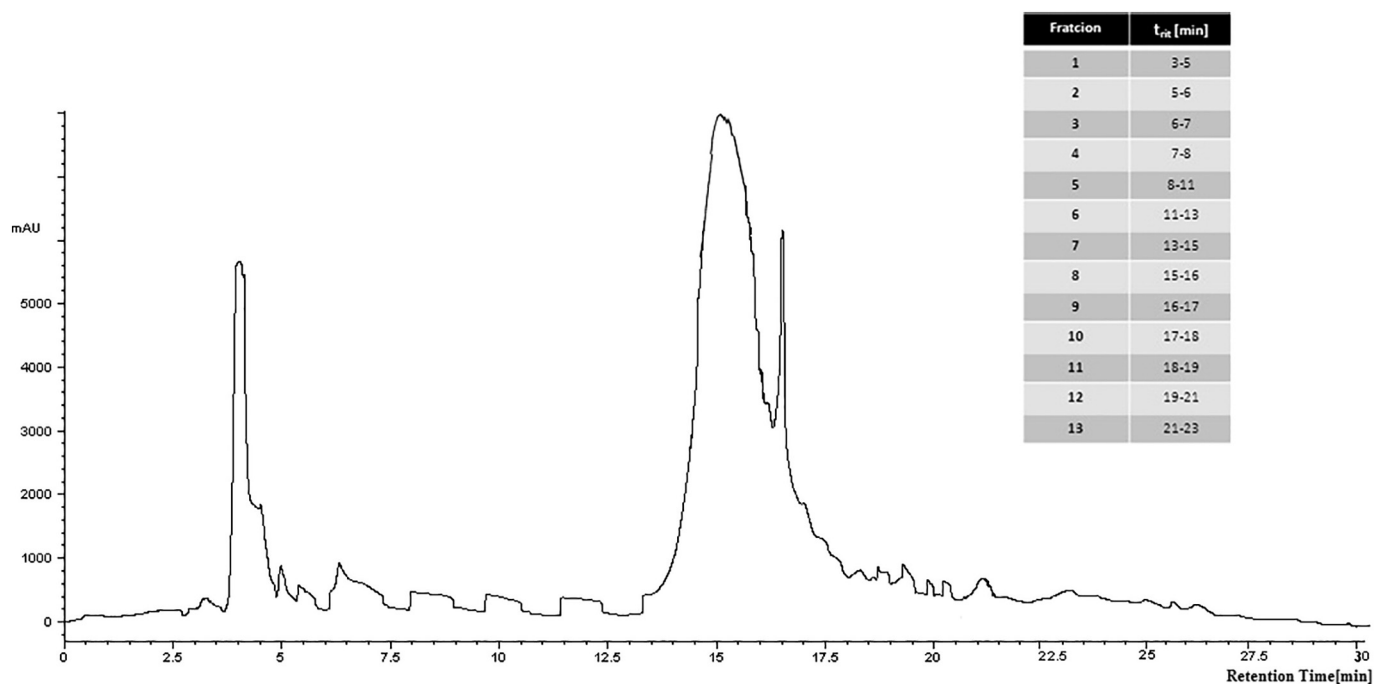| Fratcion	t_{rit} [min]
1	3-5
2	5-6
3	6-7
4	7-8
5	8-11
6	11-13
7	13-15
8	15-16
9	16-17
10	17-18
11	18-19
12	19-21
13	21-23

Fig. 4. Fractional analysis of açai extract by reverse-phase HPLC. Reverse phase C18 column was used with a linear gradient from 5% to 95% in 0.1% trifluoroacetic acid; elution was monitored at 278 nm. Thirteen fractions were manually collected.

30 mM TrisHCl pH 8.2, centrifuged at 14,000g for 5 min at 4 °C and the absorbance of the supernatant was measured at 412 nm by using a multiplate reader (Biorad). GSH levels were expressed as the percentage of TNB absorbance in the sample under test with respect to the untreated sample. Values are the mean of three independent experiments, each with triplicate determinations.

2.7. Measurement of Lipid Peroxidation

The thiobarbituric acid reactive substances (TBARS) assay was performed as described by Del Giudice et al. [13]. Briefly, after irradiation with UV-A, cells were kept for 90 min at 37 °C, then detached and suspended (5×10^4 cells) in 0.67% thiobarbituric acid (TBA) and 20% trichloroacetic acid (1:1 v/v). After heating and centrifugation at 3000g for 5 min at 4 °C, samples were read at 532 nm. Lipid peroxidation levels were expressed as the percentage of the absorbance at 532 nm of the sample under test, with respect to untreated cells (100%). Three independent experiments were carried out, each one with three determinations.

2.8. Western Blot Analyses

Cells were plated at a density of 2×10^4 cells/cm^2 in complete medium for 24 h and then treated as described above (paragraph 2.4). When açai extract was used before UV-A exposure, cells were incubated for further 90 min at 37 °C. Alternatively, following the induction of UV-A stress, cells were treated with the extract for 120 min at 37 °C. Cell lysates were then analyzed by Western blotting performed as reported by Galano et al. [37]. Phosphorylation levels of p38, MAP-KAPK-2 or HSP-27 were detected by using specific antibodies purchased from Cell Signal Technology (Danvers, MA, USA). To normalize protein intensity levels, specific antibodies against internal standards were used, i.e. anti-β-actin (Proteintech, Manchester, UK) or anti-GAPDH (ThermoFisher, Rockford, IL, USA). The chemiluminescence detection system (SuperSignal® West Pico) was from Thermo Fisher.

2.9. Reverse-phase HPLC Analysis

Methanol/water açai extract (2 g/mL, 1 mL) was fractionated by reverse-phase HPLC on a Phenomenex Jupiter C18 column (250 × 2.00 mm 5 μm, 300 Å pore size) (Phenomenex, Torrance, California, USA) with a linear gradient from 5% to 95% acetonitrile (Sigma Aldrich,Saint Louis, USA) in 0.1% trifluoroacetic acid (TFA) (Carlo Erba Reagents S.r.l, Milan, Italy) over 40 min, at a flow rate of 200 μL/min; elution was monitored at 278 nm. The eluate was collected manually in thirteen fractions (1 mL). Fractions were lyophilized and dissolved in 5% DMSO in PBS, and their content identified by MALDI-TOF (AB SCIEX, Milan, Italy) analysis.

2.10. Mass Spectrometry Analyses

Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) experiments were performed on a 5800 MALDI-TOF-TOF ABSciex equipped with a nitrogen laser (337 nm) (AB SCIEX, Milan, Italy). Aliquots of açai HPLC fractions (0.5 μL) were mixed (1:1, v/v) with 2.5 dihydroxybenzoic acid (10 mg/mL) (Sigma-Aldrich, Saint Louis, USA) in acetonitrile:water (90:10) solution. Calibration was performed by using AB SCIEX calibration mixture (Monoisotopic (M + nH)n$^+$: 904.46 Da des-Arg-Bradykinin, 1296.68 Da Angiotensin I, 1570.67 Da Glu-Fibrinopeptide B, 2093.08 Da ACTH (clip 1–17), 2465.19 Da ACTH (clip 18–39), 3657.92 Da ACTH (clip 7–38)). For polyphenols identification, MS and tandem mass (MSMS) spectra were acquired using a mass (m/z) range of 100–4000 Da.

2.11. Statistical Analyses

In all the experiments samples were analyzed in triplicate. Quantitative parameters were expressed as the mean value ± SD. Significance was determined by Student's t-test at a significance level of 0.01. Differences among fractions were determined by using SPSS (Statistical Package for Social Sciences) Package 6, version 15.0 (SSPS Inc., Chicago, IL, USA). Significance was determined by Duncan's-test at a significance level of 0.05.
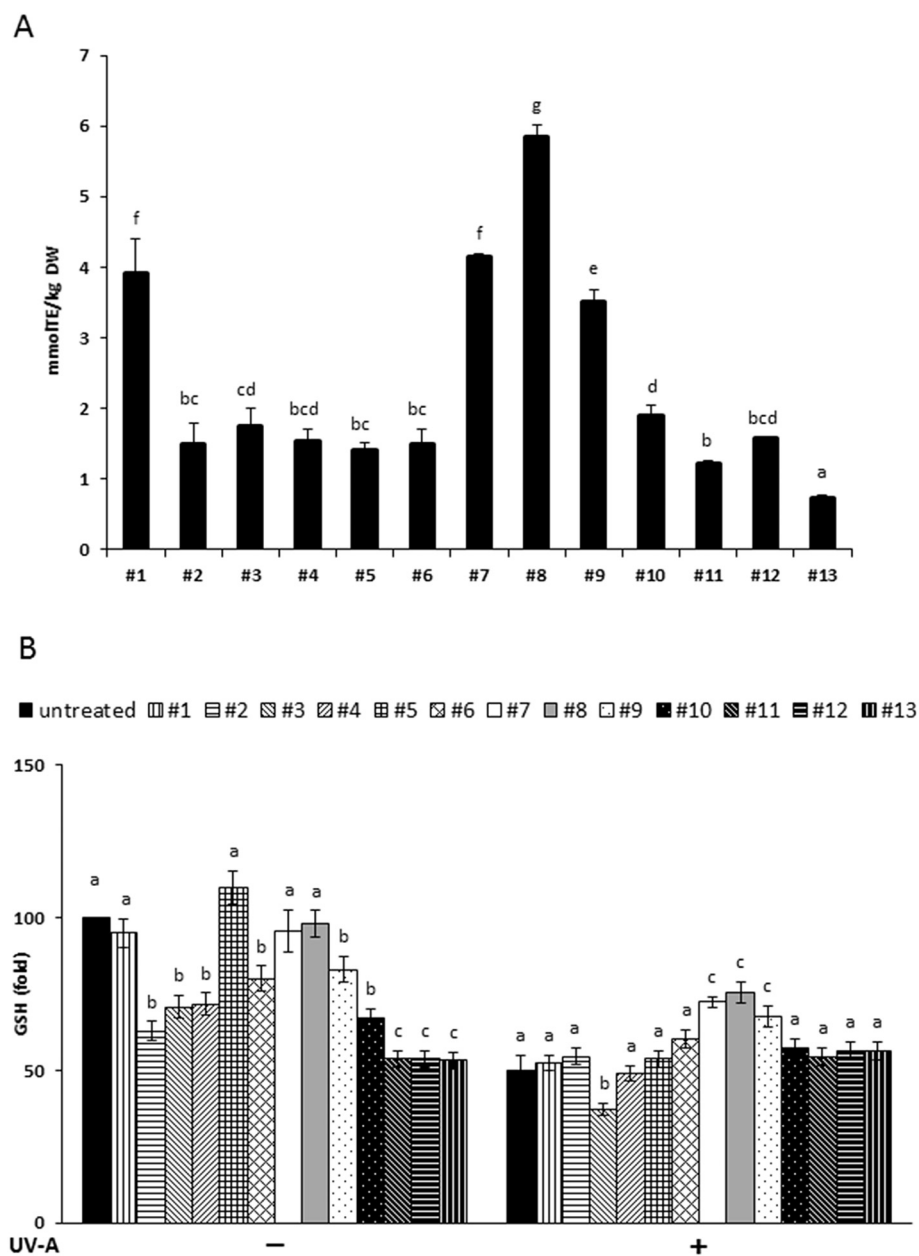
Fig. 5. Analysis of the antioxidant activity of açai fractions. A, Hydrophilic antioxidant activity (mmol TE/kg DW) in fractions evaluated by ABTS test. Data shown are the means ± S.D. of three independent experiments. Significance was determined by Duncan's-test at a significance level of 0.05. Values with different letters are significantly different. B, Changes in intracellular GSH levels of cells before (−) and after (+) UV-A treatment. Cells were incubated with açai extract, or equivalent amount of fractions obtained by HPLC, for 2 h prior to UV-A irradiation (100 J/cm^2). Data shown are the means ± S.D. of three independent experiments. Values with different letters are significantly different compared to untreated cells, indicated by black bars (p < 0.05).

3. Results

3.1. Antioxidant Compounds Determination

We first characterized the methanol/water açai extract by measuring its antioxidant compounds and determining the antioxidant activity. By colorimetric assays, the total phenol content was estimated to be 192.41 ± 10.78 mg GAE/100 g DW, whereas that of total flavonoids was found to be 159.67 ± 7.65 mg QE/100g DW. The mean antioxidant activity was 2.1 ± 0.17 mmol TE/100 g DW. Data are reported in Table 1.

3.2. Biocompatibility of Açai Extracts on Fibroblasts

The biocompatibility of the açai extract on BALB/3T3 fibroblasts

was tested by a dose-response test, in a range from 0.2 to 100 mg/mL of extract. As shown in Fig. 1, cell viability was not affected up to 40 mg/mL, whereas at the highest concentration tested (100 mg/mL) a 50% reduction of cell viability was observed. On the basis of these results, subsequent experiments were carried out at 10 mg/mL of açai extract, corresponding to a flavonoid content of 0.16 mg QE/mL and to a phenolic content of 0.19 mg GAE/mL.

3.3. Pretreatment With Açai Extracts Inhibits UV-A-Induced Damage in BALB/3T3 Fibroblasts

In all the experiments reported below, BALB/3T3 fibroblasts were pretreated with açai extract for 2 h before the induction of oxidative stress by UV-A irradiation (100 J/cm^2). Immediately after UV-A irradiation, ROS production was determined by using H$_2$DCF-DA (2,7-
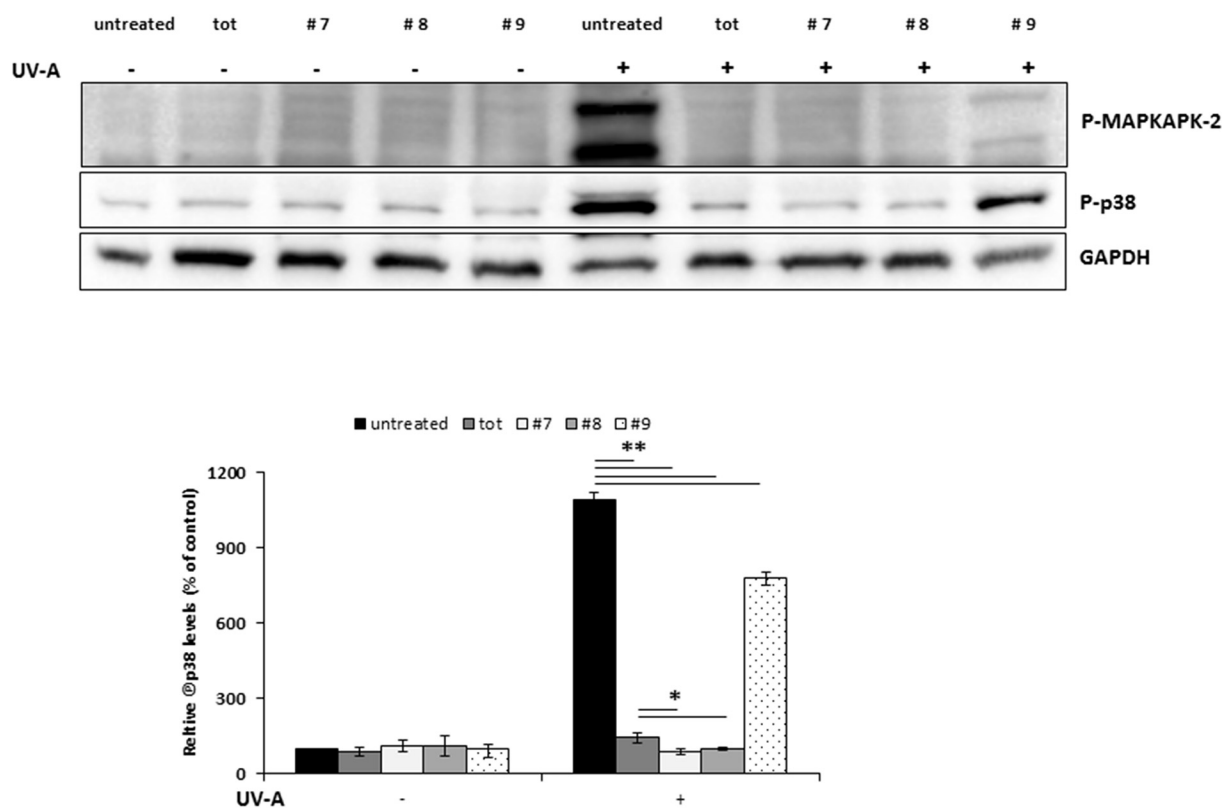
Fig. 6. Western blot analysis of UV-A-treated BALB/3T3 cells in the presence of HPLC fractions. Cells were incubated with different fractions for 2h prior to UV-A (100 J/cm^2) irradiation and then cells were incubated for further 90min at 37 °C. Representative western blots show the phosphorylation levels of p38 (A) and MAPKAPK-2 (B), with the relative densitometric analysis in the absence (black bars) or in the presence (white bars) of extract. GAPDH was used as internal standard. Data shown are the means ± S.D. of three independent experiments. * indicates p < 0.01, ** indicates p < 0.001.

dichlorofluorescin diacetate) (Fig. 2A). As expected, UV-A treatment significantly increased DCF fluorescence intensity, compared to the non-irradiated samples, whereas açai extract had no effect on ROS levels. Interestingly, pretreatment of cells with açai extract prior to UV-A exposure, resulted in the inhibition of ROS production. Moreover, following UV-A-oxidative stress induction, we found that intracellular GSH was significantly decreased with respect to untreated cells, whereas no alteration in intracellular GSH levels was observed when cells were pretreated with açai extract before being irradiated by UV-A (Fig. 2B). The protective effect of açai extract was confirmed by analyzing the lipid peroxidation levels after 90 min from irradiation. Lipid peroxidation level was measured by thiobarbituric acid reactive substances (TBARS). A significant increase in lipid peroxidation levels after UV-A treatment was observed but, noteworthy, this effect was abolished when cells were pretreated with açai extract. Again, pretreatment of cells with açai extract significantly counteracted the increase of lipid peroxidation (Fig. 2C).

The protective effect of açai extract was further confirmed by Western blot experiments, in which the phosphorylation levels of p38 and its direct target, MAPKAPK-2, were analyzed (Fig. 3). These proteins are directly involved in signaling stress pathways induced by UV-A [38]. When cells were UV-A irradiated, we observed a significant increase in the phosphorylation levels of p38 and MAPKAPK-2, as expected (Fig. 3A and B, third lane). Instead, although the treatment of non-irradiated cells with açai extract slightly altered the phosphorylation level of p38 (Fig. 3A, second lane), when cells were exposed to açai extract prior to UV-A treatment, the phosphorylation levels of p38 and MAPKAPK-2 were similar to those observed in non-irradiated cells (Fig. 3A and B, fourth lane). A similar trend was observed for HSP-27, a heat shock protein whose phosphorylation is directly related to oxidative stress [39,40] (Fig. 3C).

3.4. Antioxidant Activity of Açai Fractions

In order to identify which açai compounds were responsible for the antioxidant activity observed in irradiated fibroblasts, the extract was fractionated by HPLC on a reverse phase C18 column, from which the profile shown in Fig. 4 was obtained. 13 HPLC fractions were collected, as reported in the Figure, and their antioxidant activity was tested in vitro. The HAA analysis (Fig. 5A) revealed that the highest antioxidant activity was associated to fractions 7 and 8, for which the HAA mean values of 4.13 mmol TE/kg DW and 5.85 mmol TE/kg DW, respectively, were calculated. High antioxidant power was also identified in fractions 1 and 9, with HAA mean values of 3.92 and 3.49 mmol TE/kg DW, respectively. Lower values were detected in the other fractions, ranging from 0.72 mmol TE/kg DW (fraction 13) to 1.89 mmol TE/kg DW (fraction 10).

Then, we analyzed the intracellular GSH levels in BALB/3T3 cells pretreated with each HPLC fraction, before and after UV-A treatment. As shown in Fig. 5B, fractions 7 and 8 were able to contrast the detrimental effects of UV-A irradiation on GSH depletion, whereas fraction 9 was less efficient, in agreement with the HAA analysis. By contrast, fraction 1 did not show any protective effect from oxidative stress. Western blot experiments confirmed the strong antioxidant activity of fraction 7 and 8 on UV-A-stressed fibroblasts. In particular, the increase in p38 and MAPKAPK phosphorylation levels induced by UV-A irradiation was attenuated when cells were pretreated with açai fractions 7 and 8, whose effect was found to be slightly, but significantly, more pronounced than that of the whole extract (Fig. 6). Noticeable, a lower decrement in p38 phosphorylation level was observed when cells were pre-treated with fraction 9.
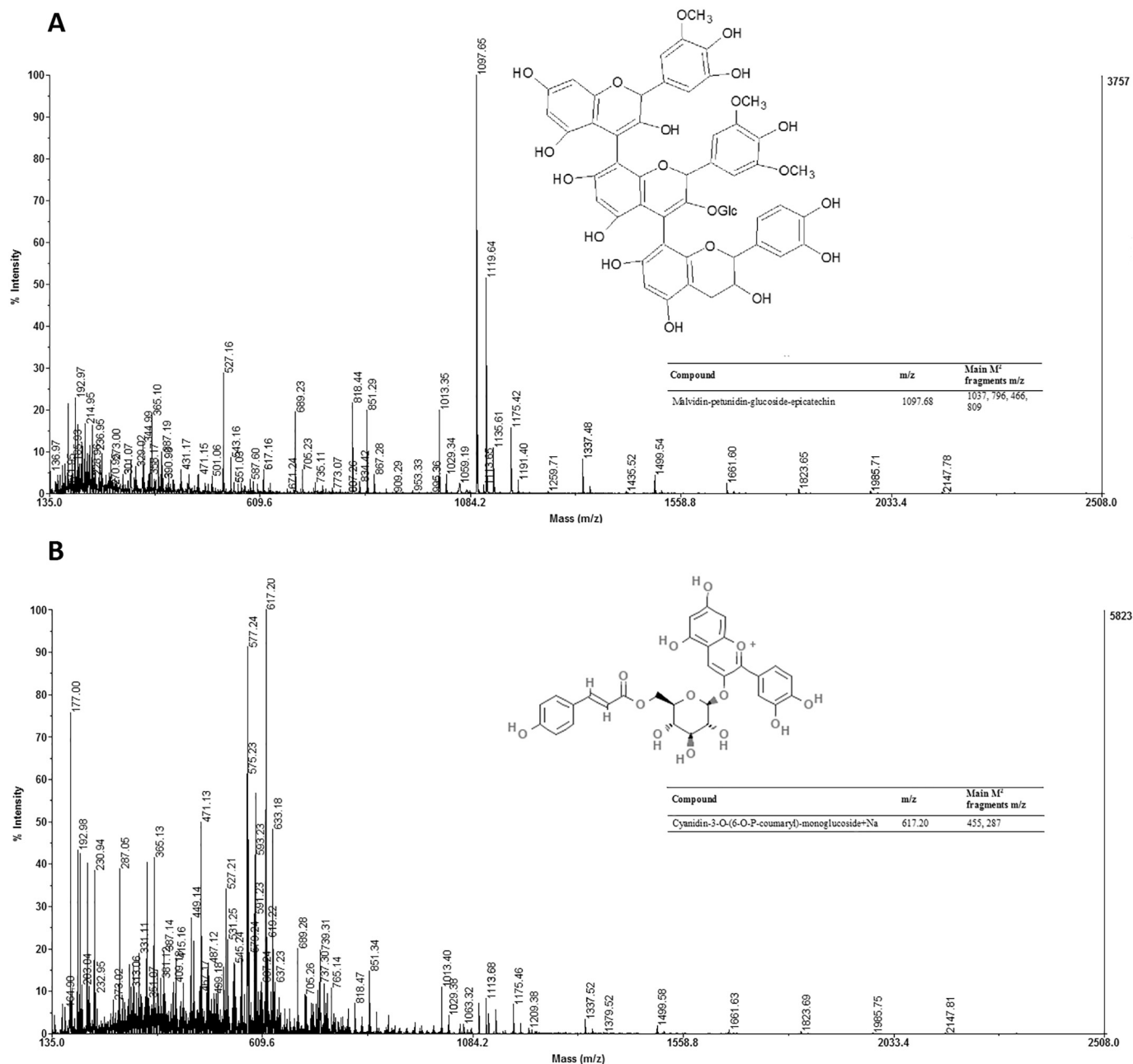
A

Compound	m/z	Main M² fragments m/z
Malvidin-petunidin-glucoside-epicatechin	1097.68	1037, 796, 466, 809

B

Compound	m/z	Main M² fragments m/z
Cyanidin-3-O-(6-O-P-coumaryl)-monoglucoside+Na	617.20	455, 287

Fig. 7. MALDI-MS analysis of selected HPLC fractions from açaí extract. The MALDIMS spectra of fractions 7 and 8 are reported in panels A and B, respectively. The attributions are reported.

3.5. Malvidin and Cyanidin are Responsible for the Açai Antioxidant Activity on Immortalized Fibroblasts

HPLC fractions 7 and 8 were submitted to MALDI-MS analysis, as described in the Material and Method section. The different species detected in the MALDI spectra (Fig. 7A-B) were identified on the basis of MSMS MALDI TOTOF analysis and compared with literature data [41]. Mass spectral identification is reported in Table 2. Qualitative and semi-quantitative differences can be appreciated in the MALDI spectra showing the presence of different flavonoids in açai fractions. Among them, malvidin-petunidin-glucoside-epicatechin was found to be the predominant compound in fraction 7, whereas cyanidin-3-O-(6-O-p-coumaryl)-monoglucoside was found to be the most abundant species in fraction 8. DCFDA assay indicated that malvidin and cyanidin enriched fractions were able to mitigate ROS production as well as the whole extract (Fig. 8A). However, the TBARS assay (Fig. 8B) revealed that the

activity of malvidin enriched fraction was comparable to that of the whole extract, whereas cyanidin enriched fraction was lower.

Finally, to evaluate if malvidin and cyanidin derivatives protect cells after UV-A injury and to avoid any sunscreen effect, fibroblasts were first irradiated by UV-A (100 J/cm²), and afterwards incubated for 2h with either açai extract or its fractions and then ROS, GSH levels and the phosphorylation levels of p38 were evaluated (Fig. 9). Interestingly, UV-A induced ROS production was mitigated by the presence of either açai extract or its active fractions. Furthermore, no alteration in intracellular GSH levels was measured and a significant decrease in the phosphorylation levels of p38 was also observed.

4. Discussion

Among the most severe environmental injuries to which humans are exposed, UV-A radiations are known to be the main cause of different

Table 2

MALDI mass spectral analysis of açai fractions after reverse-phase HPLC. The identified species and the relative percentages are reported (G = glucose).

Compound	m/z	Main M^2 fragments m/z	Fraction 7	Fraction 8
Petunidin-3-acetyl	365.10	305, 203	20%	43%
Peonidin-3-*O*-acetyl monoglucoside	527.19	467, 365, 203	32%	38%
Cyanidin-3-O-(6-O-P-coumaryl)-monoglucoside + Na	617.20	455, 287	10%	100%
Petunidin-3-acetyl diglucoside	689.23	527, 629	22%	20%
Malvidin-3-acetylglucose-ethil-catechin/ Petunidin-3-acetil triglucoside	851.31	791, 689	20%	17%
Petunidin-3-acetyl tetraglucoside	1013.35	952, 851	20%	14%
Malvidin-petunidin-glucoside-epicatechin	1097.68	1037, 796, 466, 809	100%	5%
Petunidin-3-acetyl + 5G	1175.42	1115, 1013	16%	8%
Petunidin-3-acetyl + 6G	1337.48	1277,1175	12%	5%
Petunidin-3-acetyl + 7G	1499.54	1439, 1337	8%	3%
Petunidin-3-acetyl + 8G	1661.6	1601, 1499	6%	3%
Petunidin-3-acetyl + 9G	1823.65	1763, 1661	6%	2%
Petunidin-3-acetyl + 10G	1985.71	1925, 1823	4%	1%
Petunidin-3-acetyl + 11G	2147.78	2087, 1985	2%	1%

detrimental effects on the skin, such as inflammation, premature skin aging and development of skin cancer [42,43]. In recent years, an increasing body of research has focused on antioxidants since they are able to counteract oxidative stress-induced pathological conditions. However, studies are usually conducted on commercial molecules, supposed to have antioxidant activity [43–45].

Here, we analyzed hydrophilic extracts from açai berries to evaluate their dermato-protective properties against UV-A irradiation. Açai is the name commonly used for the tree *Euterpe oleracea Martius*, a South American palm [25,46], whose berry is largely consumed since it shows several beneficial activities, including anti-inflammatory and antioxidant properties. These activities are mainly due to açai high content in polyphenol content, especially anthocyanins and flavones [20–25].

Immortalized fibroblasts, exposed to UV-A radiations, showed increased levels of intracellular ROS and lipid peroxidation, and decreased levels of intracellular GSH. Noteworthy, we found that a pre-treatment of cells with açai extracts clearly inhibited the detrimental effect of oxidative stress, since ROS generation, lipid peroxidation, intracellular GSH levels and p38 phosphorylation levels were all found to be almost unaltered with respect to UV-A unexposed cells. It is

interesting to notice that the açai extract is able to efficiently protect immortalized, non-cancer cells from the deleterious effects of UV-A irradiation when tested at concentrations compatible with cell viability, i.e. non associated to any cytotoxicity. Moreover, we demonstrated that the protective effect of açai extract is not due to a sunscreen effect, as the extract was able to protect cells also after UV-A induced damage.

Mass spectrometry analyses of HPLC fractionated extract led us to the identification of the main agents responsible for the potent antioxidant activity of açai extracts on fibroblasts. These are two flavonoids, cyanidin and malvidin, both endowed with strong antioxidant activity, with malvidin the most active.

Cyanidin is considered the widest spread anthocyanin in the plant kingdom and it is known to be a strong natural antioxidant. Amorini and colleagues [47] demonstrated that cyanidin-3-glucoside, also known as kuromanin, has a remarkable antioxidant capacity in the model of Cu^{2+}-mediated human low density lipoprotein oxidation, even higher than resveratrol and ascorbic acid and independent from pH variations (from 4 to 7.4). In addition, it is reported that cyanidin-3-glucoside may attenuate obesity-associated insulin resistance and hepatic steatosis in high-fat diet-fed mice via the transcription factor FoxO1 [48].

Malvidin, normally present at high concentration in red wine and black rice, and responsible for their color, plays an important role in protecting plants from UV irradiation [49,50]. Accordingly, it has been reported that malvidin has antioxidant activity in vitro [51], as well as in neuronal and in endothelial cells [52,53]. Recently, it has been demonstrated that malvidin is also able to inhibit aging by attenuating oxidative stress in human diploid fibroblasts [54]. Moreover, several authors have shown that malvidin inhibits the growth of different tumor cell lines in vitro or in vivo [55–57].

Polyphenols are known to inhibit free radicals' production by donating hydrogen atoms, which break oxidation chains and chelate transition metal ions. Açai polyphenols may upregulate antioxidant enzymes by activating the Nrf2/Keap1 pathway [58], downregulate the expression of the pro-inflammatory transcription factor NF-kB, which in turn targets pro-inflammatory cytokines, and reduces the accumulation of intracellular lipids in differentiated adipocytes by downregulating the expression of adipogenic genes transcription factors [59]. Additionally, different studies performed on healthy women, which consumed 200 g/day of açai pulp for 4 weeks, revealed that some antioxidant enzymes, i.e. superoxide dismutase, catalase and glutathione, were increased [59–61].

Our findings now reveal that açai extract protects BALB/3T3 cells against UV-A irradiation by neutralizing the negative effects of stress induced by UV-A, as it interferes with ROS generation and keeps intracellular GSH and lipid peroxidation close to the normal cellular levels. This work reveals, for the first time, that malvidin present in açai
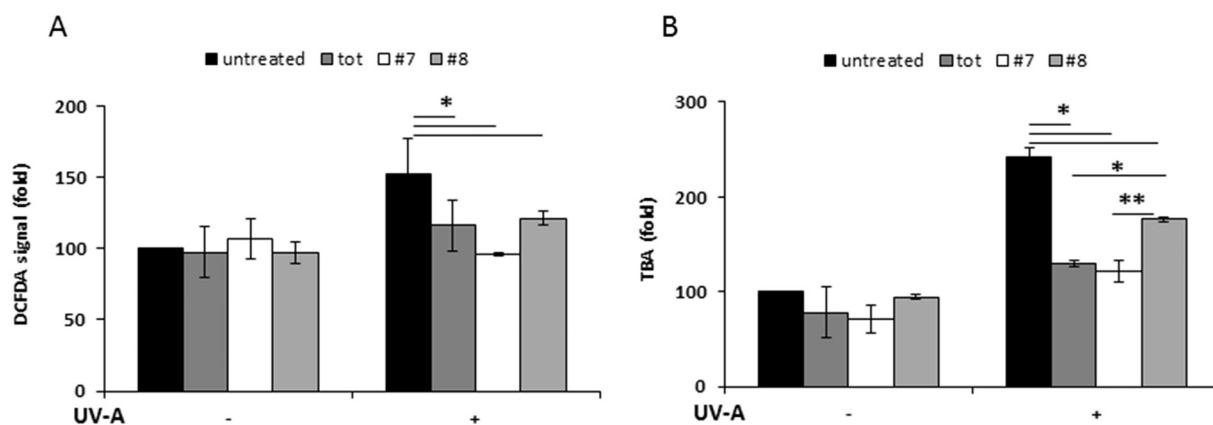

Fig. 8. ROS production and lipid peroxidation in UV-A-treated BALB/3T3 fibroblasts in the presence of açai fractions. Cells were pre-incubated with açai extract (dark grey bars), fraction 7 (white bars) and fraction 8 (light grey bars) for 2h and then irradiated by UV-A (100 J/cm²). A, intracellular ROS levels; B, lipid peroxidation levels. Values are expressed as fold increase with respect to control (i.e. untreated) cells. Data shown are the means ± S.D. of three independent experiments. * indicates p < 0.01; ** indicates p < 0.05.
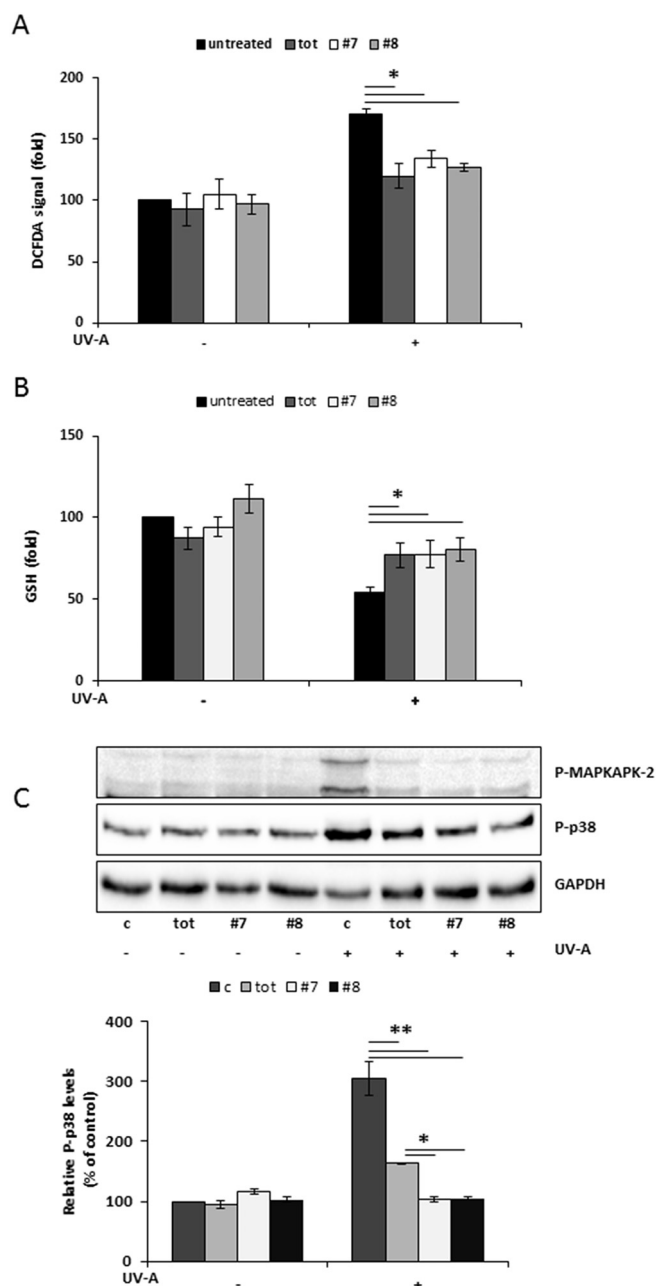
Fig. 9. Analysis of oxidative stress in BALB/3T3 fibroblasts incubated in the presence of açai fractions after UV-A treatment. Cells were irradiated by UV-A (100 J/cm^2) and then incubated with açai extract, fraction 7 or fraction 8 for 2h. A, intracellular ROS levels; B, intracellular GSH levels; C, Western blotting of the phosphorylation levels of p38 and MAPKAPK-2, with the relative densitometric analysis. GAPDH was used as internal standard. In all panels, untreated cells (named c in western blotting analysis) are represented by black bars, cells incubated with açai extract by dark grey bars, with fraction 7 by white bars and with fraction 8 by light grey bars. Values are expressed as fold increase with respect to control (i.e. untreated) cells. Data shown are the means ± S.D. of three independent experiments. * indicates $p < 0.05$; ** indicates $p < 0.01$.

extract can be considered as a good and safe candidate for dietary intervention in the prevention of age related diseases, since the use of anthocyanins as food additives (indicated as "E163") has been already approved in the European Union, Australia, and New Zealand [62].

Author contribution

G.P. and D.M.M. conceived, performed and analyzed cell biology experiments; A.R. and M.M.R. performed in vitro experiments; R.D.G.

performed spectrophluorimetric experiments; A.A and A.I. conceived, performed and analyzed the HPLC and MS experiments; R.P. analyzed and discussed the data; D.M.M. wrote the paper with the contribution of all the authors.

Conflict of interest

The authors declare no competing financial interest.

References

[1] R. Thanan, S. Oikawa, Y. Hiraku, S. Ohnishi, N. Ma, S. Pinlaor, et al., Oxidative stress and its significant roles in neurodegenerative diseases and cancer, Int. J. Mol. Sci. 16 (2014) 193–217.

[2] D.R. Bickers, M. Athar, Oxidative stress in the pathogenesis of skin disease, J. Invest. Dermatol. 126 (2006) 2565–2575.

[3] S. Briganti, M. Picardo, Antioxidant activity, lipid peroxidation and skin diseases. What's new, J. Eur. Acad. Dermatol. Venereol. 17 (2003) 663–669.

[4] J. Cadet, R. Teoule, Comparative study of oxidation of nucleic acid components by hydroxyl radicals, singlet oxygen and superoxide anion radicals, Photochem. Photobiol. 28 (1978) 661–665.

[5] J.L. Ravanat, C. Saint-Pierre, P. Di Mascio, G.R. Martinez, M.H.G. Medeiros, J. Cadet, Damage to isolated DNA mediated by singlet oxygen, Helv. Chim. Acta 84 (2001) 3702–3709.

[6] T. Douki, A. Reynaud-Angelin, J. Cadet, E. Sage, Bipyrimidine photoproducts rather than oxidative lesions are the main type of DNA damage, Biochemistry 42 (2003) 9221–9226.

[7] S. Raha, B.H. Robinson, Mitochondria, oxygen free radicals, disease and ageing, Trends Biochem. Sci. 25 (2000) 502–508.

[8] L.A. Pham-Huy, H. He, C. Pham-Huy, Free radicals, antioxidants in disease and health, Int. J. Biomed. Sci. 4 (2008) 89.

[9] L. Gil del Valle, Oxidative stress in aging: theoretical outcomes and clinical evidences in humans, Biomed. Aging Pathol. 1 (2011) 1–7.

[10] A. Scalbert, I.T. Johnson, M. Saltmarsh, Polyphenols: antioxidants and beyond, Am. J. Clin. Nutr. 81 (2005) 215S–217S.

[11] B.E. van Wyk, M. Wink, Phytomedicines, Herbal Drugs, and Poisons, first ed., (2015) (Chicago, IL).

[12] M. Wink, Modes of action of herbal medicines and plant secondary metabolites, Fortschr. Med. 2 (2015) 251–286.

[13] R. Del Giudice, G. Petruk, A. Raiola, A. Barone, D.M. Monti, M.M. Rigano, Carotenoids in fresh and processed tomato (Solanum lycopersicum) fruits protect cells from oxidative stress injury, J. Sci. Food Agric. 97 (2017) 1616–1623, http://dx.doi.org/10.1002/jsfa.7910.

[14] G. Petruk, A. Raiola, R. Del Giudice, A. Barone, L. Frusciante, M.M. Rigano, D.M. Monti, An ascorbic acid-enriched tomato genotype to fight UV-A-induced oxidative stress in normal human keratinocytes, J. Photochem. Photobiol. B 163 (2016) 284–289.

[15] H. Rogez, Açai: Preparo, Composicao e Conservacao, first ed., EDUFPA, Brazil, 2000.

[16] M.S.M. Rufino, J. Perez-Jimenez, S. Arranz, R.E. Alves, E.S. de Brito, M.S. Oliveira, F. Saura-Calixto, Açai (Euterpe oleracea) 'BRS Para': a tropical fruit source of antioxidant dietary fiber and high antioxidant capacity oil, Food Res. Int. 44 (2011) 2100–2106.

[17] J. Holderness, I.A. Schepetkin, B. Freedman, L.N. Kirpotina, M.T. Quinn, J.F. Hedges, M.A. Jutila, Polysaccharides isolated from Açai fruit induce innate immune responses, PLoS One 6 (2011) e17301.

[18] H. Peixoto, M. Roxo, S. Krstin, T. Röhrig, E. Richling, M. Wink, An Anthocyanin-rich extract of açai (Euterpe precatoria Mart.) increases stress resistance and retards aging-related markers in Caenorhabditis elegans, J. Agric. Food Chem. 64 (2016) 1283–1290.

[19] M.M. dos Santos Dias, H.S.D. Martino, G. Noratto, A. Roque-Andrade, P.C. Stringheta, S. Talcott, A.M. Ramos, S.U. Mertens-Talcott, Anti-inflammatory activity of polyphenolics from açai (Euterpe oleracea Martius) in intestinal myofibroblasts CCD-18Co cells, Food Funct. 6 (2015) 3249–3256.

[20] R. Lichtenthaler, R.B. Rodrigues, J.G.S. Maia, M. Papagiannopoulos, H. Fabricius, F. Marx, Total oxidant scavenging capacities of Euterpe oleracea Mart. (Açai) fruits, Int. J. Food Sci. Nutr. 56 (2005) 53–64.

[21] S.U. Mertens-Talcott, J. Rios, P. Jilma-Stohlawetz, L.A. Pacheco-Palencia, B. Meibohm, S.T. Talcott, H. Derendorf, Pharmacokinetics of anthocyanins and antioxidant effects after the consumption of anthocyanin-rich açai juice and pulp (Euterpe oleracea Mart.) in human healthy volunteers, J. Agric. Food Chem. 56 (2008) 7796–7802.

[22] J. Kang, C. Xie, Z. Li, S. Nagarajan, A.G. Schauss, T. Wu, X. Wu, Flavonoids from açai (Euterpe oleracea Mart.) pulp and their antioxidant and anti-inflammatory activities, Food Chem. 128 (2011) 152–157.

[23] J. Kang, K.M. Thakali, C. Xie, M. Kondo, Y. Tong, B. Ou, G. Jensen, M.B. Medina,

A.G. Schauss, X. Wu, Bioactivities of açai (*Euterpe precatoria Mart.*) fruit pulp, superior antioxidant and anti-inflammatory properties to *Euterpe oleracea Mart*, Food Chem. 133 (2012) 671–677.

[24] H.A. Favacho, B.R. Oliveira, K.C. Santos, B.J. Medeiros, P.J. Sousa, F.F. Perazzo, J.C.T. Carvalho, Anti-inflammatory and antinociceptive activities of *Euterpe oleracea Mart.*, Arecaceae, oil, Rev. Bras 21 (2011) 105–114.

[25] A.G. Schauss, X. Wu, R.L. Prior, B. Ou, D. Huang, J. Owens, A. Agarwal, G.S. Jensen, A.N. Hart, E. Shanbrom, Antioxidant capacity and other bioactivities of the freeze-dried Amazonian palm berry, *Euterpe oleracea Mart.* (açai), J. Agric. Food Chem. 54 (2006) 8604–8610.

[26] G.D. Noratto, G. Angel-Morales, S.T. Talcott, S.U. Mertens-Talcott, Polyphenolics from açai (*Euterpe oleracea Mart.*) and red muscadine grape (*Vitis rotundifolia*) protect human umbilical vascular Endothelial cells (HUVEC)from glucose and lipopolysaccharide (LPS)-induced inflammation and target micro RN A-126, J. Agric. Food Chem. 59 (2011) 7999–8012.

[27] G.D. Stoner, L.S. Wang, C. Seguin, C. Rocha, K. Stoner, S. Chiu, A.D. Kinghorn, Multiple berry types prevent N-nitrosomethylbenzylamine-induced esophageal cancer in rats, Pharm. Res. 27 (2010) 1138–1145.

[28] M. Oliveira de Souza, M. Silva, M.E. Silva, R. de Paula Oliveira, M.L. Pedrosa, Diet supplementation with açai (*Euterpe oleracea Mart.*) pulp improves biomarkers of oxidative stress and the serum lipid profile in rats, Nutrition 26 (2010) 804–810.

[29] J.F. Ferreira da Costa Guerra, C.L. Lopes de Brito Magalhaes, D.C. Costa, M.E. Silva, M.L. Pedrosa, Dietary açai modulates ROS production by neutrophils and gene expression of liver antioxidant enzymes in rats, J. Clin. Biochem. Nutr. 49 (2011) 188–194.

[30] M.M. Rigano, A. Raiola, G.C. Tenore, D.M. Monti, R. Del Giudice, L. Frusciante, A. Barone, Quantitative trait loci pyramiding can improve the nutritional potential of tomato (*Solanum lycopersicum*) fruits, J. Agric. Food Chem. 62 (2014) 11519–11527.

[31] V.L. Singleton, J.A. Rossi, Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents, Am. J. Enol. Vitic. 16 (1965) 144–158.

[32] D. Marinova, F. Ribarova, M. Atanassova, Total phenolics and total flavonoids in bulagrian fruits and vegetables, J. Univ. Chem. Technol. Metall. 40 (2005) 255–260.

[33] A. Raiola, R. Del Giudice, D.M. Monti, G.C. Tenore, A. Barone, M.M. Rigano, Bioactive compound content and cytotoxic effect on human cancer cells of fresh and processed yellow tomatoes, Molecules 21 (2016) 33.

[34] J.N. Miller, C.A. Rice-Evans, Factors influencing the antioxidant activity determined by the ABTS$^{\bullet+}$ radical cation assay, Free Radic. Res. 26 (1997) 195–199.

[35] R. Del Giudice, A. Raiola, G.C. Tenore, L. Frusciante, A. Barone, D.M. Monti, M.M. Rigano, Antioxidant bioactive compounds in tomato fruits at different ripening stages and their effects on normal and cancer cells, J. Funct. Foods 18 (2015) 83–94.

[36] D.M. Monti, D. Guarnieri, G. Napolitano, R. Piccoli, P. Netti, S. Fusco, A. Arciello, Biocompatibility, uptake and endocytosis pathways of polystyrene nanoparticles in primary human renal epithelial cells, J. Biotechnol. 10 (2015) 3–10.

[37] E. Galano, A. Arciello, R. Piccoli, D.M. Monti, A. Amoresano, A proteomic approach to investigate the effects of cadmium and lead on human primary renal cells, Metallomics 6 (2014) 587–597.

[38] L.O. Klotz, C. Pellieux, K. Briviba, C. Pierlot, J.M. Aubry, H. Sies, Mitogen-activated protein kinase (p38-, JNK-, ERK-) activation pattern induced by extracellular and intracellular singlet oxygen and UV-A, Eur. J. Biochem. 260 (1999) 917–922.

[39] W.R. Swindell, Heat shock proteins in long-lived worms and mice with insulin/insulin-like signaling mutations, Aging 1 (2009) 573.

[40] A. Strayer, Z. Wu, Y. Christen, C.D. Link, Y. Luo, Expression of the small heat-shock protein Hsp16 − 2 in Caenorhabditiselegans is suppressed by *Ginkgo biloba* extract EGb 761, FASEB J. 17 (2003) 2305–2307.

[41] R. Flamini, Mass spectrometry in grape and wine chemistry. Part I: polyphenols, Mass Spectrom. Rev. 22 (2003) 218–250.

[42] D. Kulms, T. Schwarz, Molecular mechanisms of UV-induced apoptosis, Photodermatol. Photoimmunol. Photomed. 16 (2000) 195–201.

[43] Y.-C. Hseu, C.W. Chou, K.J. Senthil Kumar, K.T. Fu, H.M. Wang, L.S. Hsu, Y.H. Kuo, C.R. Wu, S.C. Chen, H.L. Yang, Ellagic acid protects human keratinocyte (HaCaT) cells against UV-A-induced oxidative stress and apoptosis through the upregulation of the HO-1 and Nrf-2 antioxidant genes, Food Chem. Toxicol. 50 (2012) 1245–1255.

[44] Y.-C. Hseu, H.W. Lo, M. Korivi, Y.C. Tsai, M.J. Tang, H.L. Yang, Dermato-protective properties of ergothioneine through induction of Nrf2/ARE-mediated antioxidant genes in UV-A-irradiated human keratinocytes, Free Radic. Biol. Med. 86 (2015) 102–117.

[45] I.F. Almeida, A.S. Pinto, C. Monteiro, H. Monteiro, L. Belo, J. Fernandes, A.R. Bento, T.L. Duarte, J. Garrido, M.F. Bahia, J.M. Sousa Lobo, P.C. Costa, Protective effect of *C. sativa* leaf extract against UV mediated-DNA damage in a human keratinocyte cell line, J. Photochem. Photobiol. B 144 (2015) 28–34.

[46] A.G.V. Costa, D.F. Garcia-Diaz, P. Jimenez, P.I. Silva, Bioactive compounds and health benefits of exotic tropical red–black berries, J. Funct. Foods 5 (2013) 539–549.

[47] A.M. Amorini, G. Fazzina, G. Lazzarino, B. Tavazzi, D. Di Pierro, R. Santucci, F. Sinibaldi, F. Galvano, G. Galvano, Activity and mechanism of the antioxidant properties of cyanidin-3-O-βglucopyranoside, Free Radic. Res. 35 (2001) 953–966.

[48] H. Guo, M. Xia, T. Zou, W. Ling, R. Zhong, W. Zhang, Cyanidin 3-glucoside attenuates obesity-associated insulin resistance and hepatic steatosis in high-fat diet-fed and db/db mice via the transcription factor FoxO1, J. Nutr. Biochem. 23 (2012) 349–360.

[49] U. Matern, B. Grimmig, Polyphenols in plant pathology, Polyphenolic Phenomena. Paris: INRA Editions, 1993, pp. 143–147.

[50] E. Bognar, Z. Sarszegi, A. Szabo, B. Debreceni, N. Kalman, Z. Tucsek, B. Sumegi, F. GallyasJr, Antioxidant and anti-inflammatory effects in RAW264. 7 macrophages of malvidin, a major red wine polyphenol, PLoS One 8 (2013) e65355.

[51] R. Pop, M.N. Stefanut, A. Cata, C. Tanasie, M. Medeleanu, Ab initio study regarding the evaluation of the antioxidant character of cyanidin, delphinidin and malvidin, Cent. Eur. J. Chem. 10 (2012) 180–186.

[52] N. Matsunaga, S. Imai, Y. Inokuchi, M. Shimazawa, S. Yokota, Y. Araki, H. Hara, Bilberry and its main constituents have neuroprotective effects against retinal neuronal damage *in vitro* and *in vivo*, Mol. Nutr. Food Res. 53 (2009) 869–877.

[53] W. Huang, Y. Zhu, C. Li, Z. Sui, W. Min, Effect of blueberry *Anthocyanins Malvidin* and glycosides on the antioxidant properties in endothelial cells, Oxidative Med. Cell. Longev. (2016), http://dx.doi.org/10.1155/2016/1591803.

[54] H.R. Seo, M.J. Choi, J.M. Choi, J.C. Ko, J.Y. Ko, E.J. Cho, Malvidin protects WI-38 human fibroblast cells against stress-induced premature senescence, Eur. J. Cancer Prev. 21 (2016) 32–40.

[55] J.W. Hyun, H.S. Chung, Cyanidin and malvidin from *Oryza sativa* cv. Heugjinjubyeo mediate cytotoxicity against human monocytic leukemia cells by arrest of G2/M phase and induction of apoptosis, J. Agric. Food Chem. 52 (2004) 2213–2217.

[56] P.H. Shih, C.T. Yeh, G.C. Yen, Effects of anthocyanidin on the inhibition of proliferation and induction of apoptosis in human gastric adenocarcinoma cells, Food Chem. Toxicol. 43 (2005) 1557–1566.

[57] S.J. Patterson, J.G. Fischer, R.V. Dulebohn, DNA damage in HT-29 colon cancer cells is enhanced by high concentrations of the anthocyanin malvidin, FASEB J. 22 (2008) 890–900.

[58] W. Li, Y. Guo, C. Zhang, R. Wu, A.Y. Yang, J. Gaspar, A.N. Kong, Dietary phytochemicals and cancer chemoprevention: a perspective on oxidative stress, inflammation, and epigenetics, Chem. Res. Toxicol. 29 (2016) 2071–2095.

[59] H.S.D. Martino, M.M. dos Santos Dias, G. Noratto, S. Talcott, S.U. Mertens-Talcott, Anti-lipidaemic and anti-inflammatory effect of açai (*Euterpe oleracea Martius*) polyphenols on 3T3-L1 adipocytes, J. Funct. Foods 23 (2016) 432–443.

[60] P.O. Barbosa, D. Pala, C.T. Silva, M.O. de Souza, J.F. do Amaral, R.A.L. Vieira, G.A. Folly, A.C. Volp, R.N. de Freitas, Açai (*Euterpe oleracea* Mart.) pulp dietary intake improves cellular antioxidant enzymes and biomarkers of serum in healthy women, Nutrition 32 (2016) 674–680.

[61] D. Pala, P.O. Barbosa, C.T. Silva, M.O. de Souza, F.R. Freitas, A.C. Volp, R.C. Maranhão, R.N. Freitas, Açai (*Euterpe oleracea Mart.*) dietary intake affects plasma lipids, apolipoproteins, cholesteryl ester transfer to high-density lipoprotein and redox metabolism: a prospective study in women, Clin. Nutr. (2017), http://dx.doi.org/10.1016/j.clnu.2017.02.001.

[62] C.K. Singh, I.A. Siddiqui, S. El-Abd, H. Mukhtar, N. Ahmad, Combination chemoprevention with grape antioxidants, Mol. Nutr. Food Res. 60 (2016) 1406–1415.

Simultaneous production of antioxidants and starch from the microalga *Chlorella sorokiniana*

Ganna Petruk[a], Imma Gifuni[b], Anna Illiano[a], Mariana Roxo[c], Gabriella Pinto[a], Angela Amoresano[a], Antonio Marzocchella[b], Renata Piccoli[a,d], Michael Wink[c], Giuseppe Olivieri[b,e,*], Daria Maria Monti[a,d,**]

[a] *Department of Chemical Sciences, University of Naples Federico II, Complesso Universitario Monte Sant'Angelo, via Cinthia 4, 80126 Naples, Italy*
[b] *Dipartimento di Ingegneria Chimica, dei Materiali e della Produzione Industriale, Università degli Studi di Napoli Federico II, Piazzale Tecchio, 80, Naples, Italy*
[c] *Institute of Pharmacy and Molecular Biotechnology, Heidelberg University, INF 364, 69120 Heidelberg, Germany*
[d] *Istituto Nazionale di Biostrutture e Biosistemi (INBB), Rome, Italy*
[e] *Bioprocess Engineering Group, Wageningen University and Research, Droevendaalsesteeg 1, 6700AA Wageningen, the Netherlands*

ARTICLE INFO

ABSTRACT

In recent years, microalgae have gained considerable importance as potential source of biofuels and bioplastics. However, these markets are still developing, as the high costs of cultivation ask for exploiting microalgae into new areas and with a biorefinery approach towards a multicomponent cascade extraction process. Here, a sequential processing strategy was used to extract starch with high yield from *Chlorella sorokiniana* under biocompatible conditions. The extract residue was then tested as a potential source of antioxidants. We found a strong protective activity of the extract residue towards oxidative stress *in vitro* on human colon cancer cells and *in vivo* on *Caenorhabditis elegans* nematodes, by inhibiting ROS production and activating DAF-16/FOXO transcription factor pathway. A pool of molecules from three different classes (fatty acids, photosynthetic pigments and carotenoids) was identified as responsible for the antioxidant activity. To our knowledge, this is the first report on the obtainment, from a "waste" fraction, of high value products endowed with antioxidant activity tested in cell-based models and *in vivo*.

1. Introduction

The circular bioeconomy, based on the biorefinery approach, seems to be a good and long-term solution to harmonise the use of natural resources to sustain economic growth, ensure human wellness, reduce the increase of CO_2 concentration in the atmosphere and reduce waste production. Microalgae have been often regarded as a potential candidate for biorefinery as they are an incredible reservoir of compounds endowed with biological activities that could be used in different fields [1].

Microalgal biomass is renewable, as CO_2 and sun light energy are both required for their growth. As no competition exists with food culture for the use of arable land, many recent studies have been focused on the exploitation of various microalgal components in order to obtain the maximum economic benefit, as the conversion of biomass into different products produces minimal waste to the environment. Currently, the research is mainly focused on the use of microalgae as sources of lipids for biodiesel production [2], carbohydrates for methane production *via* anaerobic fermentation [3], proteins for animal nutrition [4] and food additives. Several process alternatives have been proposed to exploit microalgal components in a multi-product biorefinery approach [5]. However, most of these fractions have been tested only as raw fractions, not characterized in their single components to validate them.

In our experimental system, we found that *C. sorokiniana*, requires, for starch accumulation, nitrogen depletion [6]. This stressful condition, in turn, can be exploited to obtain pigments endowed with antioxidant activity. In principle, the cultivation conditions (light intensity and nutrient supply strategy) must be selected to optimize the accumulation of all the main products to be extracted within the biorefinery.

It is known that microalgae, when exposed to high oxygen and radical stress in their natural environment, develop several efficient protective systems against reactive oxygen species (ROS) and free radicals [7], by producing pigments, such as chlorophylls, carotenoids, xanthophylls and phycobiliproteins [8]. In particular, astaxanthin, β-carotene, zeaxanthin and lutein possess antioxidant activity exceeding that of α-tocopherol [9] and are required from food, cosmetic, nutraceutical and pharmaceutical industries. Thus, the obtainment of an active pigment fraction can significantly influence the economic balance of the whole process (cultivation plus biorefinery).

The main objective of this contribution was to obtain a second high-value class of molecules, such as antioxidants, starting from the discarded fraction after starch extraction [6]. Moreover, we identified the antioxidants obtained and validated their activity in a real applicative environment, *i.e.* both in cell-based models and *in vivo* on a *C. elegans* model, which is widely used in this context [10]. This contribution provides a clear validation of the feasibility of a multiproduct biorefinery applied to *C. sorokiniana* at lab-scale.

2. Materials and methods

2.1. Microalgal strain and culture conditions

Chlorella sorokiniana Shihiraet Krauss strain ACUF 318 (http://www.acuf.net) is a fresh water terrestrial alga. The culture medium used was Bold's Basal Medium (BBM) containing $NaNO_3$ as nitrogen source at concentration of $0.25 \, g \, L^{-1}$. The growth was carried out in inclined square bubble column photobioreactors, as reported by Olivieri et al. [11]. The culture was continuously irradiated with white fluorescent lamps on the front side of the reactors with a light intensity of $300 \, \mu mol \, m^{-2} \, s^{-1}$. Aeration and mixing was provided by feeding air supplemented with 2% of CO_2. Gas flow rate was set at 0.2 vvm. The temperature was maintained at 25 °C by air-conditioning system. Nitrogen depletion was reached after the regular uptake of $NaNO_3$ for biomass growth and the biomass was harvested in the second day of nitrogen depletion, optimal condition for starch accumulation and productivity [12].

2.2. Pigment extraction

A conventional solvent extraction method was carried out to obtain starch and antioxidants from *C. sorokiniana*. Pure ethanol (99.8%, Sigma-Aldrich) was used as solvent. The extraction protocol follows that reported by Aremu et al. [13] with some modifications. Briefly, the microalgal biomass was dried and ground to 0.150 mm powder. 200 mg of microalgal powder were suspended in 2 mL of ethanol and disrupted in bead beater (2000 g for 3 cycles of 1 min space out with 2 min breaks in ice) equipped with 0.5 mm glass beads. The disrupted biomass was moved to dark flasks and the beads were washed twice with ethanol and the final volume was led to 20 mL. The mixture was shaken for 24 h at 250 rpm in the dark on a magnetic stirrer. Then the mixture was centrifuged at 12000g for 10 min. The supernatant was dried by using N_2 stream, and represents the ethanol extract (EE), which was solubilized in 1 mL of DMSO for further analysis on a cell-based model and *in vivo*. The recovered pellet (EP) was dried to recover starch.

2.3. Biochemical composition

Initial microalgal biomass, EE and EP were characterized in terms of lipids, proteins, total sugars, starch and pigments. Lipids were extracted according to the protocol proposed by Breuer et al. [14] followed by gravimetric quantification of the chloroform extract, previously dried at 100 °C for 1 h.

Proteins were assayed by BCA Protein Assay Kit (Thermo Scientific). 10 mg of sample (biomass or dried pellet) was dissolved in 1 mL of lysis buffer (60 mM Tris, 2% SDS) and disrupted in a bead beater (700 g for

3 cycles of 4 min space out with 1 min breaks in ice) equipped with 0.5 mm glass beads. The samples were incubated at 100 °C for 30 min for the extraction of membrane proteins and then centrifuged at 2500g for 10 min. Supernatants were analyzed by BCA Kit. Total sugars concentration was measured by spectrophotometric method of Anthrone modified with respect to that described by Chen and Vaidyanathan [15]. In particular, the disruption of the biomass and of the pellet was carried out according to the procedure described for proteins. Starch was measured using Total Starch kit by Megazyme (Wicklow, Ireland) and according the manufacture's protocol with some modifications. The mass of sample was set at 10 mg and the cells disruption was carried out in the bead beater, as previously reported [16]. Pigments were directly measured by spectrophotometric measurements at 663, 646 and 470 nm. chlorophyll *a*, b and carotenoids concentration were calculated according to Wellburn [17].

Analyses were performed in triplicates. Simple sugars are calculated as difference between total sugars concentration and starch concentration. The recovery yields of the biomolecules were calculated according to mass balances. The recovery yield (*Y*, %) was defined as reported in the following formula:

$$Y = \frac{m_{x \, EP}}{m_{x \, initial}} \cdot 100$$

where $m_{x \, EP}$ is the mass (mg) of the compound *x* (proteins, lipids, total sugars, starch) in EP and $m_{x \, initial}$ is the mass (mg) of the compound *x* in the initial biomass.

2.4. Cell culture and cell survival assay

Human epithelial colorectal adenocarcinoma cells (LoVo) and human immortalized keratinocytes (HaCaT) were obtained from ATCC and cultured in Dulbecco's Modified Eagle's Medium (Sigma-Aldrich), supplemented with 10% foetal bovine serum (HyClone), 2 mM L-glutamine and antibiotics, all from Sigma-Aldrich, in a 5% CO_2 humidified atmosphere at 37 °C. Every 3–4 days the culture medium was removed and cells were rinsed with PBS, detached with trypsinEDTA and diluted in fresh complete growth medium for sub-culturing.

For dose-dependent survival assays cells were seeded in 96-well plates (100 μL/well) at a density of $4 \times 10^4 \, cells \, per \, cm^2$. 24 h after seeding, increasing amount of microalgae extract (from $20 \, \mu g \, mL^{-1}$ to $2 \, mg \, mL^{-1}$) were added to the cells for 24–48 h. At the end of treatment, cells were incubated with 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) reagent and the cells were incubated for 4 h at 37 °C in a humidified 5% CO_2 incubator. At the end of the incubation, medium was removed and the converted dye was solubilized with isopropanol containing $0.01 \, mol \, L^{-1}$ HCl (100 μL per well). Absorbance was measured at a wavelength of 570 nm using an automatic plate reader (Microbeta Wallac 1420, PerkinElmer, Waltham, MA, USA). Cell survival was expressed as percentage of viable cells in the presence of the microalgae extract under test compared with control cells grown in the absence of the extract. The assay was carried out in triplicates for at least 3 times. Control experiments were performed either by growing cells in the absence of the extract or by adding to the cell cultures identical volumes of DMSO. The method used avoids any possibility of a DMSO effect on the final results.

2.5. Oxidative stress

To analyze if EE was able to protect cells from oxidative stress, LoVo cells were plated at a density of $4 \times 10^4 \, cells \, per \, cm^2$. 24 h after seeding, cells were incubated for different length of time (24 or 48 h) in the presence or absence of EE ($200 \, \mu g \, mL^{-1}$), and then incubated in the presence of 300 μM SA for 45 min at 37 °C. At the end of treatment, cell viability was assessed by the MTT assay as reported above.

To detect the minimal pre-treatment time, cells were plated as reported above and then incubated for different length of time

(5–120 min) with EE (200 μg mL^{-1}). Then, cells were exposed to oxidative stress as reported above and a DTNB assay was performed (described more in detail in paragraph 2.6). The result of the experiment is documented in Fig. S1.

Once defined the minimum pre-treatment, in all further experiments cells were pre-treated with EE for 30 min and then exposed to SA for 45 min. After incubation, different assays were performed, as reported below.

2.6. Measurement of intracellular total GSH levels

Total reduced glutathione levels were estimated by using the procedure previously described [18]. Briefly, at the end of incubation, cells were detached by trypsin, lysed and protein concentration was determined by the Bradford assay. Then, 50 μg of proteins were incubated with 3 mM EDTA, 144 μM 5,5′-dithiobis-2-nitrobenzoic acid (DTNB) in 30 mM TrisHCl pH 8.2, centrifuged at 14,000g for 5 min at 4 °C and the absorbance of the supernatant was measured at 412 nm by using a multiplate reader (Biorad). GSH levels were expressed as the percentage of TNB absorbance in the sample under test with respect to the untreated sample. Values are the mean of three independent experiments, each with triplicate determinations.

2.7. Measurement of intracellular ROS levels in cells

ROS intracellular levels were evaluated by using a cells permeable probe, 2′,7′-dichlorodihydrofluorescein diacetate (H$_2$-DCFDA, Sigma-Aldrich). In particular the protocol described in [19] was followed. Briefly, at the end of incubation cells were incubated with the non-fluorescent H$_2$DCFDA becomes fluorescent product, 2′,7′-dichlorofluorescein (DCF), in the presence of different species of ROS. Fluorescence intensity was measured by a Perkin-Elmer LS50 spectrofluorimeter (525 nm emission wavelength, 488 nm excitation wavelength, 300 nm/min scanning speed, 5 slit width for both excitation and emission). ROS production was expressed as percentage of DCF fluorescence intensity of the sample under test, with respect to the untreated sample. Each value was assessed by three independent experiments, each with three determinations.

2.8. Western blot analyses

LoVo cells were plated at a density of 4×10^4 cellspercm^2 in complete medium for 24 h and then treated as described above (paragraph 4.5). After treatment time, cells were detached and lysed in 100 mM Tris-HCl, 300 mM sucrose with addiction of inhibitors of proteases and phosphatases. These lysates were analyzed by Western blotting as reported by Del Giudice et al. [20]. Phosphorylation levels of p38 and MAPKAPK-2 were detected by using specific antibodies purchased from Cell Signal Technology (Danvers, MA, USA). To normalize protein intensity levels, a specific antibody against anti-GAPDH (ThermoFisher, Rockford, IL, USA) was used. The chemiluminescence detection system (SuperSignal® West Pico) was from Thermo Fisher.

2.9. Caenorhabditis elegans strains and culture conditions

Nematode strains were obtained from the Caenorhabditis Genetic Center (CGC). All C. elegans strains were cultured on solid nematode growth media (NGM) plates fed with living E. coli OP50 and maintained at 20 °C in a temperature-controlled incubator. Prior to each assay eggs were isolated to obtain age synchronized cultures and kept in M9 buffer (3 g KH$_2$PO$_4$, 6 g Na$_2$HPO$_4$, 5 g NaCl, 1 mL 1 M MgSO$_4$, H$_2$O to 1 L) for hatching, as described by Stiernagle [21].

2.10. Brood size

Synchronized N2 worms at L4 stage were sorted and placed one by one on individual NGM agar plates. In the treatment group, the bacterial lawn was supplemented with microalgae extract (200 μg mL^{-1}). 100 μg mL^{-1} of epigallocatechin gallate (EGCG, Sigma-Aldrich GmbH, Steinheim, Germany) was used as positive control, whereas DMSO was used to verify that buffer was not responsible for side effects. Adult worms were transferred daily to fresh medium to separate them from their progeny. Eggs were counted every day for 5 days using a dissecting microscope. The results are presented as mean brood size.

2.11. Caenorhabditis elegans survival assay under oxidative stress

24 h after bleaching, N2 wild-type worms and transgenic worms (CF1038, GR1307), in L1 larval stage, were pre-treated with 200 μg mL^{-1} EE and maintained 48 h at 20 °C. EGCG and DMSO were used as indicated above. In the adult stage, worms were transferred to fresh media and grouped in populations of approximately 80 individuals and treated with 80 μM juglone (5-hydroxy-1,4-naphthalenedione) [22]. After 24 h, dead and live worms were counted. The worms were scored dead when they did not respond to a gentle touch stimulus. The results are expressed as percentage of living worms.

2.12. Intracellular ROS accumulation in Caenorhabditis elegans

Synchronized N2 worms at L1 larval stage were pre-treated with 200 μg mL^{-1} EE and maintained 48 h at 20 °C. EGCG and DMSO were used as indicated above. At the end of the treatment, worms were washed with M9 buffer and then incubated with 50 μM H$_2$DCFDA solution for 1 h at 20 °C. Worms were then mounted on a glass slide with a drop of 10 mM sodium azide (used to paralyze them) and photographed by the BIOREVO BZ-9000 fluorescent microscope (Keyence Deutschland 200 GmbH, Neu-Isenburg, Germany) equipped with a mercury lamp. Images were taken for at least 30 worms at constant exposure time (λ Ex 480/20 nm; λ Em 510/38 nm) using a 10 X objective lens. The relative fluorescence of the whole body was determined densitometrically by using ImageJ software 6.01. The results are expressed as percentage of fluorescence intensity of the worms under test compared to untreated worms.

2.13. Quantification of hsp-16.2::GFP expression

The expression of hsp-16.2p::GFP construct was quantified in the strain TJ375. Worms at larval stage L1 were pre-treated with 200 μg mL^{-1} EE and maintained 72 h at 20 °C. EGCG (50 μg mL^{-1}) and DMSO were used as indicated above. Then, worms were exposed to 20 μM juglone for 24 h and analyzed as described above.

2.14. Quantification of sod-3::GFP expression

The worms of transgenic strain CF1553 at larval stage L1 were pretreated with 200 μg mL^{-1} EE and maintained 72 h at 20 °C. EGCG (100 μg mL^{-1}) was used as positive control, whereas DMSO was used as negative control. After treatment, the worms were analyzed as described above.

2.15. Quantification of daf-16::GFP localization

The nematode worms of transgenic strain TJ356 at larval stage L1 were pre-treated with EE (200 μg mL^{-1}) and maintained 24 h at 20 °C and analyzed as described above. EGCG and DMSO were used as indicated above.

2.16. Reverse-phase HPLC analyses

Metabolites from ethanol extracts were separated by reverse-phase HPLC on a Phenomenex Jupiter C18 column (250 × 2.00 mm 5 μm, 300 Å pore size) (Phenomenex, Torrance, California, USA) at a flow rate

of 200 μL min^{-1}. A linear gradient from 5% to 95% acetonitrile in 0.1% formic acid (Sigma Aldrich, Milan, Italy) was used over a time of 40 min and a wavelength of 278 nm was used for monitoring the elution. The eluate was collected in thirteen fractions to be directly analyzed by MALDI-TOF and LC-MS.

2.17. Mass spectrometry analyses

Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) experiments were performed on a 5800 MALDI-TOF-TOF ABSciex equipped with a nitrogen laser (337 nm) (AB SCIEX, Milan, Italy). Aliquots of HPLC fractions (0.5 μL) were mixed (1:1, v/v) with a solution of 2.5 dihydroxybenzoic acid (DHB, Sigma Aldrich, Milan, Italy) at a concentration of 10 mg mL^{-1} in acetonitrile:water (90:10) solution. Standards Kit for Calibration of AB SCIEX MALDI-TOF was from AB SCIEX calibration mixture. For metabolite identification, MS and tandem mass (MSMS) spectra were acquired in reflector positive mode, by using a mass (m/z) range of 100–4000 Da. Laser power was set to 3500 V for MS spectra acquisition. Each spectrum represents the sum of 3000 laser pulses from randomly chosen spots per sample position. For CID experiments, ambient air was used as collision gas with medium pressure of 10^{-6} Torr. The data were reported as monoisotopic masses. LCMS analyses were performed by LC-MS equipped by an Agilent HPLC system (1260 Series) on a reverse-phase C18 column (Agilent Life Sciences Extend-C18, 2.1 × 50 mm, 1.8 μm) coupled to an Agilent 6230 TOF mass spectrometer. The HPLC separation was carried out by using water and acetonitrile as mobile phases A and B, respectively, both acidified with 0.1% formic acid. A linear gradient was employed over 40 min (0–2 min: 2% B, 2–8 min: 2–40% B, 8–20 min: 40%–60% B, 20–30 min: 60–80%, 30–40 min, 80–95) at a flow rate of 0.3 mL min^{-1}. The injection volume was 20 μL and the MS source was an electrospray ionization (ESI) interface in the positive ion mode with capillary voltage of 3000 V, gas temperature at 325 °C, dry gas (N$_2$) flow at 5 L min^{-1} and the nebulizer pressure at 35 psi. The MS spectra were acquired in a mass range of 150–1000 m/z with a rate of 1 spectrum/s, time of 1000 ms/spectrum and transient/spectrum of 9961.

2.18. Statistical analyses

In all the experiments, samples were analyzed in triplicate. The results are presented as mean of results obtained after three independent experiments (mean ± SD) and compared by one-way ANOVA according to the Bonferroni's method (posthoc) using Graphpad Prism for Windows, version 6.01.

3. Results

Recently, Gifuni and colleagues reported a purification protocol addressed to primarily recover starch from *C. sorokiniana* [6]. Authors highlighted the need of removing pigments form the microalgal biomass before recovering starch. It must be underlined that in a biorefinery perspective each component of the microalgae should be valorized. The extraction of pigments from the biomass was performed not to obtain the highest amount of these molecules, but with the aim of extracting interesting antioxidant compounds without damaging, or drastically reducing, the starch recovery yield. For the same reason, ethanol, which does not damage starch, was used as a solvent. Moreover, ethanol belongs to the permitted solvents exploitable for the extraction of compounds used in food industries (law 2009/32/CE), which would allow the use of the extracted antioxidants, or the molecules recovered from the pellet, also as food additives. Finally, ethanol has a low evaporation temperature (78 °C) and it may be easily recovered by evaporation at low specific energetic request.

Table 1
Total composition of biomass, of the extracted biomass pellet (EP) and of ethanol extracted fraction (EE).

	Biomass composition (%DW)	EP composition (% DW)	EE composition (mg)
Simple sugars	10.0 ± 0.3	13.7 ± 0.4	4.9 ± 0.8
Starch	39.2 ± 1.0	26.4 ± 0.6	–
Proteins	29.1 ± 1.1	18.5 ± 0.1	50.2 ± 1.3
Lipids	18.8 ± 0.4	4.1 ± 0.1	65.8 ± 3.6
Weight loss	0	35	–
Chl a	–	0	0.79 ± 0.23
Chl b	–	0	0.66 ± 0.11
Carotenoids	–	0	0.17 ± 0.09

3.1. Pigment extraction

The microalgal biomass was produced in photobioreactors [12], harvested at the onset of nitrogen depletion, and dried. Pigment extraction was performed by using ethanol from the dried biomass, as reported in Materials and Methods section. The composition in starch, simple sugars, proteins and lipids was assessed for the raw biomass and for the extracted biomass pellet (herein denoted as EP). Results are reported in Table 1. The microalgal biomass contained starch (39.2 ± 1.0%), proteins (29.1 ± 1.1%), lipids (19.8 ± 0.4%) and simple sugars (10.0 ± 0.3%). After pigments extraction, the weight loss of the EP was about 35%, with respect to the initial biomass. The EP composition revealed a starch content of 26.4 ± 0.6%, 18.5 ± 0.1% of proteins, 13.7 ± 0.4% of simple sugars and 4.1 ± 0.1% of lipids, thus indicating that lipids were almost completely extracted by the solvent. With this established extraction protocol, it was possible to recover 82.9% of the initial starch, and a significant amount of proteins (78.3%). Only 17.0% of the initial lipid content remained in the pellet after the extraction. Therefore, starch recovery was considered satisfying. Simple sugars seemed to increase in the pellet with respect to the untreated biomass, probably because part of the initial starch is reduced in simple sugars during the treatment. The composition in proteins and lipids of EP suggests applications as feedstock for bioplastics or as feed or food complements without further treatments. By introducing further separation steps, all the components of the pellet could be separated and exploited as biochemicals in different industrial fields with higher profit.

When the ethanol extracted fraction (herein denoted as EE) was analyzed, besides the presence of lipids (29.5 ± 3.6 mg), proteins (12.5 ± 1.3 mg) and sugars (4.9 ± 0.8 mg), carotenoids, chlorophyll *a* and *b* were detected (0.17 ± 0.23 mg, 0.79 ± 0.11 mg and 0.66 ± 0.09 mg, respectively). Thus, the EE was selected to test a putative antioxidant activity on human cells against oxidative stress injury.

3.2. In vitro biocompatibility and antioxidant activity of the ethanol extract from microalgae

Human cancer colon cells (LoVo) were chosen to study the protective effects of EE as a proof of concept. For this reason, we first evaluated the biocompatibility of EE by treating cells for 24 and 48 h with increasing amounts (from 20 μg mL^{-1} to 2 mg mL^{-1}) of the extract and the results are reported in Fig. 1A. In particular, no toxicity was observed up to 200 μg mL^1 of microalgal extract at any time point analyzed, whereas, at higher concentrations (1 and 2 mg mL^{-1}), cell viability was found to be significantly reduced in a dose- and time-dependent manner. These results are in line with literature data, as it has been reported that low doses of antioxidants may have a protective effect on cancer cells, whereas toxicity is observed at high concentration [23, 24].

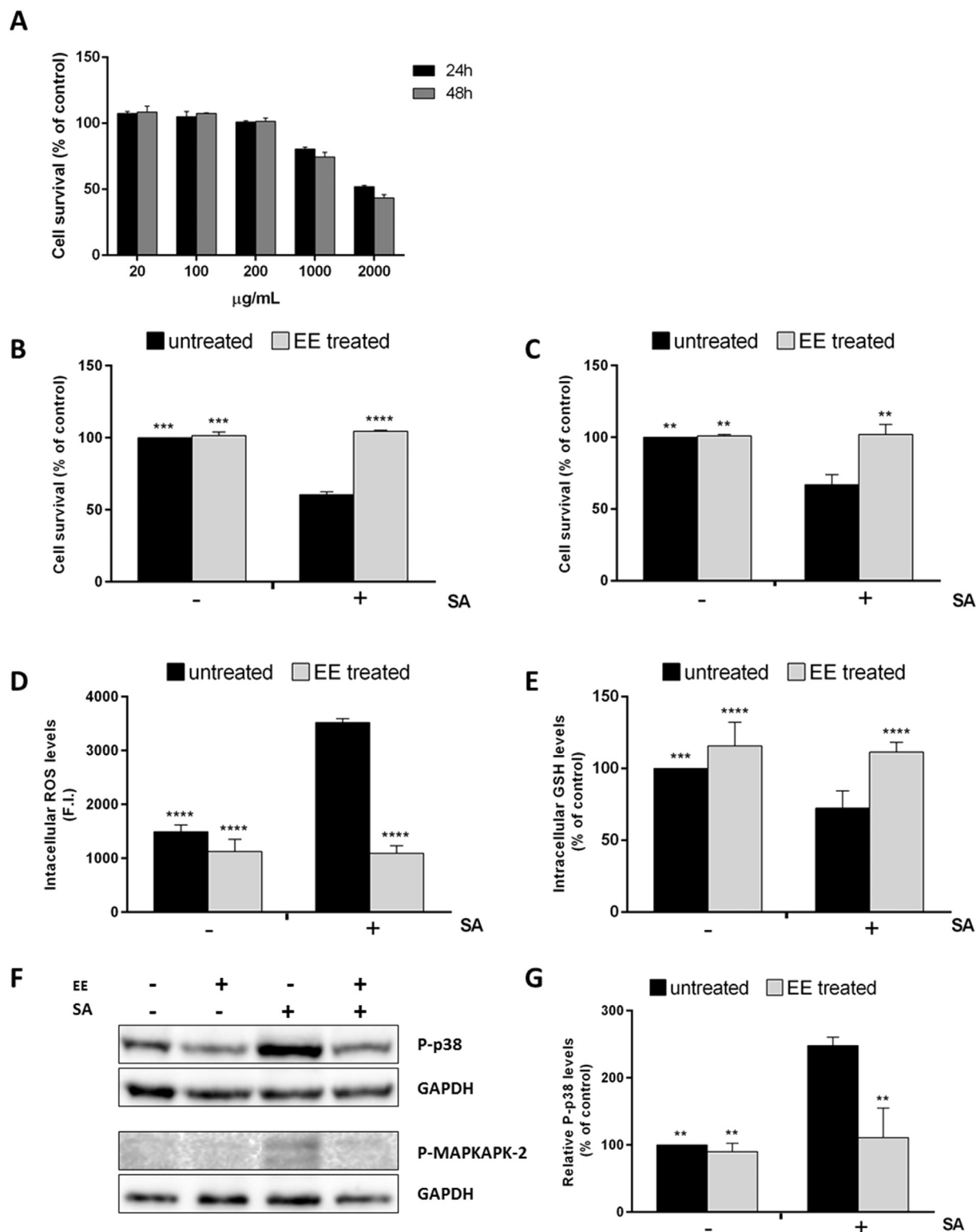Based on these results, we selected 200 μg mL^{-1} microalgal extract

Fig. 1. *In vitro* antioxidant activity of microalgal ethanol extract.
A, LoVo cells were treated with increasing concentrations of EE (from 20 to 2000 μg mL^{-1}) for 24 (black bars) and 48 h (grey bars). Cell viability was assessed by the MTT assay. B, C, LoVo cells were pre-incubated with 200 μg mL^{-1} EE (grey bars) for 24 (B) and 48 h (C) and then treated with 300 μM SA for 45 min. Cell viability was assessed by the MTT assay. D, E, LoVo cells were incubated with 200 μg mL^{-1} EE prior to be stressed by 300 μM sodium arsenite (SA) for 45 min at 37 °C. D, Intracellular ROS levels were determined by DCFDA assay. Values are expressed as fold increase with respect to control (*i.e.* untreated) cells. E, intracellular GSH levels determined by DTNB assay. F, Representative images of Western blots of cells incubated with EE. Cells were treated as described above and, after SA-treatment, incubated for 90 min at 37 °C. In Western blots the phosphorylation level of P-p38 and P-MAPKAPK-2 is reported. GAPDH was used as internal standard. In B, C, D, E, G histograms, black bars refer to control cells, grey bars refer to cells pre-treated with EE, untreated (−) or treated with SA (+).
Values are expressed as fold increase with respect to control (*i.e.* untreated) cells. Data shown are the means ± S.D. of three independent experiments. ** indicates $p < 0.005$; *** indicates $p < 0.001$; **** indicates $p < 0.0001$.
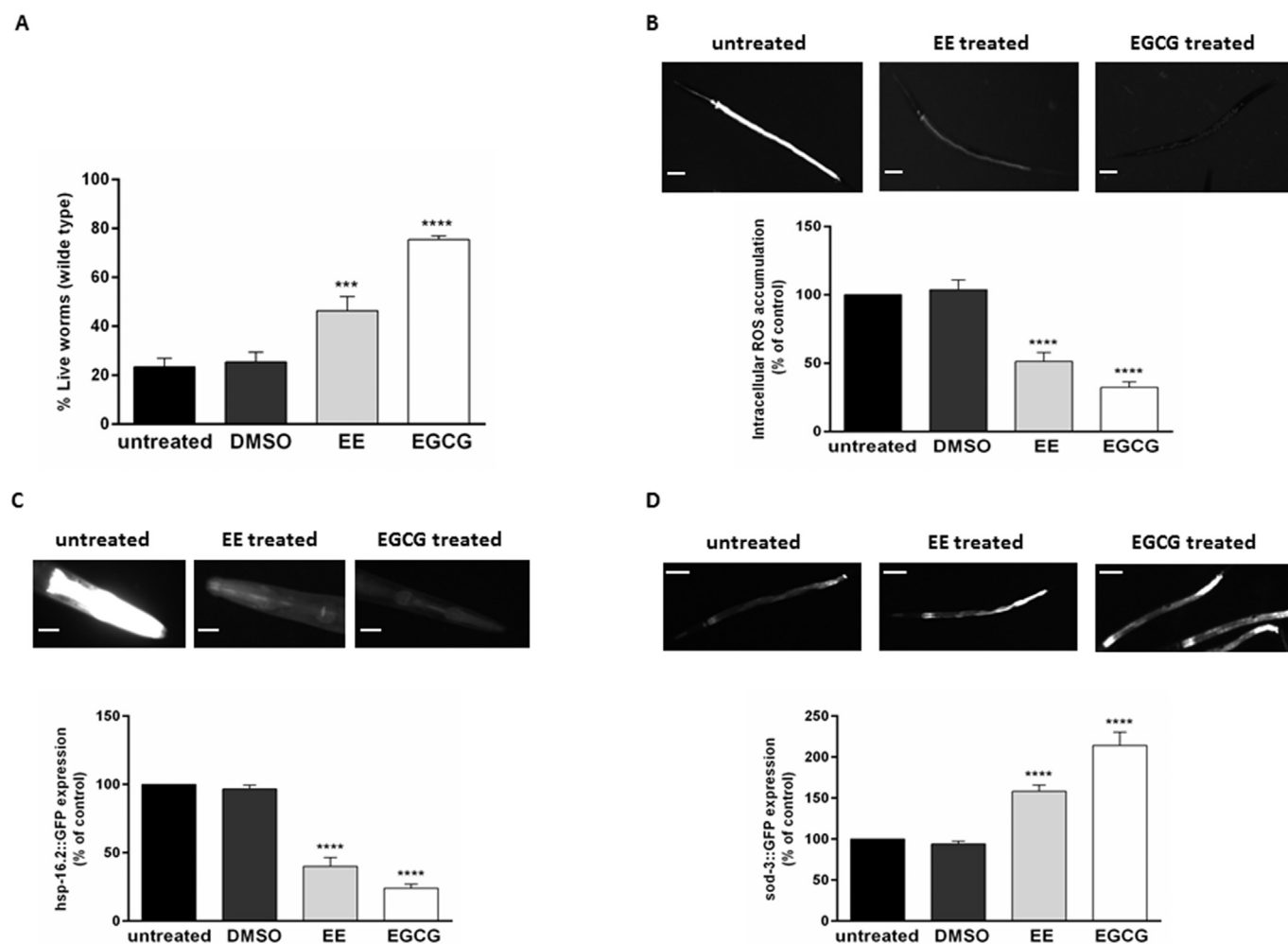
Fig. 2. *In vivo* antioxidant activity of microalgal ethanol extract.
A, wild type N2 worms were treated with $200\,\mu g\,mL^{-1}$ EE extract for 48 h and then exposed to lethal dose of juglone ($80\,\mu M$). After 24 h, alive and dead worms were counted and survival rate were reported in histograms. B, wild type N2 worms were treated with $200\,\mu g\,mL^{-1}$ EE extract for 48 h and then incubated with DCF fluorescent probe. After 1 h incubation, accumulation of intracellular ROS levels was evaluated. On the top, representative images of DCF fluorescence in untreated worms and in worms treated with EE or EGCG; on the bottom, quantification of the fluorescence intensity of the whole body determined densitometrically by using ImageJ software. C, transgenic TJ375 worms were treated with $200\,\mu g\,mL^{-1}$ EE extract for 72 h and then exposed to juglone ($20\,\mu M$). After 24 h, the expression of GFP::*hsp-16.2* was evaluated. On the top, representative images of GFP fluorescence in untreated worms and in worms treated with extract or EGCG; On the bottom, the relative fluorescence of the heads determined densitometrically by using ImageJ software. D, transgenic CF1553 worms were treated with $200\,\mu g\,mL^{-1}$ EE extract for 72 h and then the expression of GFP::*sod-3* was evaluated. On the top, representative images of GFP fluorescence in untreated worms and in worms treated with EE or EGCG; on the bottom, the relative fluorescence of the tails determined densitometrically by using ImageJ software. In the Figure, black bars refer to control worms, light grey bars refer to worms treated with EE, white bars refer to worms treated with EGCG (used as positive control), whereas dark grey bars refer to worms treated with buffer (DMSO). In B, C and D histograms data are presented as percentage of the mean pixel intensity of each treatment, with respect to control worms (mean $\pm$ SD, $n = 40$, replicated at least 3 times). Images were taken with BZ9000 from Keyence, scale bar $= 100\,\mu m$. *** indicates $p < 0.001$, **** indicates $p < 0.0001$, compared to the untreated worms by one-way ANOVA followed by Bonferroni (post-hoc).

as the optimal concentration to analyze the ability of the extract to protect cells from oxidative stress induced by sodium arsenite (SA). Cells were pre-treated with $200\,\mu g\,mL^{-1}$ EE for 24 h (Fig. 1B, grey bars) or 48 h (Fig. 1C, grey bars) in the presence or absence of $300\,\mu M$ SA (45 min incubation) and then cell viability was measured. Cells incubated in the presence of SA, showed a low survival rate (67 and 61% after 24 and 48 h, respectively) when compared to untreated cells (Fig. 1B and C, black bars). Interestingly, the pre-treatment of cells with EE before inducing oxidative stress was able to maintain unaltered cell viability, comparable to that of untreated cells (Fig. 1B and C, grey bars).

As it is well documented that SA induces high ROS production and consequently GSH oxidation in eukaryotic cells [25, 26], we analyzed EE ability to maintain unaltered intracellular ROS and GSH levels. As shown in Fig. 1D, the pre-treatment of LoVo cells with the extract

reduced ROS levels of about 3.5 times with respect to cells exposed to SA. Similarly, total intracellular GSH levels were significantly higher (about 40%) in cells pre-treated with microalgal extract, compared to cells exposed to SA (Fig. 1E). The protective effects of EE against oxidative stress were further confirmed by using human immortalized keratinocytes (HaCaT cells) stressed by SA. Indeed, SA is known to induce oxidative stress also on these cells [27]. By using the same experimental procedure used for LoVo cells, we found that ROS and GSH levels were kept at physiological levels, as reported in Fig. S2.

Finally, we analyzed the phosphorylation levels of p38 and its direct target MAPKAPK-2 (Fig. 1F–G). These proteins belong to the family of mitogen-activated protein kinases (MAPK) and are directly involved in oxidative signaling stress pathways [18]. Consistently with the results reported above, we observed a significant increase in the phosphorylation level of both proteins after oxidative stress (Fig. 1F, third lane).
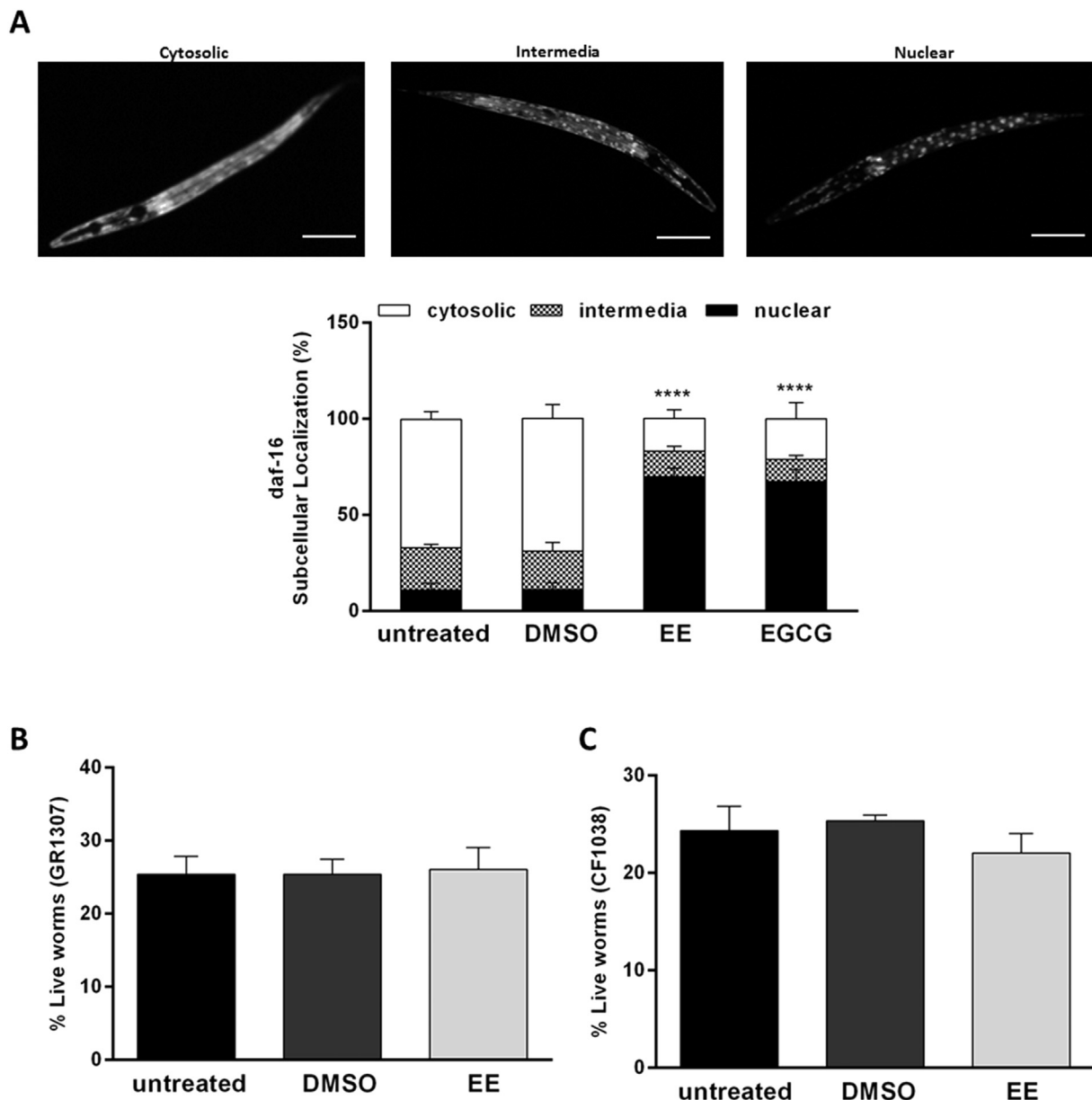
Fig. 3. The DAF-16 pathway activated by EE in *C. elegans*.
A, transgenic TJ356 worms were treated with 200 µg mL^{-1} EE extract for 24 h and then GFP::*daf-16* localization was evaluated. On the top, representative images of different GFP localization. Below, histograms reporting the percentage of worms exhibiting a DAF-16 sub-cellular localization pattern, namely cytosolic (white bars), intermediate (checked bars) and nuclear (black bars). Data are presented as mean ± SD (*n* = 40, replicated at least 3 times). Images were taken with BZ9000 from Keyence, scale bar = 100 µm. **** indicates *p* < 0.0001, compared to the untreated worms by one-way ANOVA followed by Bonferroni (post-hoc). B–C, a *daf-16* null mutant (GR1307) and *daf-16* loss-of-function mutant (CF1038) worms were treated with 200 µg mL^{-1} EE extract (light grey bars) for 48 h and then exposed to lethal dose of juglone (80 µM). After 24 h, alive and dead worms were counted and the survival rate was reported. Data are presented as mean ± SD (*n* = 40, replicated 3 times).

On the other hand, the pre-treatment of cells with the microalgal extract prior to SA resulted in the inhibition of the phosphorylation of the above mentioned markers (Fig. 1F, fourth lanes). It is worth to notice that, in all the experiments described, the microalgal extract alone did not induce oxidative stress.

3.3. In vivo biocompatibility and antioxidant activity of the ethanol extract from microalgae

The protective activity of EE against oxidative stress was further evaluated in an *in vivo* system, *i.e.* Caenorhabditis elegans. We first analyzed the variation of brood size, widely accepted as a toxicity marker [28], after treatment with the microalgal extract, and we did

not find any significant difference with respect to untreated worms (169 ± 7 *vs* 177 ± 6 offspring, respectively).

Then, the antioxidant activity of EE was evaluated by treating N2 wild-type worms, in the L1 larval stage, with EE (200 µg mL^{-1}), for 48 h at 20 °C. The antioxidant epigallocatechin-3-gallate (EGCG), from green tea (100 µg mL^{-1}) was used as a positive control [22], whereas DMSO was used to verify that buffer was not responsible for side effects.

As showed in Fig. 2A, worms pre-treated with EE showed a significantly higher survival rate (46 ± 7 worms, light grey bar) with respect to worms treated with juglone (23 ± 4 worms, black bar) or juglone and DMSO (25 ± 4 worms, dark grey bar).

In the next step, the ability of EE to counteract ROS production was evaluated. As reported in the histograms of Fig. 2B, a significant lower
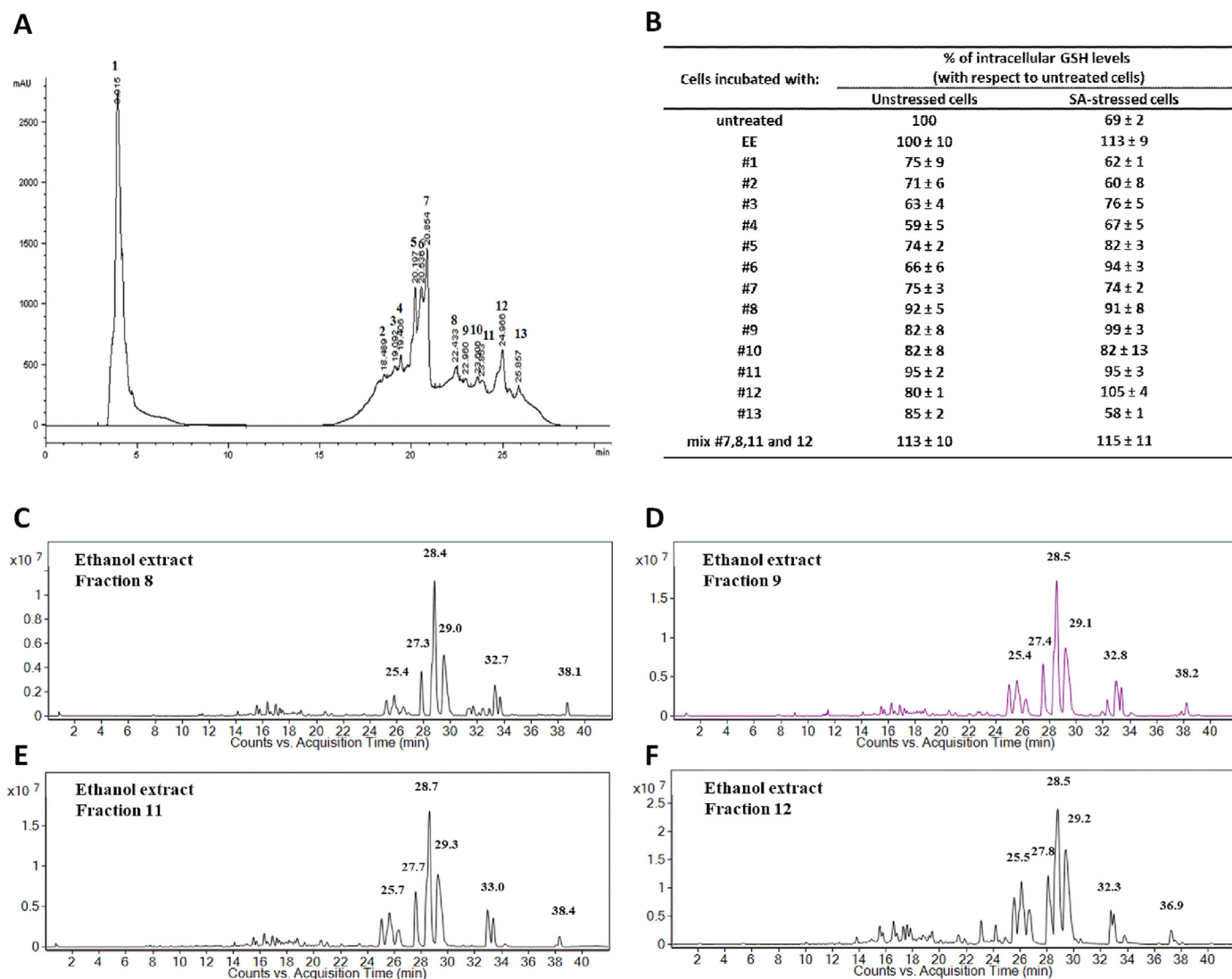
Fig. 4. Characterization of EE.

A, HPLC chromatogram of EE, in which eluted fractions are marked in the profile. B, Analysis of the antioxidant activity of microalgae fractions by DTNB assay. LoVo cells were pre-incubated in the presence of 200 µg mL^{-1} EE or equivalent amount of each fraction after HPLC for 120 min and then cells were stressed by SA treatment (300 µM) for 45 min and intracellular GSH levels determined, as described in Materials and Methods section. Values are reported as % of untreated cells. C–F, TIC chromatogram of the fractions with the highest biological activity.

fluorescence intensity was observed in nematodes (N2 strain) treated with EE (about 50%, light grey bar) with respect to the untreated control group (black and darker grey bars).

Subsequently, the expression of heat shock protein *hsp-16.2* was evaluated. The family of HSPs proteins is found in almost all living organisms and its expression is mainly induced by heat shock or oxidative stress [29]. For this purpose, the mutant strain TJ375, which has a *hsp-16.2* promoter fused with a GFP reporter, was incubated with EE (72 h) in the presence of 20 µM juglone (24 h). Worms pre-treated with the extract showed extremely low expression of *hsp-16.2* (about 60%, Fig. 3C, light grey bar) when compared to stressed worms (Fig. 2C, black and dark grey bars).

As SOD is a known first line defense enzyme in the elimination of ROS, the effect of EE (72 h) on *sod-3* expression was analyzed by using the mutant strain CF1553, which has a *sod-3* promoter fused to a GFP reporter. A significant higher expression of *sod-3*::GFP was detected in worms treated with EE (about 42% increase, Fig. 2D, light grey bar) in comparison with control worms (Fig. 2D, black and dark grey bars).

Therefore, these results indicate that EE extract exerts its antioxidant activity not only by scavenging ROS, but also by modulating

the expression of stress response genes, such as *sod-3* and *hsp-16.2*.

Finally, we analyzed the subcellular localization of DAF-16/FOXO, which is the main transcription factor involved in the regulation of stress response genes. Under physiological conditions, DAF-16/FOXO remains inactive in the cytosol, whereas, some environmental or stress conditions can stimulate its translocation to the nucleus, where it induces the expression of different genes involved in stress response, metabolism and longevity.

The mutant strain TJ356 has a *daf-16* fused to a GFP reporter and was used to detect *daf-16* localization. As shown in Fig. 3A, worms pretreated with EE (24 h) showed a higher percentage of *daf-16*::GFP nuclear localization (about 70%), with respect to untreated worms (12%). This pathway was further confirmed by performing survival assays with a *daf-16* loss-of-function mutant (CF1038) and a *daf-16* null mutant (GR1307) grown in the presence of EE (48 h). In both cases, the extract had no effect on the survival rate of the worms (Fig. 3B and C, light grey bars), in contrast to the results obtained with N2 worms (Fig. 2A).

Table 2

Identification of each peak of TIC chromatogram with relative retention time and relative abundance.

Species	Measured m/z	Detected ion	MALDI-TOF/TOF fragments	Rt (min)	Relative abundance (%)			
					8	9	11	12
Fatty acid C18:3	278.2	M·$^+$		15.4	1.5	3.1	2.4	2.5
Dimeric form of C18:3	555.5	[M + H]$^+$						
Fatty acid C19:2	294.2	M·$^+$		16.2/16.7	3.1	3.7	3.6	2.8
Dimeric form of fatty acid C19:2	587.4	[M + H]$^+$,						
	609.3	[M + Na]$^+$						
Dimeric form of C19:0	595.5	MH$^+$		17.1	1.9	1.7	0.7	1.8
	617.5	[M + Na]$^+$						
Fatty acid C19:1	296.2	M·$^+$		18.6	0.8	2.4	1.4	0.3
Dimeric form of fatty acid C19:1	591.5	[M + H]$^+$,						
		[M + Na]$^+$						
Fatty acid C18:3	279.3	[M + H]$^+$		18.9	0.4	n.d.	n.d.	n.d.
S Porphyrin (SEt)	611.3	[M + H]$^+$	567.4 Da, 547.3 Da, 537.3 Da, 493.3 da, 465.3 Da	19.4	1.7	n.d.	1.4	n.d.
Luteolin dihexoside	611.4	[M + H]$^+$	579.4 Da, 565.4 Da					
Delphinidine 3-O p-coumaroyl-monoglucoside	611.4	[M + H]$^+$	519.3 Da, 465.31, 493.3 (malvidine 3–0 monoglucoside)					
Fatty acid C18:4	277.2	[M + H]$^+$		19.9	0.4	n.d.	1.4	n.d.
Fatty acid C19:2	295.2	[M + H]$^+$						
Fatty acid C22:5	330.2	M·$^+$						
S-Porphyrin (SEt)	611.4	[M + H]$^+$	567.4 Da, 547.3 Da, 537.3 Da, 493.3 da, 465.3 Da					
Luteolin dihexoside	611.4	[M + H]$^+$	579.4 Da, 565.4 Da					
Delphinidine 3-O p-coumaroyl-monoglucoside	611.4	[M + H]$^+$	519.3 Da, 465.31, 493.3 (malvidine 3–0 monoglucoside)					
Fatty acid C19:0	298.2	M·$^+$		20.3	0.5	1.9	1.1	1.3
Dimeric form of fatty acid C19:0	595.5	[M + H]$^+$						
	617.5	[M + Na]$^+$						
Fatty acid C14:0	228.2	M·$^+$		21.9	0.3	0.3	0.3	2.3
Dimeric form of C14:0	455.4	[M + H]$^+$						
Fatty acid C16:1	254.2	M·$^+$		23.4	0.3	0.4	0.6	1.9
		M·$^+$						
Fatty acid C22:4	332.2	M·$^+$						
Dimeric form of C16:1	507.5	[M + H]$^+$						
Fatty acid C18:2	280.2	M·$^+$		24.8	2.3	4.8	4.3	5.2
Dimeric form of C18:2	559.5	[M + H]$^+$		25.4	5.2	8.6	8.7	9.6
		[M + Na]$^+$		26.2	2.4	3.7	4.0	5.2
Fatty acid C16:0	256.3	M·$^+$		27.5	7.7	7.3	8.5	9.7
Dimeric form of C16:0	511.5	[M + H]$^+$						
	533.5	[M + Na]$^+$						
Fatty acid C18:1	282.3	M·$^+$		28.3	32.0	27.6	28.5	23.2
Dimeric form of C18:1	563.5	[M + H]$^+$		29.0	20.3	16.6	20.9	18.6
	585.5	[M + Na]$^+$						
Fatty acid C17:0	270.3	M·$^+$		32.6	n.d.	n.d.	n.d.	0.7
Dimeric form of C17:0	539.5	[M + H]$^+$						
	561.5	[M + Na]$^+$						
Quercetin 3-(2″-p-hydroxybenzoyl-4″-p-coumarylrhamnoside)	715.3	[M + H]$^+$	699.1 Da, 669.1 Da, 561.1 Da, 311.4 Da, 283.4 Da	30.9	0.4	n.d.	n.d.	n.d.
Chlorophyll c2	609.3	[M + H]$^+$	579.4 Da, 549.4 Da, 455.3 Da, 429.05 Da	31.1	5.7	0.5	n.d.	1.3
Pheophorbide a	593.3	[M + H]$^+$	533.3 Da, 459.3 Da	32.8	n.d.	n.d.	n.d.	0.7
				33.8				0.3
Fatty acid C18:0	284.4	M·$^+$		32.7	6.4	5.8	4.8	3.2
Dimeric form of C18:0	567.6	[M + H]$^+$						
	589.5	[M + Na]$^+$						
Fatty acid C20:1	310.3	M·$^+$		33.2	3.7	4.2	n.d.	2.8
Pheophorbide a	593.3	[M + H]$^+$	533.3 Da, 459.3 Da					
Dimeric form of C20:1	619.6	[M + H]$^+$						
	641.6	[M + Na]$^+$						
Pheophytin a′	871.2	[M + H]$^+$	593.2 Da, 714.9 Da					
Hydroxylated chlorophyll a	908.9	[M + H]$^+$	752.9 Da, 311.3 Da, 283.2 Da					
Fatty acid C20:1	310.3	M·$^+$		34.0	0.4	0.6	0.4	0.9
Astaxanthin	596.4	[M + H]$^+$						
Dimeric form of C20:1	619.6	[M + H]$^+$						
	641.6	[M + Na]$^+$						
TG(16:0/18:1(9Z)/18:2	874.7	[M + NH$_4$]$^+$		35.5	n.d.	1.9	n.d.	n.d.
TG(18:1(9Z)/18:1(9Z)/18:1(9Z)	902.7	[M + NH$_4$]$^+$						
Fatty acid C22:1	338.3	M·$^+$		38.0	0.8	1.9	1.6	1.8
Dimeric form of C22:1	675.7	[M + H]$^+$						
	697.5	[M + Na]$^+$						

3.4. Identification of active compounds

EE was then fractionated by HPLC and the elution profile is shown in Fig. 4A. Thirteen fractions were manually collected and tested for antioxidant activity in LoVo cells by measuring intracellular GSH in the presence or absence of oxidative stress (300 μM SA). As shown in Fig. 4B, fractions 8, 9, 11 and 12 had no effect on GSH levels when tested alone, but showed a high protective effect against oxidative stress. To verify the occurrence of additive effects, the active fractions were combined and tested on LoVo cells, in the presence or absence of SA damage, and the results, shown in Fig. 4B, suggested the presence of additive effects.

The active fractions were further analyzed by LC-MS and MALDI-TOF/TOF for the identification of the main agents responsible for their bioactivity. The LC-MS TIC chromatogram of each fraction is reported in Fig. 4C–F. A similar profile for the four fractions was observed, although relevant differences in their relative abundance were recorded. The base peak chromatogram of an ethanol extract with the identification of the most abundant peaks is reported in the Supporting information (Fig. S3) while mass spectral identification of each peak is reported in Table 2. The assignment of species to each mass peak is based on the interpretation of MALDI-TOF/TOF spectra and on the comparison with literature data [30–33]. By LC-TOF analysis, we found several saturated and polyunsaturated chain fatty acids, consisting of their radical species, dimeric forms and sodium adducts (Table 2), as shown as an example in Fig. S4.

Besides FAs, different antioxidant molecules, well documented in the literature, were identified by LC-TOF analysis and confirmed by tandem MS. In particular, carotenoids, such as astaxanthin and its sodium salt (596.4 Da,), and porphyrins, as chlorophyll c2 (609.3 Da), pheophorbide a (593.3 Da) and thioether porphyrin (611.4 Da) were found to be present [10, 34]. A MALDI-TOF/TOF spectrum is reported for chlorophyll C2, with the structures drawn for the most intense fragments (Fig. S5). Interestingly, the chlorophylls were esterified with unsaturated long chain fatty acid, e.g. pheophytin a′ (871.2 Da, Fig. S6) and hydroxylated chlorophyll a (908.9 Da, Fig. S7), were detected only by MALDI.

The other fragmentation spectra were documented in Supporting information (Figs. S8–11). Finally, a derivate of the flavonoid quercetin and some triacylglycerols were found exclusively in fractions 8 and 9, respectively.

4. Discussion and conclusions

A new strategy for a sequential recovery of starch and antioxidants was proposed and validated. We found that, after starch extraction from *C. sorokiniana*, the waste product was able to counteract oxidative stress, both on cell-based models and *in vivo*, as many carotenoids, chlorophylls and FAs are present. Although a variable content of proteins and lipids depending on the specific algal strain and growth conditions was detected in microalgae [35], lipids are always present at high levels [36]. Recently, Wang and Wink reported the beneficial effect of chlorophyll in *C. elegans*, demonstrating that this antioxidant molecule is absorbed by worms, thus becoming bioavailable and promoting longevity [10]. In general, the antioxidant activity of carotenoids is well documented; instead, the literature is still quiescent on the FAs antioxidant activity. However, some studies are reported on the impact of polyunsaturated fatty acids (PUFAs) intake on oxidative stress, as omega-3 PUFAs have been shown to reduce lipoperoxidation levels, advance glycation end products, increase SOD/CAT enzymatic ratio in the liver of diabetic rats fed with a high fat thermolyzed diet (rich in advanced glycation end-products) [37, 38]. These properties, together with the reported inhibition of hepatic lipogenesis afforded by EPA and DHA [39], suggest a multifaceted healthful activity of omega-3 fatty acids supplementation in liver disorders. We assume that chlorophylls and carotenoids, which are known antioxidants, are important

in EE. We still do not know whether additional substances of EE have antioxidant activity, or if there is a synergistic effect among them. Further analysis will be performed to shed light on this issue. It has to be taken into account, however, that, to date, papers claim the multi-product biorefinery of algae, but no cascade or validation of the obtained products are reported [5, 40, 41]. The physico-chemical characterization of the starch granules, extracted with the method here reported, was already carried out [6] and its properties suggest interesting industrial applications: emulsifier for food and pharmaceuticals, bioplastics, textiles and paper preservation. The novelty of the present work concerns the proved feasibility of the cascade approach to isolate two different high value products: starch and antioxidants.

Conflict of interest

All the authors declare no competing interests.

Ethical statement

This article does not contain any studies with human participants or animals performed by any of the authors.

Author contribution

G.P. and I.G. conceived the idea; G.P., D.M.M., M.W. conceived and analyzed the experiments; G.P. performed the cell biology and *in vivo* experiments; M.R. helped in *in vivo* experiments; I.G. performed the extraction experiments with the supervision of G.O.; A.A., A.I., G.P. performed and analyzed the HPLC and mass spectrometry analyses; R.P. and A.M. discussed the results; All authors contributed to write the paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.algal.2018.07.012.

References

[1] K.W. Chew, J.Y. Yap, P.L. Show, N.H. Suan, J.C. Juan, T.C. Ling, D.-J. Lee, J.-S. Chang, Microalgae biorefinery: high value products perspectives, Bioresour. Technol. 229 (2017) 53–62.
[2] C. Monari, S. Righi, S.I. Olsen, Greenhouse gas emissions and energy balance of biodiesel production from microalgae cultivated in photobioreactors in Denmark: a life-cycle modeling, J. Clean. Prod. 112 (2016) 4084–4092, https://doi.org/10.1016/j.jclepro.2015.08.112.
[3] D.U. Santos-Ballardo, S. Rossi, C. Reyes-Moreno, A. Valdez-Ortiz, Microalgae potential as a biogas source: current status, restraints and future trends, Rev. Environ. Sci. Biotechnol. 15 (2016) 243–264, https://doi.org/10.1007/s11157-016-9392-z.
[4] S. Carboni, S.H. Clegg, A.D. Hughes, The use of biorefinery by-products and natural detritus as feed sources for oysters (*Crassostrea gigas*) juveniles, Aquaculture 464 (2016) 392–398, https://doi.org/10.1016/j.aquaculture.2016.07.021.
[5] J. Ruiz, G. Olivieri, J. de Vree, R. Bosma, P. Willems, J.H. Reith, M.H.M. Eppink, D.M.M. Kleingris, R.H. Wijffels, M.J. Barbosa, Towards industrial products from microalgae, Energy Environ. Sci. 9 (2016) 3036–3043, https://doi.org/10.1039/C6EE01493C.
[6] I. Gifuni, G. Olivieri, I.R. Krauss, G. D'Errico, A. Pollio, A. Marzocchella, Microalgae as new sources of starch: isolation and characterization of microalgal starch granules, Chem. Eng. Trans. 57 (2017) 1423–1428, https://doi.org/10.3303/CET1757238.
[7] O. Pulz, W. Gross, Valuable product from biotechnology of microalgae, Appl. Microbiol. Biotechnol. 65 (2004) 635–648, https://doi.org/10.1007/s00253-004-1647-x.
[8] K. Goiris, W. Van Colen, I. Wilches, F. León-Tamariz, L. De Cooman, K. Muylaert, Impact of nutrient stress on antioxidant production in three species of microalgae, Algal Res. 7 (2015) 51–57.
[9] N. Shimidzu, M. Goto, W. Miki, Carotenoids as singlet oxygen quenchers in marine

organisms, Fish. Sci. 62 (1996) 134–137, https://doi.org/10.2331/fishsci.62.134.

[10] E. Wang, M. Wink, Chlorophyll enhances oxidative stress tolerance in *Caenorhabditis elegans* and extends its lifespan, PeerJ 4 (2016) e1879 https://peerj.com/articles/1879/.

[11] G. Olivieri, I. Gargano, R. Andreozzi, R. Marotta, A. Marzocchella, G. Pinto, A. Pollio, Effects of photobioreactors design and operating conditions on *Stichococcus bacillaris* biomass and biodiesel production, Biochem. Eng. J. 74 (2013) 8–14, https://doi.org/10.1016/j.bej.2013.02.006.

[12] I. Gifuni, A. Pollio, G. Olivieri, A. Marzocchella, Identification of an industrial microalgal strain for starch production in biorefinery context: the effect of nitrogen and carbon concentration on starch accumulation, New Biotechnol. 41 (2018) 46–54, https://doi.org/10.1016/j.nbt.2017.12.003.

[13] A.O. Aremu, N.A. Masondo, Z. Molnár, W.A. Stirk, V. Ördög, J. Van Staden, Changes in phytochemical content and pharmacological activities of three *Chlorella* strains grown in different nitrogen conditions, J. Appl. Phycol. 28 (2016) 149–159, https://doi.org/10.1007/s10811-015-0568-7.

[14] G. Breuer, W.A.C. Evers, J.H. de Vree, D.M.M. Kleinegris, D.E. Martens, R.H. Wijffels, P.P. Lamers, Analysis of fatty acid content and composition in microalgae, J. Vis. Exp. (2013), https://doi.org/10.3791/50628.

[15] Y. Chen, S. Vaidyanathan, Simultaneous assay of pigments, carbohydrates, proteins and lipids in microalgae, Anal. Chim. Acta 776 (2013) 31–40, https://doi.org/10.1016/j.aca.2013.03.005.

[16] I. Gifuni, G. Olivieri, A. Pollio, T.T. Franco, A. Marzocchella, Autotrophic starch production by *Chlamydomonas* species, J. Appl. Phycol. 29 (2017) 105–114, https://doi.org/10.1007/s10811-016-0932-2.

[17] A.R. Wellburn, The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution, J. Plant Physiol. 144 (1994) 307–313, https://doi.org/10.1016/S0176-1617(11)81192-2.

[18] G. Petruk, A. Raiola, R. Del Giudice, A. Barone, L. Frusciante, M.M. Rigano, D.M. Monti, An ascorbic acid-enriched tomato genotype to fight UVA-induced oxidative stress in normal human keratinocytes, J. Photochem. Photobiol. B Biol. 163 (2016) 284–289, https://doi.org/10.1016/j.jphotobiol.2016.08.047.

[19] R. Del Giudice, G. Petruk, A. Raiola, A. Barone, D.M. Monti, M.M. Rigano, Carotenoids in fresh and processed tomato (*Solanum lycopersicum*) fruits protect cells from oxidative stress injury, J. Sci. Food Agric. 97 (2017) 1616–1623, https://doi.org/10.1002/jsfa.7910.

[20] R. Del Giudice, A. Raiola, G.C. Tenore, L. Frusciante, A. Barone, D.M. Monti, M.M. Rigano, Antioxidant bioactive compounds in tomato fruits at different ripening stages and their effects on normal and cancer cells, J. Funct. Foods 18 (2015) 83–94, https://doi.org/10.1016/j.jff.2015.06.060.

[21] T. Stiernagle, Maintenance of *C. elegans*, (2006), https://doi.org/10.1895/wormbook.1.101.1.

[22] S. Abbas, M. Wink, Green tea extract induces the resistance of *Caenorhabditis elegans* against oxidative stress, Antioxidants 3 (2014) 129–143, https://doi.org/10.3390/antiox3010129.

[23] M.M. Rigano, A. Raiola, G.C. Tenore, D.M. Monti, R. Del Giudice, L. Frusciante, A. Barone, Quantitative trait loci pyramiding can improve the nutritional potential of tomato (*Solanum lycopersicum*) fruits, J. Agric. Food Chem. 62 (2014) 11519–11527, https://doi.org/10.1021/jf502573n.

[24] M.V. Eberhardt, C.Y. Lee, R.H. Liu, Antioxidant activity of fresh apples, Nature 405 (2000) 903–904, https://doi.org/10.1038/35016151.

[25] M. Akanda, H. Tae, I. Kim, D. Ahn, W. Tian, Hepatoprotective role of hydrangea macrophylla against sodium arsenite-induced mitochondrial-dependent oxidative stress via the inhibition of MAPK/caspase-3 pathways, Int. J. Mol. Sci. 18 (2017)

1–15, https://doi.org/10.3390/ijms18071482.

[26] N. Maheshwari, F. Khan, R. Mahmood, Sodium meta-arsenite induced reactive oxygen species in human red blood cells: impaired antioxidant and membrane redox systems, haemoglobin oxidation, and morphological changes, Free Radic. Res. 51 (2017) 1–49, https://doi.org/10.1080/10715762.2017.1327714.

[27] G. Petruk, G. Donadio, M. Lanzilli, R. Isticato, D.M. Monti, Alternative use of *Bacillus subtilis* spores: protection against environmental oxidative stress in human normal keratinocytes, Sci. Rep. 8 (2018), https://doi.org/10.1038/s41598-018-20153-2.

[28] L. Bischof, D. Huffman, R. Aroian, Assays for toxicity studies in *C. elegans* with Bt crystal proteins, *C. elegans*, PCR Methods Appl. (2006) 139–154.

[29] M.G. Santoro, Heat shock factors and the control of the stress response, Biochem. Pharmacol. 59 (2000) 55–63 (doi:S0006-2952(99)00299-3 [pii]).

[30] C. Juin, A. Bonnet, E. Nicolau, J. Bérard, R. Devillers, UPLC-MSE profiling of phytoplankton metabolites: application to the identification of pigments and structural analysis of metabolites in porphyridium, Mar. Drugs 13 (2015) 2541–2558.

[31] B. Gilbert-López, J. Mendiola, J. Fontecha, Downstream processing of *Isochrysis galbana*: a step towards microalgal biorefinery, Green Chem. 17 (2015) 4599–4609.

[32] I. Lang, L. Hodac, T. Friedl, Fatty acid profiles and their distribution patterns in microalgae: a comprehensive analysis of more than 2000 strains from the SAG culture collection, BMC Plant Biol. 11 (2011) 124, https://doi.org/10.1186/1471-2229-11-124.

[33] R. Flamini, P. Traldi, Mass Spectrometry in Grape and Wine Analysis, (2010).

[34] M.N. Serrano, A. Gustavo, Natural astaxanthin, antioxidant protection power for healthy eyes, Agro. Food Ind. Hi Tech (2014) 25.

[35] V.T. Duong, F. Ahmed, S.R. Thomas-Hall, S. Quigley, E. Nowak, P.M. Schenk, High protein- and high lipid-producing microalgae from northern Australia as potential feedstock for animal feed and biodiesel, Front. Bioeng. Biotechnol. 3 (2015) 53, , https://doi.org/10.3389/fbioe.2015.00053.

[36] K. Sharma, Y. Li, P. Schenk, UV-C-mediated lipid induction and settling, a step change towards economical microalgal biodiesel production, Green Chem. 16 (2014) 3539–3548 http://pubs.rsc.org/-/content/articlehtml/2014/gc/c4gc00552j.

[37] D. Richard, K. Kefi, U. Barbe, P. Bausero, F. Visioli, Polyunsaturated fatty acids as antioxidants, Pharmacol. Res. 57 (2008) 451–455, https://doi.org/10.1016/j.phrs.2008.05.002.

[38] A.M. de Assis, A. Rech, A. Longoni, L.N. Rotta, C.C. Denardin, M.A. Pasquali, D.O. Souza, M.L.S. Perry, J.C. Moreira, Ω3-Polyunsaturated fatty acids prevent lipoperoxidation, modulate antioxidant enzymes, and reduce lipid content but do not alter glycogen metabolism in the livers of diabetic rats fed on a high fat thermolyzed diet, Mol. Cell. Biochem. (2011) 1–10, https://doi.org/10.1007/s11010-011-1099-4.

[39] A. Lamaziere, C. Wolf, U. Barbe, P. Bausero, F. Visioli, Lipidomics of hepatic lipogenesis inhibition by omega 3 fatty acids, Prostaglandins Leukot. Essent. Fat. Acids 88 (2013) 149–154, https://doi.org/10.1016/j.plefa.2012.12.001.

[40] L.M.L. Laurens, J. Markham, D.W. Templeton, E.D. Christensen, S. Van Wychen, E.W. Vadelius, M. Chen-Glasser, T. Dong, R. Davis, P.T. Pienkos, Development of algae biorefinery concepts for biofuels and bioproducts; a perspective on process-compatible products and their impact on cost-reduction, Energy Environ. Sci. 10 (2017) 1716–1738, https://doi.org/10.1039/C7EE01306J.

[41] C. Safi, D.Z. Liu, B.H.J. Yap, G.J.O. Martin, C. Vaca-Garcia, P.-Y. Pontalier, A two-stage ultrafiltration process for separating multiple components of *Tetraselmis suecica* after cell disruption, J. Appl. Phycol. 26 (2014) 2379–2387, https://doi.org/10.1007/s10811-014-0271-0.

FAAH-Catalyzed C−C Bond Cleavage of a New Multitarget Analgesic Drug

Alessia Ligresti,*,[†] Cristoforo Silvestri,[†] Rosa Maria Vitale,[†] Jose L. Martos,[‡] Fabiana Piscitelli,[†] Jenny W. Wang,[§,○] Marco Allarà,[†] Robert W. Carling,[‡] Livio Luongo,[∥] Francesca Guida,[∥] Anna Illiano,[#] Angela Amoresano,[#] Sabatino Maione,[∥] Pietro Amodeo,[†] David F. Woodward,[§] Vincenzo Di Marzo,*,[†,⊥] and Gennaro Marino[#,∇]

[†]Institute of Biomolecular Chemistry, Endocannabinoid Research Group, National Research Council of Italy, Pozzuoli 80078, Italy

[‡]Discovery Department, Selcia Limited, Ongar CM5 0GS, United Kingdom

[§]Department of Biological Sciences, Allergan Inc., Irvine, California 92623, United States

[∥]Department of Experimental Medicine, Pharmacology Division, University of Campania, Naples 80138, Italy

[⊥]Canada Excellence Research Chair on the Microbiome-Endocannabinoidome Axis in Metabolic Health, Universitè Laval, Quebec City G1V 0A6, Canada

[#]Department of Chemical Sciences, University of Naples "Federico II", Naples 80126, Italy

[∇]University "Suor Orsola Benincasa", Naples 80132, Italy

S *Supporting Information*

ABSTRACT: The discovery of extended catalytic versatilities is of great importance in both the chemistry and biotechnology fields. Fatty acid amide hydrolase (FAAH) belongs to the amidase signature superfamily and is a major endocannabinoid inactivating enzyme using an atypical catalytic mechanism involving hydrolysis of amide and occasionally ester bonds. FAAH inhibitors are efficacious in experimental models of neuropathic pain, inflammation, and anxiety, among others. We report a new multitarget drug, AGN220653, containing a carboxyamide-4-oxazole moiety and endowed with efficacious analgesic and anti-inflammatory activities, which are partly due to its capability of achieving inhibition of FAAH, and subsequently increasing the tissue concentrations of the endocannabinoid anandamide. This

inhibitor behaves as a noncompetitive, slowly reversible inhibitor. Autoradiography of purified FAAH incubated with AGN220653, opportunely radiolabeled, indicated covalent binding followed by fragmentation of the molecule. Molecular docking suggested a possible nucleophilic attack by FAAH-Ser241 on the carbonyl group of the carboxyamide-4-oxazole moiety, resulting in the cleavage of the C−C bond between the oxazole and the carboxyamide moieties, instead of either of the two available amide bonds. MRM-MS analyses only detected the Ser241-assisted formation of the carbamate intermediate, thus confirming the cleavage of the aforementioned C−C bond. Quantum mechanics calculations were fully consistent with this mechanism. The study exemplifies how FAAH structural features and mechanism of action may override the binding and reactivity propensities of substrates. This unpredicted mechanism could pave the way to the future development of a completely new class of amidase inhibitors, of potential use against pain, inflammation, and mood disorders.

KEYWORDS: *FAAH mechanism, C−C bond cleavage, multitarget inhibitors*
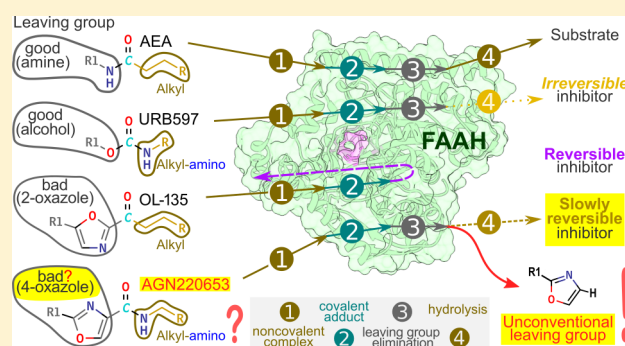
■ INTRODUCTION

Fatty acid amide hydrolase (FAAH) is an atypical serine hydrolase, which embodies a catalytic mechanism for the hydrolysis of amide and ester substrates. As the enzyme responsible for the hydrolysis of the endocannabinoid *N*-arachidonoyl-ethanolamine (anandamide, AEA) and related anti-inflammatory molecules,[1] FAAH has been proposed as a therapeutic target for the treatment of neuropathic pain, inflammation, and mood disorders. Indeed, AEA, via different mechanisms including

peripheral and/or spinal activation of cannabinoid CB1 and CB2 receptors, or activation and/or desensitization of transient receptor potential vanilloid type-1 (TRPV1) channels, exerts antinociceptive and antihyperalgesic actions in various animal models of inflammatory and neuropathic pain.[2] Furthermore,

AEA also inhibits anxiety and depression in experimental models.[3] Accordingly, selective FAAH inhibitors have been developed as potential analgesics, anxiolytics and antidepressants.[4] However, when tested in phase II clinical trials, an irreversible FAAH inhibitor failed to alleviate pain in patients with osteoarthritis.[5] Recently, new strategies focusing on targeted antagonism of selected prostanoid receptors have been developed[6,7] resulting in efficacious compounds acting through the simultaneous blockade of the action of proinflammatory prostanoids at DP_1, DP_2, EP_1, EP_4, FP, and TP receptors, while sparing the more benign EP_2, EP_3, and IP receptors. Based on this background, and on the increasing demand for rationally designed multitarget drugs, we aimed at including the chemical requirements for FAAH inhibition into the aforementioned prostanoid pan-antagonists through their progressive chemical modification. Having achieved this aim, we discovered, using molecular modeling, biochemistry, quantum mechanics, and mass spectrometry approaches, that these new multitarget molecules inhibit FAAH by favoring an unprecedented catalytic behavior. This unique mechanism may now open previously unexplored avenues in the fields of biochemistry and medicinal and physical chemistry, with potential impact on therapeutic drug development.

■ RESULTS AND DISCUSSION

Rational Design of Combined Pan-PG Antagonists Hosting the Capability of Inhibiting FAAH in a Slowly Reversible Manner. We previously patented[8] and reported[7] a novel scaffold (AGN211377) with a unique polypharmacology and therapeutically ideal prostanoid receptor pan-antagonism profile at FP, TP, DP, EP_1, and EP_2 receptors that spares EP_2, EP_3, and IP receptors.[6]

This compound (AGN211377) possesses relatively weak FAAH inhibitory activity (medium μmolar range), which however could be significantly improved by the introduction of amide substituents in the molecule. AGN220653 (IC_{50} = 93.6 nM against rat brain FAAH, Tables S1 and S2) was thus identified as a mixed prostanoid pan-antagonist/FAAH inhibitor, which produced analgesic activity in vivo via the expected molecular targets (Figures 1a−d and S1a-b). Our efforts were next directed to understanding AGN220653 selectivity and mode of action. First, we verified, using an affinity-based proteomic profiling assay, that AGN220653 does not hit several "off-target" serine hydrolases present in a mouse brain proteome (Figure 1e,f). Furthermore, unlike some other previously reported FAAH inhibitors, AGN220653, up to 10 μM, showed little affinity for CB1 and CB2 cannabinoid receptors and did not activate nor antagonize TRPV1 channels (Table S3). Next, we assessed whether or not AGN220653 acts via a slowly reversible mechanism, by comparing it to other well-known FAAH irreversible inhibitors (URB579 and MAFP) incubated with rat brain membranes for up to 24 h. A gradual reduction of AGN220653 potency was found over time with a corresponding recovery of enzyme activity. Instead, as previously reported,[9,10] MAFP retained its activity for up to 24 h, while the effect of URB597 was comparable with that of AGN220653 but showed a reduction of its inhibition potency only at the latest time point measured, suggesting a difference in the type of interaction with the enzyme (Figures S1c−e and S2).

Mechanistic Characterization of FAAH Inhibition by AGN220653. Next, the following questions were addressed: on which of the two carbonyl groups of AGN220653 does the nucleophilic attack by Ser241 (one of the residues of the unusual FAAH catalytic triad) occur? Is the nucleophilic attack followed by elimination of a leaving group? Which of the two expected leaving groups (the acyl-sulfonamine or the cyclohexyl-butyl-amine) is preferentially eliminated? To address these questions, we radiolabeled either of the two putative leaving groups of AGN220653. Furthermore, AGN220346, an analogue with comparable activity at FAAH but presenting an isopropyl sulfonamide group instead of the phenyl sulfonamide moiety, was also utilized. Both compounds were [^{14}C]-radiolabeled at the 1-position of the cyclohexyl-butyl amine portion. Alternatively, the 1-position of the isopropyl group in AGN220346, and the phenyl group in AGN220653, of the sulfonamide were [^{14}C]-radiolabeled. The [^{14}C]-AGN compounds (75 000 cpm; 30 μM) were then incubated with custom highly purified humanized FAAH (0.2 μg; Medicilon Ltd., Shangai, China) at pH 9 in the presence and absence of an excess of the corresponding nonradiolabeled AGN compounds (700 μM).

The denatured protein was electrophoresed through a 10% SDS polyacrylamide gel before exposing the films. Only when labeled on their cyclohexyl-butyl amine portion did the inhibitors also label full length FAAH, with a signal that was displaced by the initial coincubation of the enzyme with the corresponding nonradiolabeled compounds (Figure 2). These results supported a mechanism involving the elimination of the acyl sulfonamide group, with Ser241 attacking the "upper" carbonyl that is adjacent to this group.

However, molecular docking suggested a completely different scenario. In fact, on the basis of the representative poses, i.e., those forming the best ligand-protein polar interactions and endowed with the most favorable binding energies, both AGN220653 and AGN220346 share a similar energetically favored pose within the FAAH binding site (Figure 3a). In this pose, the carbonyl group is close to Ser241 and the group interacting with the oxyanion hole is, instead, the carboamide, that is, the one in the "lower" chain of the compounds.

As shown in Figure 3b, neither the *cis* nor the *trans* conformations of the sulfonamide group affect the overall arrangement of the two ligands, supporting the binding mode prediction. Accordingly, the cyclohexyl-butyl chain and the oxazole moiety were hosted in the acyl-binding pocket (ABP) and in the cytosolic port (CP), respectively, and the sulfonamide group protruded toward the protein surface, thus being too distant from the active site. Intriguingly, the only potential cleavage site on the inhibitor that is compatible with such orientation is, quite unexpectedly, the bond between the carbonyl group and the oxazole moiety. The attack by Ser241 on this carbonyl group and the subsequent, and quite unexpected, cleavage of the C−C bond instead of the "upper" amide bond, would still explain the autoradiography results.

Mass Spectrometric Characterization of the AGN220653 Moiety Covalently Binding to FAAH. To gain experimental evidence in favor of this unexpected mechanism for FAAH reaction with AGN220653, we first verified the formation of the hypothesized adduct between Ser241 and the inhibitor by means of MALDI mass spectrometry analyses. MALDI mapping of purified humanized FAAH-AGN220653 tryptic hydrolysate covered more than 82.4% of the amino acid sequence. This, most interestingly, included a signal occurring at m/z 2851.87, corresponding to the FAAH peptide including the 213−243 amino acids (and containing Ser241), bound with the cyclohexyl-butyl-carbonyl group (Δ = 181.17), as compared to the expected value Δ = 181.27), as shown in Figure 4).
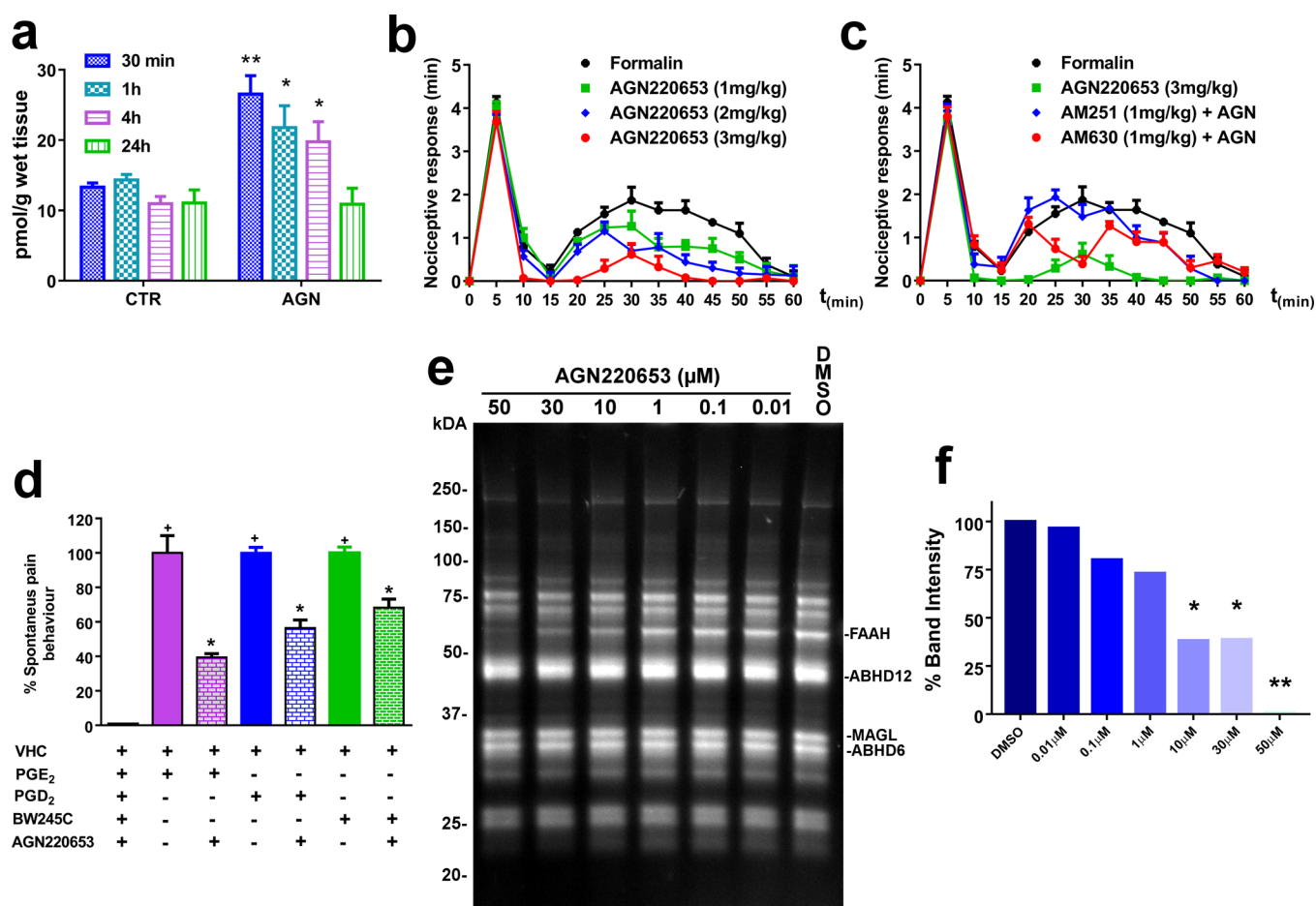
Figure 1. (a) AEA levels (pmol/g of wet tissue) in mouse brains treated with AGN220653 (3 mg/kg, i.p.) for different time points (30 min, 1 h, 4 h, and 24 h). Data are expressed as mean ± SEM ($n = 4$ per group). Student's t test $*p < 0.5$; $**p = 0.0079$ was performed for treated vs untreated samples at each time point. (b,c) Effect of AGN220653 (1, 2, and 3 mg/kg, i.p.) alone or in combination, at the dose of 3 mg/kg, with AM251 (1 mg/kg, i.p.) or AM630 (1 mg/kg, i.p.) in formalin-injected mice. The total time of the nociceptive response was measured every 5 min and expressed as the total time of the nociceptive responses in minutes. Results are mean ± SEM ($n = 6-8$ for each group). (d) Effect on spontaneous pain behavior after subcutaneous injection of PGE$_2$ (10 μg/20 μL/paw) or PGD$_2$ (150 μg/20 μL/paw) or selective DP$_1$ receptor agonist BW245C (200 μg/20 μL/paw) in the presence of vehicle (0.5% DMSO in saline, i.p.) or AGN220653 (3 mg/kg i.p.). The total time of the nociceptive response was measured within 20 min and expressed as percentage of PG-induced effects (100%). Results are mean ± SEM ($n = 6-8$ for each group). Plus signs (+) denote statistically significant differences ($P < 0.05$) vs vehicle, and asterisks (*) denote statistically significant differences ($P < 0.05$) between each group and relative control (one-way ANOVA, Tukey's post hoc). (e) Selectivity profiles of AGN220653 (0.01−0.1−1−10−30−50 μM), as judged by competitive ABPP analysis in the rat brain membrane proteome. (f) Bars represent densitometric analysis of band intensity for AGN220653 effect (0−50 μM) on FAAH activity in the rat brain membrane proteome. Results represent the average values ± SEM of three independent experiments. Student's t test ($*p < 0.05$; $**p < 0.001$) was performed for inhibitor versus vehicle (DMSO) treated samples.

This finding confirmed the cleavage of AGN220653 according to the new mechanism. It is worth noting that a signal at m/z 2670.70, corresponding to the unmodified tryptic peptide, was also identified, whereas the putative adduct resulting from the "classical" cleavage mechanisms, corresponding to the elimination of the acyl sulfonamide group and the labeling of the 213−243 FAAH peptide with the remainder of the molecule (m/z = 3182.75), was undetectable (Figure S3). In order to unequivocally detect the modified Ser241, a multiple reaction monitoring (MRM) tandem mass spectrometry experiment, using an ion spray source operated in positive mode, was set up to monitor the modified 213−243 FAAH peptide. Selected transitions and precursor ion-daughter ions for the 213−243 FAAH-AGN220653 peptide were obtained using the Skyline software. MRM analysis revealed the presence of the adduct between the enzyme and the cyclohexyl-butyl-carboxyamide portion at the level of Ser241 (Figure 5). These findings confirm the unprecedented C−C bond

cleavage by FAAH subsequent to the enzyme nucleophilic attack on the carbonyl group adjacent to the oxazole moiety.

AGN220653 Interaction with FAAH Reveals a Non-amidase Mechanism of Action for the Enzyme. Since the mechanism suggested by the above experiments implies the cleavage of a theoretically very stable C−C bond, further support of this hypothesis was sought from computational techniques.

Because of the varied catalytic behavior and inhibition mechanisms of FAAH and the number of enzyme residues affecting it,[1,4,11] several methodological approximations and simplifications are mandatory to successfully apply computational methods to this complex protein and its complexes. As a consequence, many different approaches have been developed and applied[12] to optimally target specific structural or functional features of the studied systems.

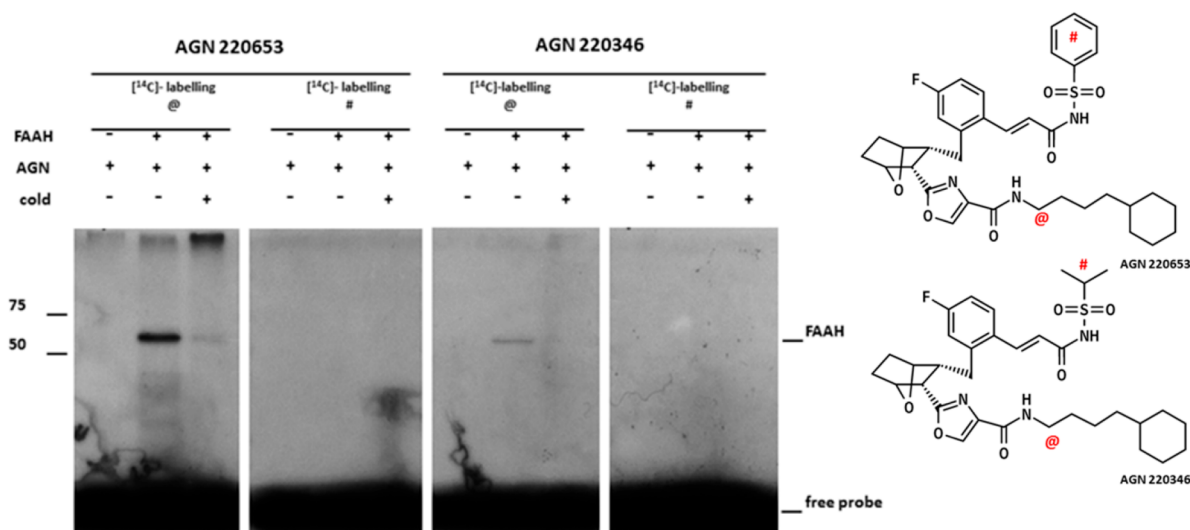In the case of the FAAH-AGN220653 complex, we sought for a Quantum Mechanical (QM) computational approach (I)

Figure 2. SDS-PAGE of purified humanized FAAH (Medicilon, Ltd., Shangai, China) incubated with AGN220653 or AGN220346 showing ^{14}C-radiolabeling of cyclohexyl-butyl amine (@) and sulfonamide (#) groups. Purified FAAH incubated either with or without radiolabeled compound in the absence or presence of nonradiolabeled compound and subjected to SDS-PAGE before exposure to the radiographic film. Protein ladder (left), the position of FAAH, and unbound radiolabeled compound (free probe) are indicated.
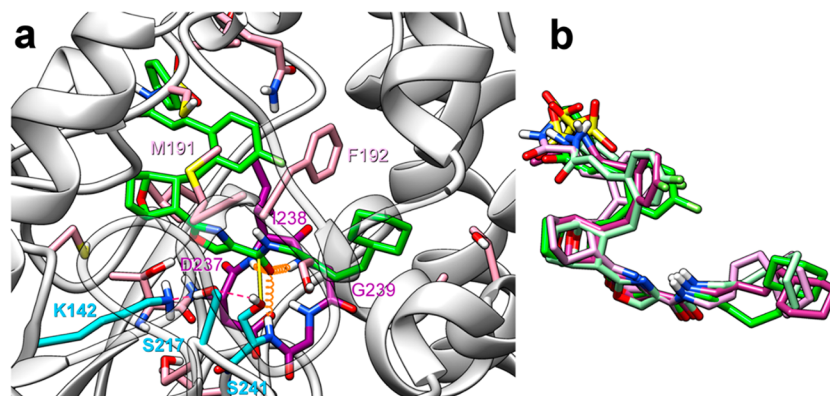

Figure 3. (a) Representative docking model of the FAAH (gray ribbons)-AGN220653$_{cis}$ complex (stick ligand, ligand-contacting (<4 Å) side chain and oxyanion-hole backbone bonds, with green C atoms and other elements in standard colors). C atoms colored cyan/magenta/pink for catalytic-triad/oxanionic-hole/other residues. (b) Ligand superposition from FAAH complexes of AGN220653$_{cis}$ (green), AGN220653$_{trans}$ (light-green), AGN220346$_{cis}$ (pink), and AGN220346$_{trans}$ (dark magenta).

providing a satisfying structural description of a model system comparable to that under investigation; (II) reproducing the functional behavior of the model system; (III) correctly predicting the lack of C−C bond cleavage in a negative control system similar to the complex under investigation; and (IV) providing results substantially independent upon the exact choice of QM methods/setups within a reasonable range of conditions, with the latter being another important criterion to evaluate the results on a new, unexpected mechanism.

The first step in the choice of the methods and the definition of the systems to be simulated is the selection of a system to be used to guide and validate the system setup according to the above-defined criteria (I)-(III). AGN220653 arrangement in FAAH active site, as assessed in docking experiments, closely resembles that found in PDB entry 2WJ1 for the complex between FAAH and α-ketooxazole inhibitor (therein named "S99"), where a slowly reversible covalently bound tetrahedral intermediate (TI) is formed.[13] However, while bearing an oxazole moiety like AGN220653, S99 features a simple carbonyl instead of an amide group, with a different attachment to the

oxazole ring (C2 atom vs C4 in AGN220653). In this view, the FAAH-S99 complex, in addition to guide the choice of the computational methods and opportunely size the system to be studied, was also employed to validate the computational approach and study the structure−activity relationships of these compounds. In particular, since no C−C bond cleavage is experimentally observed for FAAH-S99 complex, we reasoned that this can also represent the proper "negative control" necessary to test possible biases of the selected computational protocol toward C−C bond cleavage (criterion III).

To derive a QM model, we subdivided the ligand in four regions, including atoms: a) contacting FAAH residues substantially contributing to enzyme activity in experimental/computational studies[11] (labeled in bold magenta in Figure 6a); b) forming interactions rigidly or semirigidly anchoring the ligand to the protein in regions adjacent to a); c) contacting the protein in points distant from a) region; d) forming week or no interaction with FAAH. Only ligand atoms in regions a) and b) were retained (Figure 6b), together with the corresponding contacting residues or side chains, shown in Figure 6a. The selected
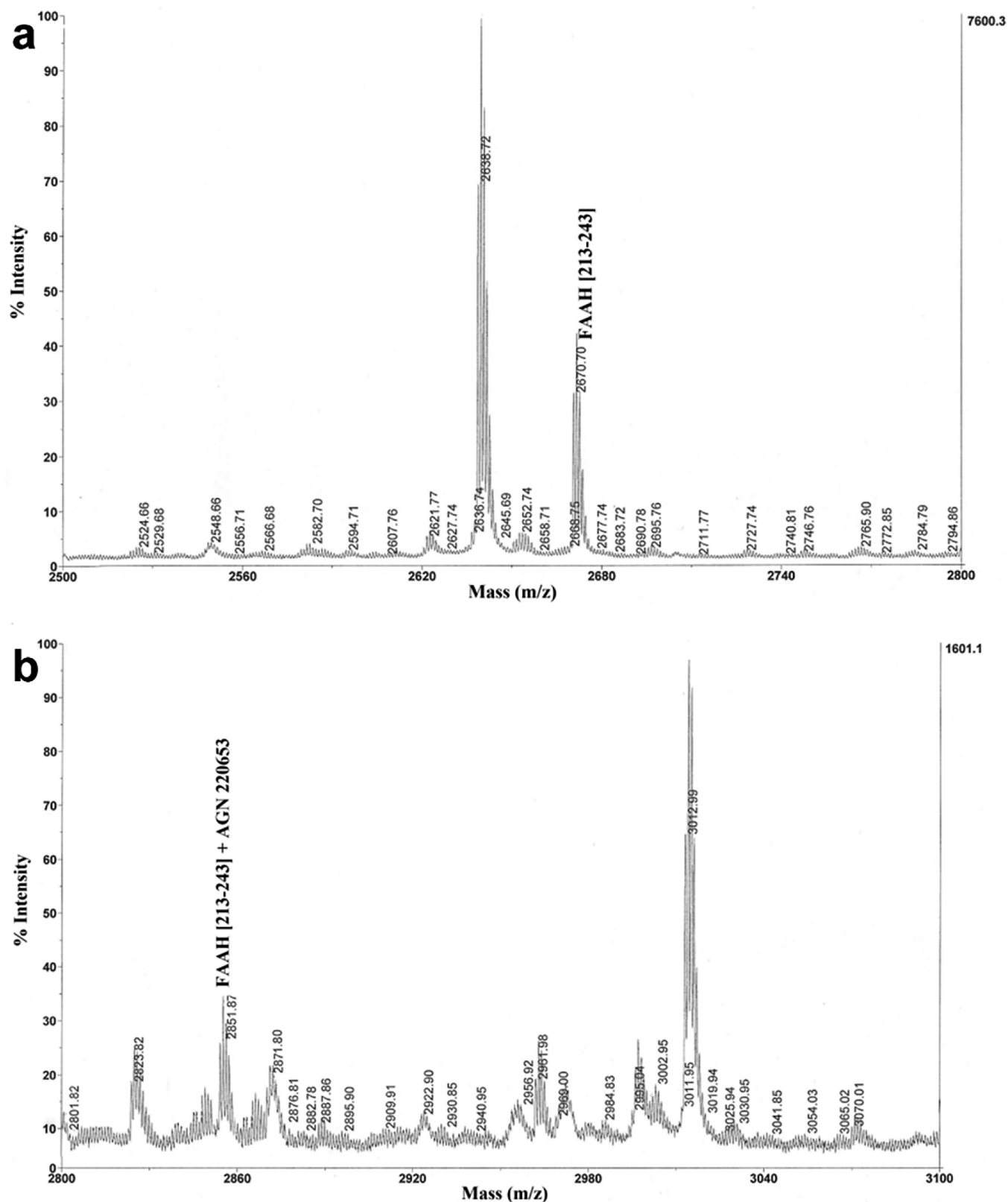
Figure 4. Partial MALDI-MS spectrum of th e tryptic digest of rhFAAH incubated with AGN220653 highlighting the peptides corresponding to the unbound (a) and bound (b) state.

protein residues include all the FAAH residues identified by Chudyk et al.[11] as critical or relevant for FAAH activity. A noticeable feature is that Ser193, the only (moderately) important residue in ref 11 not forming relevant contacts with S99, will show its importance for the C−C bond cleavage in FAAH-

AGN220653. The only significant exception to this set is represented by Arg243, critical to ensure the proper charge and orientation of Asp237, in turn required to obtain the right structure and reasonable energies for ligands in the oxanionic hole (Ile238-Gly239-Gly240-Ser241). This feature was also
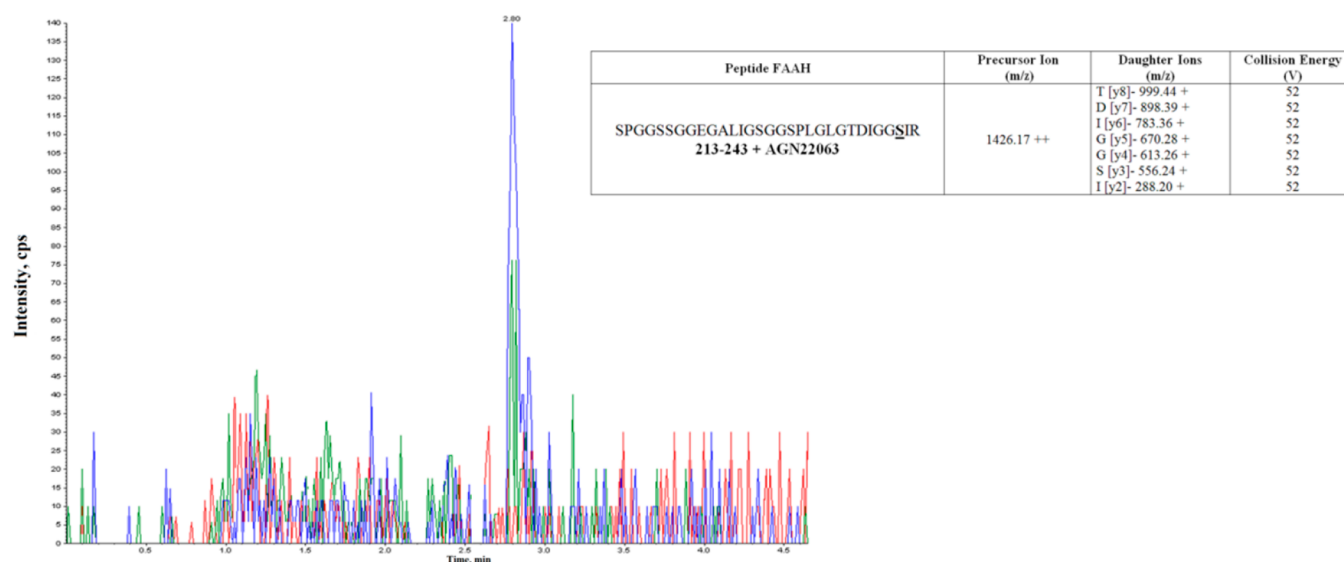
E

Peptide FAAH	Precursor Ion (m/z)	Daughter Ions (m/z)	Collision Energy (V)
SPGGSSGGEGALIGSGGSPLGLGTDIGG**S**IR 213-243 + AGN22063	1426.17 ++	T [y8]- 999.44 + D [y7]- 898.39 + I [y6]- 783.36 + G [y5]- 670.28 + G [y4]- 613.26 + S [y3]- 556.24 + I [y2]- 288.20 +	52 52 52 52 52 52 52

Figure 5. Multiple reaction monitoring tandem mass spectrometry analysis performed on the trypsin digest of FAAH incubated with AGN. The transitions and the mass spectral parameters used for the selective isolation of modified FAAH peptide 213−243 are reported in the inset. The MRM current corresponding to the transitions detected for the peptide 213−243 carrying the AGN moiety at level of Ser241 and coeluting at r.t. for 2.80 min is reported.
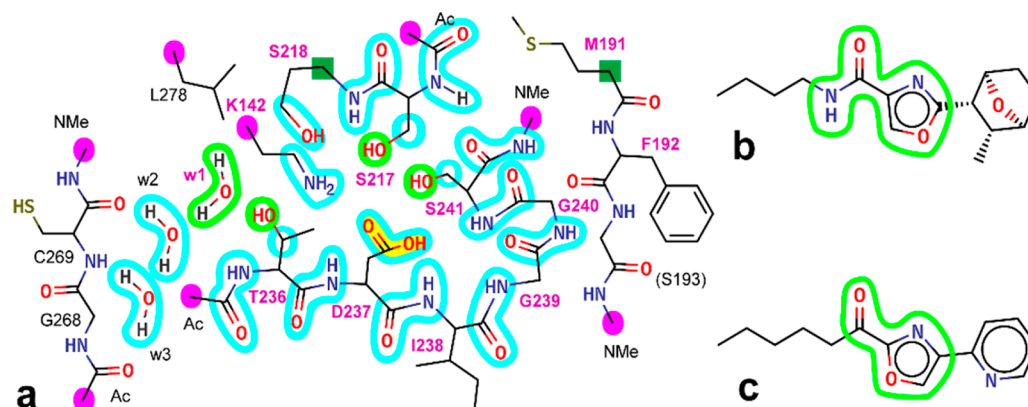

Figure 6. QM model of FAAH-complexes: truncated FAAH (FAAH^T, panel a), AGN220653 (AGN220653^T, panel b), and S99 (S99^T, panel c) structures used in QM calculations are shown omitting nonpolar hydrogen atoms. FAAH residues substantially contributing to enzyme activity in experimental/computational studies are labeled in bold magenta. Cyan and green lines delimit atoms described by "improved" and "maximum" basis set functions, respectively. The area enclosing Asp237 side chain atoms with restrained torsional angles is colored in yellow. Positionally restrained C atoms are marked with magenta circles, when also the three dihedral angles ending on the three H atoms adjacent to the marked atom are restrained, or dark green squares, when only the position of the C atom is fixed.

confirmed by preliminary calculations on either the nonbonded, or the covalently bonded form of the FAAH-S99 complex (not shown). In both cases, the ligand carbonyl oxygen atom was unable to assume the right position inside the oxanionic hole when Asp237 was omitted, misoriented, or set in its charged form in the absence of Arg243. These calculations showed that the electronic and geometrical effects of the Arg residue can be recovered by adopting a protonated form for Asp237 side chain and restraining its torsional degrees of freedom (four dihedral constraints, with ending atoms enclosed in a yellow area in Figure 6a). In this view, Arg243 and Asn498 were deleted from the truncated protein model (named FAAH^T). FAAH^T forms exhibiting neutral Lys142/Ser241 or protonated Lys142/deprotonated Ser241, i.e., corresponding to the end states of the FAAH catalytic proton relay, are indicated as nFAAH^T and cFAAH^T, respectively. Further details on the models are described in Methods. The "truncated" model of the S99 ligand from PDB: 2WJ1 was named S99^T. FAAH^T-S99^T included 262 atoms (124 heavy atoms, see Figure 6c). Similarly, a reduced

model of the ligand was named AGN220653^T. The FAAH^T-AGN220653^T complex included 270 atoms (126 heavy atoms).

The large size of the QM systems to be simulated required some compromise between the QM level adopted to describe the atoms and the completeness of the approach used to sample the potential energy surface of the systems. In this case, because of the four criteria described above, and in the absence of quantitative data on energetics and kinetics of FAAH-AGN220653, we decided to maximize the QM level, while employing simple unidimensional (and limited bidimensional) adiabatic mapping (AM) scans of the main degrees of freedom of the three reaction steps under investigation: Ser241 → Ser217 → Lys142 proton transfer; Ser241 Oγ → AGN220653 carbonyl-C nucleophilic attack with tetrahedral covalent intermediate (TCI) formation; and AGN220653(C4) protonation by Ser217(Hγ) with simultaneous AGN220653 C4-carbonyl C bond-breaking and final carbamate formation. This choice limited the number of degrees of freedom systematically explored, the characterization of their couplings and the completion of such explorations.

Table 1. Energetics of the Main Minima of FAAH^T-AGN220653^T and FAAH^T-S99^T Complex QM Models^a

system	QM setup	ΔE(MinNBneu) (kcal mol^{-1})	ΔE(MinNBchg) (kcal mol^{-1})	ΔE(MinCB) (kcal mol^{-1})	ΔE(MinDis) (kcal mol^{-1})
FAAH^T-AGN220653^T	MB-SV/MB-SV	1.8	4.9	11	0
	MB-TZ/MB-TZ	4.3	nc	nc	0
	EMBTZVPP/MB-SV	1.3	nc	nc	0
	EMBTZVPP/MBTZ	2.3	nc	nc	0
	DLPNO-CCSD(T) cc-pVTZ/ MB-SV	3.1	9.4	13.5	0
FAAH^T-S99^T	MB-SV/MB-SV	15.6	uns	**0**	11.7
	DLPNO-CCSD(T) cc-pVTZ/ MB-SV	16.5	uns	**0**	7.5

^{a}Energy differences (ΔE, in kcal mol^{-1}) calculated from the minimum of each structure set for a given QM setup (energy calculation/geometry). MB-SV and MB-TZ indicate "MixBasSV" and "MixBasTZ" setups, respectively. MinNBneu, MinNBchg, MinCB, and MinDis indicate nFAAH^T-Lig^T noncovalent, cFAAH^T-Lig^T noncovalent, cFAAH^T-Lig^T covalent, and nFAAH^T-LigFrag^T covalent + LigLG^T complexes, respectively ("Lig" = "AGN220653" or "S99", "LigFrag" = ligand fragment still covalently bonded to FAAH after elimination of the LigLG protonated leaving group), nc = not calculated, uns = unstable.

In this view, no dynamic study of the system was attempted, and only an approximation to the "real" energy barriers and intrinsic reaction coordinates was obtained. The implications of such approach for the different degrees of freedom involved in the three steps are discussed in the "Quantum Mechanics (QM) Calculations" section of the Supporting Information (page S8).

While the exact QM setups for the different methods and protocols are described in Methods, a noticeable feature of our models is the adoption of three levels of basis set functions in DFT+D3 calculations: "standard", for atoms only contributing to the overall system stability and ligand orientation within the site; "improved", for atoms forming relevant polar protein−ligand interactions; and "maximum", with inclusion of diffuse functions, for atoms potentially involved in the reaction mechanism, possibly bearing at some stage a negative charge surplus. The assignment of the different protein and ligand atoms to the three groups is shown in Figure 6 where cyan and green lines delimit atoms described by "improved" and "maximum" sets, respectively, and listed in Methods, where the different basis sets are also described. Note that for S99 complexes, atoms from oxazole and the adjacent 2-pyridyl group were included in the "maximum" set.

The reference calculations fully validated the required features of the computational approach of the system both functionally and structurally. Functionally (criteria II and III): in fact, S99^T, in perfect agreement with available data, easily produces a covalent unreactive complex with FAAH, whose formation already starts at the end of the relay protein transfer stage of the FAAH catalytic mechanism. In particular, AM calculations on the nFAAH^T-S99^T complex showed an intramolecular proton transfer to cFAAH^T associated with a very small energy barrier (1−2 kcal mol^{-1}) and, already before the end point, a barrierless addition of the nascent −O$^-$ moiety of S241 to the S99^T carbonyl C atom. This process led to the formation of the experimentally observed tetrahedral covalent adduct,[13] characterized by an oxyanion O ligand atom stabilized by its interaction with the protein oxyanion hole. The exploration by AM of a possible further step in FAAH-S99 reaction involving cleavage of the C−C bond between the ligand oxazole ring and carbonyl group resulted in a high estimated barrier (about 45 kcal mol^{-1}) and a final dissociated system less stable than the covalent adduct (ΔE = 11.7 kcal mol^{-1} in MixBasSV/MixBasSV, 7.5 kcal mol^{-1} in DLPNO−CCSD(T) cc-pVTZ/MixBasSV, see Table 1). Structurally (criterion I): in fact, the final complex fits the corresponding experimental structure (PDB: 2WJ1), with a rmsd of less than 0.3 Å on all heavy atoms excluding side chain

atoms beyond C$_\beta$ of "frontier" residues Phe192 and Leu278 and the last three atoms of the residual aliphatic chain of S99^T (thus including 114 heavy atoms out of 124). In addition, except for two backbone oxygen atom (belonging to Gly239 and Gly268) and protruding out the simulated region (pointing toward FAAH counterparts missing in the truncated model), no atom in the FAAH^T selected set was distant more than 0.4 Å from its FAAH counterpart.

Consequently, we applied the same AM protocol to nFAAH^T-AGN220653^T complex, adopting the "reaction coordinates" described and illustrated in Figure 7 for the AM scans, obtaining a substantially different picture of the evolution of the two complexes. In fact, AGN220653^T gives rise to an unstable TI that undergoes dissociation of the carbonyl-4-oxazole bond, producing a slowly hydrolyzable carbamate derivative of Ser241, with elimination of the substituted oxazole ring (Figure 7). In particular, the proton transfer, monitored using the Ser241 H$_\gamma$−Ser241 O$_\gamma$ distance as reaction coordinate (distance d_1 in Figure 7a), exhibited a barrier of about 5 kcal mol^{-1}. The cFAAH^T-AGN220653^T minimum was about 3 kcal mol^{-1} higher in energy than the starting structure (Figure 7d) and did not show any spontaneous tendency to evolve into the covalent adduct (Figure 7e). The latter, in turn, as resulting from AM scan of the Ser241 O$_\gamma$−AGN220653^T carbonyl C distance (distance d_2 in Figure 7b), resulted about 9 kcal mol^{-1} higher in energy than the starting structure (and 4 kcal mol^{-1} higher than cFAAH^T-AGN220653^T), with a barrier of about 10.5 kcal mol^{-1} over the starting structure (5.5 kcal mol^{-1} over cFAAH^T-AGN220653^T). This marginally stable structure can either revert to the nonbonded complex, or slowly evolve toward the "dissociated" form (Figure 7f) hypothesized from isotopic labeling experiments and docking, which is about 2 kcal mol^{-1} lower in energy than the starting nonbonded complex. The dissociation occurs by protonation of the C4 oxazole atom of AGN220653^T by Ser241 H$_\gamma$, (distance d_3 in Figure 7c) and simultaneous breaking of the oxazole C4-amide C bond (distance d_4 in the same figure), accompanied by a back-transfer of the proton from Lys142 to Ser217 (pink and magenta lines/bonds in figures) and rotations of oxazole ring and the amide moiety of the carbamate group resulting after dissociation. This latter motion, in particular, results in the formation of a stabilizing H-bond between ligand amide H and Ser193 backbone O atoms (red dashed line in Figure 7f). This interaction contributes by about 2−3 kcal mol^{-1} to the overall stability of the dissociated form, as estimated by comparison with other AM
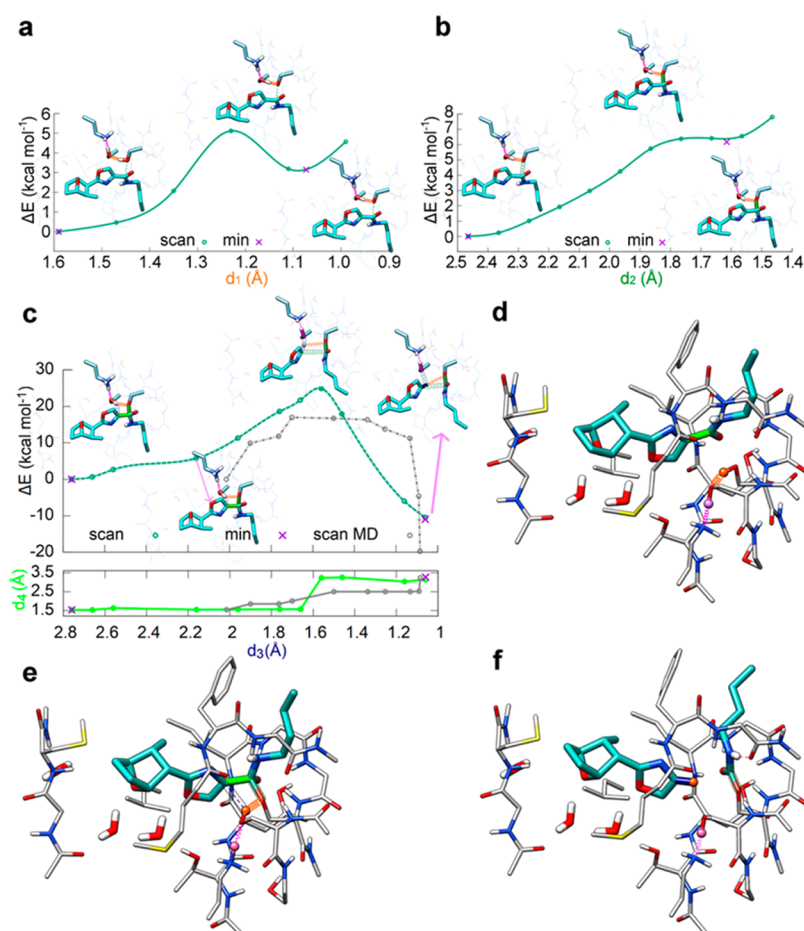
Figure 7. (a−c) QM reaction curves for Ser241 → Ser217 → Lys142 proton transfer (a), Ser241 $O\gamma$ → AGN220653^T carbonyl-C nucleophilic attack (TCI formation) (b), protonation of AGN220653^T(C4) by Ser217(H$_\gamma$), simultaneous AGN220653^T C4-carbonyl C bond-breaking and final carbamate formation (c). Energy differences (kcal mol^{-1}) calculated from the starting structure of each map. The most-representative distances used as "reaction coordinate": Ser241(H$_\gamma$)−Ser241(O$_\gamma$) (d_1, orange label/"springs") (a), Ser241(O$_\gamma$)−AGN220653^T(carbonyl C) (d_2, dark-green label and distance dashed-line plus transparent stick/opaque stick → before/after full covalent-bond formation) (b), Ser217(H$_\gamma$)−AGN220653^T(C4) (d_3, dark-blue label and distance dashed-line plus transparent stick/opaque stick → before/after full covalent-bond formation), plus AGN220653^T(C4)−(carbonyl C) (d_4, green label, plot line and distance opaque stick/dashed-line plus transparent stick → before/after full covalent-bond breaking), where dashed-lines remark that no simple "reaction coordinate" exists for this step, the curves resulting from several one-/two-dimensional scans on d_3 and/or d_4 (c). Other relevant distances: Lys142(N$_\zeta$)−(H) (pink), Lys142(N$_\zeta$H)−Ser217(O$_\gamma$) (magenta), Ser217(O$_\gamma$)−(H pointing toward AGN220653^T(C4)) (medium-blue). Ligand plus FAAH catalytic-triad side chains/other bonds shown as sticks/lines. Gray dash-and-dot/solid lines show the QM energy vs d_3/d_4 vs d_3 profiles for calculations started from the selected MD frame of the TCI complex. (d−f) detail of starting noncovalent (d), the unstable TCI (e) and the final dissociated (f) complexes. Representation details: stick bonds (bold: ligand and water); ligand/other C atoms colored cyan/very-light-blue; relevant distances, bonds and H atoms: see panels (a)−(c).

calculations ending with the ligand amide NH pointing toward Met191 backbone O atom, with formation of a weaker, suboptimal H-bond. Interestingly, this stabilizing interaction for the final carbamate is also found in PDB entry 3LJ7, an experimental structure featuring a carbamate adduct on FAAH Ser241. However, the influence of this interaction on energy is already seen during C−C bond cleavage, since it also reduces the barrier of a comparable amount to that found on final minima in comparison with the aforementioned calculations in which the ligand amide moves toward Met191 backbone.

The upper limit to the energy barrier predicted for C−C bond dissociation in TI (turquoise dashed line in Figure 7c), amounts to about 25 kcal mol^{-1}, but about one-third of it (i.e. 10 kcal mol^{-1}) is just required to reduce the relatively large distance between HO217 and C4 atoms (r_{HC}) in the starting structure (~2.8 Å) to 2.0 Å, where they really start to interact. Since we expected that, similarly to anandamide, this kind of distance reduction[14] or a distortion from planarity of the center

adjacent to the ligand carbonyl group[15,16] could further reduce the barrier, we investigated if overall protein motions at room temperature could reduce r_{HC} and/or decrease the planarity around C4. These motions involve the whole (dimeric) protein, depend on FAAH embedding into cell membrane, and occur on a relatively large time scale (>10^{-9} s). Therefore, we recurred to standard classical molecular dynamics (MD) of a covalently bonded complex of the AGN220653 fragment used in QM calculations with each monomer of FAAH dimer embedded in membrane and solvated.

The MD trajectory resulted quite stable in terms of backbone fluctuations both of the overall dimeric structure (Figure S3a) and of each FAAH monomer over itself (Figure S3b) or over the other monomer (Figure S3c), as well as for the embedding of the transmembrane helix in the lipid bilayer (Figure S3d). This result shows that neither large interdomain relative motions, nor overall major drifts or rearrangements occurred over the simulated time.

However, local breathing motions are observed, also involving the active site region. The resulting 400 ns MD trajectories show that these motions also affect ligand-protein interatomic distances. In particular, in ~10% of the sampled structures r_{HC} is significantly shorter (2.4−2.6 Å) than in our starting model, in ~1% it is very short (2.2−2.4 Å) and in ~0.1% it is shorter than 2.2 Å. In addition, several structures with r_{HC} < 2.2 Å also exhibit a significant distortion (12°) of the planarity around C4 (Figure 8a,b). This
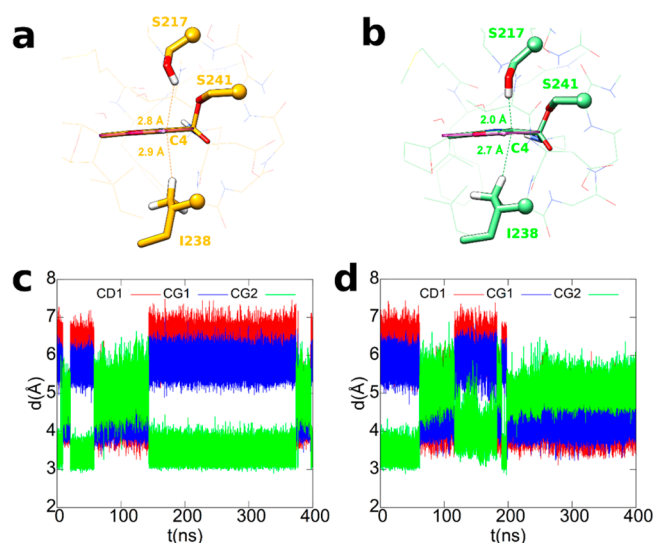
Figure 8. Influence of protein motions on the structure of FAAH-AGN220653[T] TCI complex. Details of the structures of QM-refined FAAH-AGN220653[T] TCI complexes obtained from the selected docking model (a) and "shrunk-site" frame of MD simulation (b) show the variation in distances between AGN220653[T](C4) and either Ser217(H$_\gamma$) or Ile238(H$_{\gamma21}$) atoms. The distortion of the planarity of the oxazole ring at C4 on going from panel (a) to (b) is highlighted by trasparent magenta discs enclosing the ring plus directly bonded atoms except C4. Ser217, Ile238, and Ser 241 side chain bonds, the corresponding C$_\alpha$ atoms, AGN220653[T] bonds are represented as thick sticks, spheres, and thin sticks, respectively. The constant presence of a close contact between AGN220653[T](C4) and Ile238 side chain is shown by the time-series of the distances between AGN220653[T](C4) and either Ile238(C$_{\delta1}$) (red), Ile238(C$_{\gamma1}$) (blue), or Ile238(C$_{\gamma2}$) (green) atoms for FAAH A (c) and B (d) monomers from the MD trajectory of the dimeric FAAH-AGN220653[T] covalent complex.

distortion can be attributed to the relative higher rigidity of the residues involved in the oxyanion hole (Ile238-Gly239-Gly240-Ser241), through a close contact between AGN220653[T] C4 atom and Ile238 side chain. When the site-shrinking motions of the rest of the active site are propagated to the ligand, reducing the r_{HC} distance by up to 0.8 Å, AGN220653[T] is unable to further transmit the motion to the close-contacting side chain of Ile238. The latter, in fact, is already at van der Waals distance from C4 (maximum distance decrease of about 0.2 Å) and can just undergo a limited rotation, which, however, only cyclically permutes over time the specific Ile238 atom in closest contact with AGN220653[T] C4 among CG1, CD1, and CG2, thus preserving the steric interaction with C4 atom during the whole MD simulation (Figure 8c,d). Globally, this stable interaction determines a deformation of AGN220653[T] ring, Ile238 side chain behaving like a rigid "anvil" and the rest of the site acting as a "hammer" hitting AGN220653[T] oxazole ring at C4 atom, thus distorting it from planarity.

QM energy minimization (EM) of the MD structure with the shortest r_{HC} distance confirmed the deformation of the oxazole

ring observed in MD (Figure 8). Application of the above QM protocol to this EM structure reduced the barrier by ~10 kcal mol^{-1} (Figure 7f) and confirmed that Ile238 side chain facilitates C−C bond cleavage by pushing the oxazole ring toward Ser217 HO atom and deforming it (Figure 8b).

EM and single-point calculations on the main FAAH[T]-AGN220653[T] and FAAH[T]-S99[T] minima performed with larger basis sets or higher levels of theory (Tables 1, S4, and S5) fully supported the MixBasSV/MixBasSV results so far discussed, showing a substantial independence of the overall predicted behavior of AGN220653 and S99 from the specific QM setup, thus satisfying criterion (IV) for the computational approach. In particular, relative energies (Table 1) are very moderately (less than 4 kcal mol^{-1}), and geometries (Table S5) are only marginally (less than 0.12 Å), affected by increasing basis set size, as shown by the comparison of energies and atom rmsd values between the MixBasSV/MixBasSV and MixBasTZ/MixBasTZ calculations in Tables 1 and S5, or by energies between the values calculated with EMBTZVPP basis sets on either MixBasTZ or MixBasSV geometries. Even moving to higher levels of theory, relative energies are still scarcely affected (less than 4 kcal mol^{-1}) (comparison of DFT+D3 with DLNPO-CSSD(T) results). These results are confirmed by those on FAAH[T]-S99[T], where a difference of 3.4 kcal mol^{-1} is observed for the relative stability of covalent vs nonbonded complexes. Thus, the tendency of FAAH[T]-AGN220653[T] to dissociate cannot be apparently ascribed to biases in the computational procedures, since FAAHT-S99[T], in perfect agreement with experimental evidence, exhibits a considerable preference for the TCI form vs the nonbonded complex and no tendency to dissociate, and these trends are even more pronounced when considering the results at the higher level of theory (CCSD(T)).

In summary, the overall picture of the reactive behavior of the FAAH-AGN220653 complex emerging from the QM analysis (Figure 9) is in agreement with the biochemical, docking, and MS data presented above, and shows the initial formation of a noncovalent protein−ligand interaction in the FAAH binding site. As suggested by docking, the oxazole ring occupies a region and exhibits an orientation in the cytosolic port similar to the corresponding groups in ketone FAAH inhibitors OL-135/S99, and the adjacent amide group and hydrophobic tail dock in the acyl-binding pocket. This initial arrangement is rather stable, but can slowly evolve toward dissociation according to a "substrate-like" mechanism, albeit involving the atypical oxazole leaving group. The instability of the "reversible-like" FAAH-AGN220653 tetrahedral covalent intermediate, in spite of the apparent strong similarity with the structure observed in FAAH-S99, mainly depends on the presence of an amide group. This implies the destruction of the amide conjugation upon formation of the TCI, associated with a high energetic cost, increased by minor contributions from less-favorable steric interactions determined by the different oxazole attachment between AGN220653 and S99. Thus, the unstable TCI can dissociate with the assistance of Ser217 (and the final contribution of Ser193 backbone O atom described above), resulting in the prediction of an overall energetically favorable process. It is noteworthy that these QM results have been obtained in spite of the very conservative conditions adopted in the computational approach. In particular, favorable entropic and enthalpic effects expected after departure of the dissociated substituted oxazole from the final QM complex have not been accounted for, and the coupling between the motions of the QM-simulated region and the overall enzyme has been only partially characterized by MD plus QM
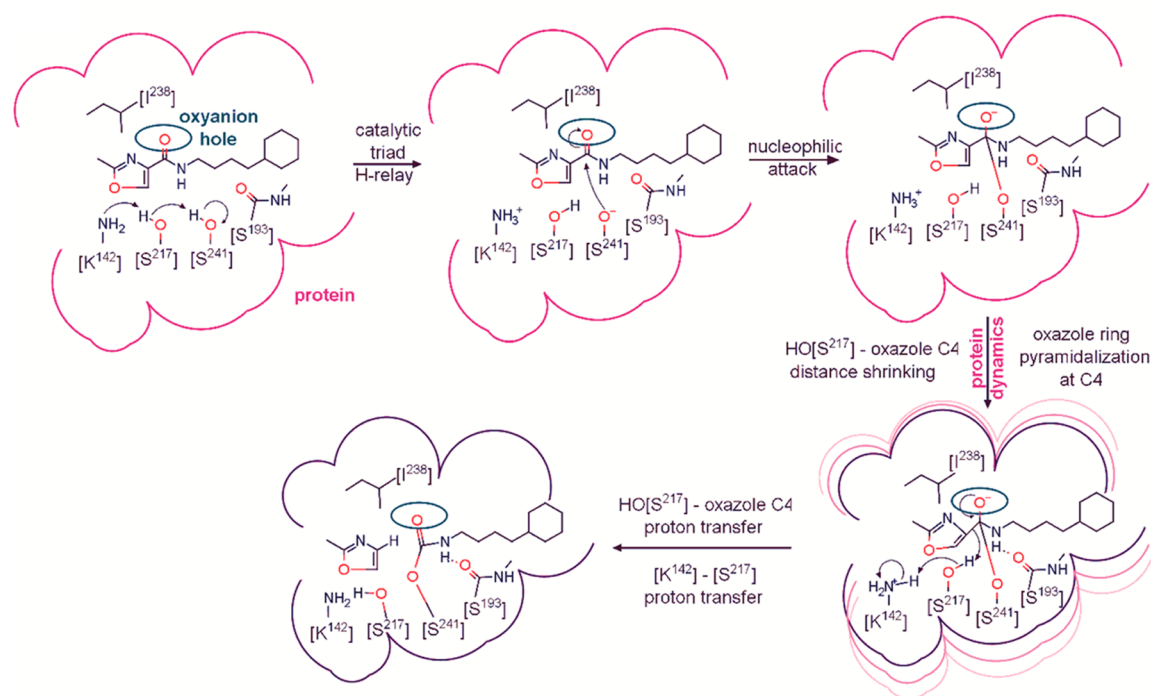
Figure 9. FAAH-AGN220653 reaction scheme up to carbamate formation and detachment of the substituted oxazole. The relevant FAAH side chain/backbone atoms and ligand AGN220653[T] fragment are shown. Lines represent FAAH oxyanion hole (blue) and active site walls before (magenta) and after (purple) the site "shrinkage" recurring in MD simulations (pink-to-purple concentric lines).

calculations. This coupling typically helps the reactions by "shrinking" the paths associated with the highest energy steps[14] and/or decreasing the planarity of unsaturated groups therein involved,[15,16] thus decreasing the corresponding energy barriers.

CONCLUSIONS

The catalytic mechanism of FAAH is unique. Rather than the more common serine−histidine−aspartate triad found in classical serine hydrolases, it involves a catalytic triad consisting of two serine residues and one lysine residue.[1] A multitarget drug containing a carboxyamide-4-oxazole moiety was found here to exhibit inhibition of FAAH by exploiting a novel catalytic mechanism of this enzyme. Catalytic promiscuity, i.e., the tendency to catalyze different reactions with the same machinery, depends on (1) the ability to bind different substrates, (2) the requirements of these substrates for different chemistry, and (3) changes in hydrogen bond partners or Bronsted acids/bases in the active sites, which create new binding pathways.[17] From a mechanistic point of view, the unique scaffold of AGN220653, with its "swapped substrate" amide configuration, in which the alkyl chain is bound to the amide nitrogen rather than the carbonyl atom as in FAAH substrates, resembles that of previously reported carbamate or urea FAAH inhibitors. These compounds also behave as slowly reversible inhibitors.[18,19] However, in AGN220653, the amide carbonyl group is also bound to a carbon instead of the heteroatom (O or N) that provides a good leaving group in the carbamylation step of the FAAH inhibition mechanism followed by ureas or carbamates. The occurrence of a CO-C bond in this "swapped substrate" motif also recalls reversible ketone FAAH inhibitors, in which, however, the carbonyl group is involved in two CO−C bonds.[13] Thus, while the new motif was likely to interact with FAAH, its mechanism of action was not easily predictable in advance.

Indeed, the operating mechanism of FAAH on AGN220653 reported here following the use of a variety of biochemical, computational, mass spectrometric, and quantum mechanics methodologies is quite different from the one normally used by the enzyme on its substrates or most of its inhibitors. We propose that this difference is due to two main related reasons: (1) the substrate exhibits an opposite orientation of its amide bond within the site in comparison with conventional substrates; and (2) as a consequence, the resulting mechanism now requires the protonation of the oxazole C4 atom, an acceptor of the Ser217 proton far less effective than the nitrogen atom typically involved. These findings expand to a new level the already wide array of catalytic capabilities of serine hydrolases and amidases, particularly of an atypical one such as FAAH. They illustrate a case in which the structural features and working mechanism of the "enzyme machinery" override the binding and reactivity propensities of the substrates, resulting in the unusual cleavage of the C−C bond of the substrate-inhibitor. An alternative interpretation from the "ligand's point of view" is that, while many inhibitors work by counteracting/blocking FAAH catalytic mechanism at the early binding or covalent adduct stages, AGN220653 instead acts by "seconding" it down to the substrate protonation and dissociation step, albeit with a noncanonical orientation within the binding site. This results in the overall reduction/temporary block of FAAH activity through the formation of a slowly hydrolyzable covalent final derivative.

Serine hydrolases with "classical" serine−histidine−aspartate catalytic triads are not new to catalyze the hydrolysis of C−C bonds.[17,20] Thus, it is perhaps not too surprising that FAAH, although being characterized by an unusual catalytic triad, is also able to use such a mechanism of action. Nevertheless, given the crucial role of this enzyme in controlling the levels of neuroactive mediators such as the endocannabinoids, the N-acylethanolamines

and other long chain fatty acid amides, including the recently discovered N-oleoyl-glycine,[21] the present findings have obvious implications for the development of new inhibitors of potential therapeutic importance for the treatment of neuropathic pain and neuropsychiatric and neurological disorders.

■ METHODS

Chemical Synthesis. Molecules were synthesized at Selcia Laboratories (Ongar, Essex, United Kingdom) starting from the oxabicycloheptane scaffold as previously described.[8] Selcia Laboratories also performed the [14C]-labeling of compounds. Further details on chemical syntheses as well as compound characterization data by mass spectrometry and NMR are available in the Supporting Information.

In Vitro Analyses of Inhibitor Potency on Different FAAH Preparations. AEA hydrolysis was measured by incubating the compounds with different enzyme preparations (10 000g membrane fraction of rat brain [70 μg per sample], rat basophilic leukemia cells [RBL-2H3, 100 μg per sample], or mice neuroblastoma cells [N18TG2, 100 μg per sample]). Two commercially available enzymatic sources, the human recombinant FAAH (Cayman, 2 μg per sample) and the custom and highly purified humanized FAAH (Medicilon, Ltd., Shangai, China 0.2 μg per sample) were tested. Compounds were incubated with the enzyme in Tris−HCl 50 mM, at pH 9.5 at 37 °C for 30 min, with synthetic N-arachidonoyl-[14C]-ethanolamine (55mCi per mmol, ARC St. Louis, MO) properly diluted with AEA (Tocris Bioscience, Avonmouth, Bristol, UK). Reactions were blocked by organic extraction with ice-cold chloroform/methanol (1:1, vol:vol), and the aqueous phase containing [14C]-ethanolamine produced by AEA hydrolysis was measured by scintillation counting. For time-course experiments, compounds were preincubated 20 min with FAAH (rat brain membrane) before the addition of the [14C]-substrate. After a further 30 min incubation, AEA hydrolysis was measured as previously described.[22] In a different set of experiments rat brain membrane (70 μg per sample) was incubated in the presence and absence of increasing concentrations of AGN220653, MAFP or URB-597 (10−50−100−500−1000 nM) with N-arachidonoyl-[14C]-ethanolamine (55mCi per mmol, ARC St. Louis, MO, USA) for different time points (0.5−1−2−4−24 h). After each time point, the reaction was terminated by organic extraction with ice-cold chloroform/methanol (1:1, vol:vol) and the aqueous phase containing [14C]-ethanolamine was measured by scintillation counting as above-mentioned. All data are expressed as the concentration exerting 50% inhibition of [14C]-AEA hydrolysis (IC_{50}). Values indicate averages ± SD of three independent experiments conducted in triplicate.

Competitive Activity-Based Protein Profiling (ABPP). Rat brain membranes were prepared according to previously reported methods[23,24] and diluted to 1 mg/mL prior to use. Proteomes (50 μL) were preincubated with either DMSO or 1−1000 nM concentrations of inhibitor at 37 °C. After 20 min, fluorophosphonate-rhodamine (FP-Rh) (1.0 μL, 50 μM in DMSO, a kind gift of Prof. Benjiamin Cravatt) was added and the mixture was incubated for another 30 min at 37 °C. Reactions were quenched with SDS loading buffer (12.5 μL, 5×) and run on SDS-PAGE. Following gel imaging, serine hydrolase activity was determined by measuring fluorescent intensity of gel bands corresponding to MAGL, ABHD6, and FAAH using ImageJ 1.43u software.

Detection and Quantitation of Radiolabeled Complexes between Inhibitors and Humanized FAAH by Autoradiography. AGN 220653 and AGN 220346 were [14C]-radiolabeled as above-described at the 1-position of the cyclohexyl-butyl amine portion. The 1-position of the isopropyl group in AGN220346 and phenyl group of the sulfonamide ring in AGN220653 were also [14C]-radiolabeled. [14C]-AGNs (75000 cpm; 30 μM) with custom and highly purified humanized FAAH (0.2 μg/s; Medicilon, Ltd., Shangai, China, 5 min in 50 mM Tris-HCl buffer, pH 9.5) were incubated in the presence and absence of an excess of the corresponding nonradiolabeled AGNs (700 μM) in a final volume of 20 μL. Reactions were terminated by boiling samples in SDS-sample buffer (5 min 95 °C). Denatured samples were next electrophoresed through a

10% SDS polyacrylamide gel and the gels were dyed with Comassie Blue (45 min) before a further incubation in a destaining-fixing solution (Kodak; 45 min). In order to improve detection sensitivity, before drying gels at room temperature, gels were exposed to Amplify Fluorographic Reagent (GE Healthcare, Amersham; 30 min) and finally to a radiographic film (Hyperfilm, Amersham) for 78 days at −80 °C. Autoradiographic images were acquired by ChemiDoc (Bio-Rad Laboratories).

Molecular Docking. Starting ligands geometries were built with Ghemical 2.99.2,[25] followed by energy minimization (EM) at molecular mechanics level first, using Tripos 5.2 force field parametrization,[26] and then at AM1 semiempirical level. For docking calculations, both cis and trans conformations of the acyl sulfonamide group of AGN220346 and AGN220653 were taken into account. Docking studies were performed with both AutoDock 4.2[27] and AutoDock Vina 1.1.2[28] by using two crystallographic structures of FAAH, complexed with two different inhibitors, namely, URB597 (PDB entry 3LJ7) and OL-135 isomer (PDB entry 2WJ1). UCSF Chimera 1.10.1[29] was used to add hydrogens to protein structures, keeping the catalytic Lys142 in the nonprotonated form. Both proteins and ligands were processed with AutoDock Tools (ADT) package version 1.5.6rc12[27] to merge non polar hydrogens, calculate Gasteiger charges and select the rotatable side-chain bonds. Protein side-chains of Phe192 and Met191 and all rotatable bonds of both ligands were made flexible. The grid for docking evaluation with a spacing of 0.375 Å and 70 × 70 × 60 points, centered on the catalytic site, was generated using the program AutoGrid 4.2 included in Autodock 4.2 distribution. A 100 molecular docking run was performed adopting a Lamarckian Genetic Algorithm (LGA) and the following associated parameters: 100 individuals in a population with a maximum of 15 million energy evaluations and a maximum of 37 000 generations, followed by 300 iterations of Solis and Wets local search. Default parameters were used for AutoDock Vina.

Quantum Mechanical (QM) Calculations. *FAAH-Ligand Reduced Complexes.* To simulate the presence of the full protein and to avoid end effects at the protein cut points, the position of the terminal atoms of each sequence of the FAAH^T model was rigidly restrained. Thus, one positional (on C atom) plus three dihedral (ending on the three H atoms) constraints were applied for each terminal methyl group (magenta circles in Figure 6a), while only the positional constraint was used for terminal methylene groups (dark green squares in Figure 6a). The systems were completed by addition of three water molecules from PDB 2WJ1, which are also conserved in other FAAH-ligand complex experimental structures.

Setup and Protocol. QM calculations were run either at the DFT+D3 level, i.e., DFT with inclusion of the DFT-D3 atom-pairwise dispersion correction according to the Becke−Johnson damping scheme (D3BJ),[30] or at the DLPNO-CCSD(T) level.[31] The B3LYP functional[32,33] was used in DFT+D3. Calculations were run with ORCA program (version 3.0.3[34] for DFT and 4.0.0[35] for DLPNO-CCSD(T) calculations), and include single-point energy evaluation (DFT+D3 or DLPNO-CCSD(T)), energy minimization (EM) (DFT+D3), and adiabatic mapping (AM) approach of single geometric parameters with full optimization of the remaining degrees of freedom, starting from geometry and wave function of the previous scan point (DFT+D3).

DFT+D3 calculations were performed with different basis set ensembles (Figure 6):

(i) "MixBasSV" featuring def2-SV(P)[36−38] for "standard" hydrogen, carbon, nitrogen, oxygen; def2-TZV(P)[36,39] for "standard" sulfur; def2-SVP[39] for "improved" atoms (peptide −CONH− of Ser217, Ser218, Thr236-Ile242, side chain −CH2OH of Ser218, side chain −COOH of Asp237, side chain −CH2−NH2 of Lys142, βCH2 of Ser217 and Ser241, βCH of Thr236); and ma-def2-SVP[40] for the "maximum" set (−OH of Ser217, Thr236, Ser241, AGN220653 amide −NHCO− and oxazole atoms, and the water molecule H-bonded to Thr236).

(ii) "MixBasTZ" using the MixBasSV setup for "standard" atoms; def2-TZV(d)[36,39] for "improved"; and ma-def2-TZVP[40] for "maximum" atom sets.

(iii) "EMBTZVPP", ORCA 3.0.3 DefBas-3 (TZV(d) basis set) setting[36,39] for "standard" atoms; def2-TZVPP[36,39,41] for "improved"; and ma-def2-TZVPP[40,41] for "maximum" atom sets.

In DFT+D3 calculations, the RIJCOSX approximation[42] was used to the Hartree−Fock exchange term, employing the auxiliary basis sets automatically set by ORCA.

For DLPNO-CCSD(T) calculations, since preliminary tests on FAAH complexes performed at the "double-ζ" level with either cc-pVDZ basis set[43,44] alone or cc-pVDZ for "standard" and "improved" atoms, plus aug-cc-pVDZ[43,44] for the "maximum" set calculations showed small differences between the results obtained with the two setups, the final "triple-ζ" level calculations were run with the cc-pVTZ[43,44] set on all atoms.

For DLPNO-CCSD(T) calculations, only the resolution of identity approximation for the resolution of MP2 integrals, mandatory in the ORCA implementation of this method, was applied, using the corresponding cc-pVnZ/C auxiliary basis set[45] with n = D or T for double- or triple-ζ functions.

The whole systems were "immersed" in a low polarity (chloroform) implicit solvent bath, to simulate protein bulk effects. The COSMO[46] and C-PCM[47] implicit solvent implementations into ORCA were used in DFT+D3 and CCSD(T) calculations, respectively.

Classic Molecular Dynamics (MD). *System Preparation and Parametrization.* The modified residue (hereinafter termed SAG) formed by truncated ligand AGN220653[T] covalently bonded to catalytic Ser241 was fully optimized using the GAMESS program[48] at the Hartree−Fock level with STO-3G basis set and subjected to HF/6-31G*/STO-3G single-point calculations to derive its partial atomic charges by the RESP procedure.[49] The dimeric X-ray structure 2WJ1 was completed with the residues encompassing the N-terminal transmembrane helix (9−33) and the C-terminal residues (578−579) by using the program Modeller v9.15.[50] The resulting protein complex, after addition of all hydrogen atoms, underwent 500 steps of steepest descent followed by 5000 steps of conjugate gradient EM, using ff14SB version of AMBER[51] force field for the protein and gaff parameters for the ligand adduct.[52] Then, the dimeric protein was embedded in a pre-equilibrated POPC bilayer and solvated with TIP3P water molecules by using the "Membrane Builder" plug-in available at the CHARMM-GUI Web site (http://www.charmm-gui.org). Potassium and chloride ions were added to ensure electric neutrality and to reproduce an ionic strength of 0.15 M. Charmmlipid2amber.py script available in AmberTools v16 was used to convert the resulting files to the Amber Lipid14 force field naming convention.[53]

Setup and Protocol. Standard MD simulations were carried out with Amber16 pmemd.cuda module.[54−56] Temperature was controlled by a Langevin thermostat using a coupling constant of 1.0 ps^{-1}. Bonds involving hydrogen were constrained using the SHAKE algorithm,[57] thus allowing for a 2.0 fs time step. A 10.0 Å cutoff was adopted to truncate van der Waals interactions. The particle mesh Ewald method was used to evaluate electrostatic interactions. MD equilibration phase was performed as follows: (i) 5000 steps of steepest descent followed by 5000 steps of conjugate gradient EM with positional restraints on protein complex using a force constant of 10 kcal mol^{-1} Å^{-2}. (ii) A heating phase from 0 to 100 K over 500 ps in the NVT ensemble and from 100 to 300 K over 1 ns in NPT ensemble, harmonically restraining the whole system (protein+lipid) using a force constant of 10 kcal mol^{-1} Å^{-2}. The Langevin thermostat was used with the Berendsen barostat[58,59] with a target pressure of 1.0 bar and a relaxation time of 1.0 ps. (iii) 25 ns NPT with positional restraints only on the protein with a force constant of 5 kcal mol^{-1} Å^{-2}, followed by 5 ns NPT with positional restraints only on protein Cα atoms, to ensure the lipid bilayer equilibration. MD production run was carried out in NPT ensemble for 400 ns with the catalytic Lys142 in the protonated form.

Mass Spectrometric Analysis. PD-10 prepacked column was used to desalt the solution, highly purified humanized FAAH (Medicilon, Ltd., Shangai, China) labeled with AG220653, as described above, was eluted in 10 mM AMBIC. Dried fractions containing protein were dissolved in denaturation buffer: guanidine 6 M, Tris HCl 0.3 M, EDTA 10 mM at pH 8.0. Reduction was carried out by using a 10:1 DTT/cysteine molar ratio. After incubation at 37 °C for 2 h, iodoacetamide was added to perform carboamidomethylation using an excess of alkylating agent of 5:1 with respect to the moles of thiol groups. The mixture was then incubated in the dark at room temperature for 30 min.

The alkylation reaction was stopped by addition of formic acid, in order to achieve an acidic pH. The product was purified by precipitation, 100 μL volume of sample was added with 400 μL of CH$_3$OH, 100 μL of CHCl$_3$, and 300 μL of water. The resulting solution was centrifuged for 5 min at 12 000 rpm. The protein precipitated at the interface organic−inorganic phases. The supernatant was removed, and 300 μL of CH$_3$OH was added. The resulting solution was centrifuged for 5 min at 12 000 rpm, and the protein pellet it was clearly visible on the bottom of the eppendorf. The supernatant was removed, and the sample was dried. The samples were then diluted with ammonium bicarbonate 10 mM pH 7.5, and enzymatic digestion was carried out by addition of trypsin at a 50:1 protein/enzyme mass ratio at 37 °C for 16 h. The supernatants were then recovered by centrifugation, filtered on 0.22 μm PVDF membrane (Millipore), concentrated, and purified using a reverse-phase C18 Zip Tip pipet tip (Millipore). Peptides were eluted with 20 μL of a solution made of 50% acetonitrile, 0.1% formic acid in Milli-Q water and analyzed (1 μL) by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). MALDI-MS was carried out on a 4800 Plus MALDI TOF/TOF mass spectrometer (Applied Biosystems, Framingham, MA) equipped with a nitrogen laser (337 nm). The peptide mixture (1 mL) was mixed (1:1, v/v) with a 10 mg/mL solution of α-cyano-4-hydroxycinnamic acid in acetonitrile/50 mM citrate buffer (70:30 v/v). Mass calibration was performed using external peptide standards purchased from Applied Biosystems. Spectra were acquired using a mass (m/z) range of 400−4000 amu, and raw data were analyzed using Data Explorer Software provided by the manufacturer. Peptide mixtures were analyzed by MRM LC-MS/MS analysis using a 4000 QTRAP system (AB SCIEX) equipped with a HPLC EKsigent binary pump. For each run, 5 μL peptide mixture were injected and separated on a HALO C18 column 2.7 μm, 90A, 1.0 × 50 mm (AB SCIEX, Milford, MA,) at 60 °C with a flow rate of 30 μL/min. Peptides were eluted (starting 1 min after injection) with a linear gradient consisting of eluent A (0.1% formic acid in water) and eluent B (0.1% formic acid in 100% ACN) from 7% to 95% in 5 min. The column was re-equilibrated at initial conditions for 4 min with eluent A. MRM mass spectrometry analyses were performed in positive ion mode. Skyline Software drew the choice for the selection of virtual best transitions and collision energy calculated to generate maximal fragmentation intensities. The developed MRM method was used with MRM detection window set to 0.5−1.6 min per peptide, the duty cycle was set to automatic and dwell times were minimal 5 ms. Cone voltage was set to 35 V. The mass spectral MRM parameters obtained using software Skyline are reported in the Figure 5b.

■ ASSOCIATED CONTENT

⑤ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acschemneuro.8b00315.

> Synthesis and characterization of compounds; functional assay for prostanoid receptors; details of animal studies; analyses of inhibitor activity in intact cells and inhibition mechanism; additional QM, MD and MS details (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: aligresti@icb.cnr.it.
*E-mail: vdimarzo@icb.cnr.it.

ORCID ⓘ

Alessia Ligresti: 0000-0003-1787-3900
Rosa Maria Vitale: 0000-0001-9243-1307
Fabiana Piscitelli: 0000-0001-9343-4622
Anna Illiano: 0000-0003-1491-8966
Pietro Amodeo: 0000-0002-6439-7575
Vincenzo Di Marzo: 0000-0002-1490-3070

■ REFERENCES

(1) McKinney, M. K., and Cravatt, B. F. (2005) Structure and function of fatty acid amide hydrolase. *Annu. Rev. Biochem.* 74, 411–32.

(2) Maione, S., Costa, B., and Di Marzo, V. (2013) Endocannabinoids: a unique opportunity to develop multitarget analgesics. *Pain* 154 (Suppl 1), S87–S93.

(3) Patel, S., and Hillard, C. J. (2009) Role of Endocannabinoid Signaling in Anxiety and Depression. *Curr. Top. Behav. Neurosci.* 1, 347.

(4) Blankman, J. L., and Cravatt, B. F. (2013) Chemical probes of endocannabinoid metabolism. *Pharmacol. Rev.* 65 (2), 849–71.

(5) Huggins, J. P., Smart, T. S., Langman, S., Taylor, L., and Young, T. (2012) An efficient randomised, placebo-controlled clinical trial with the irreversible fatty acid amide hydrolase-1 inhibitor PF-04457845, which modulates endocannabinoids but fails to induce effective analgesia in patients with pain due to osteoarthritis of the knee. *Pain* 153 (9), 1837–46.

(6) Wang, J. W., Woodward, D. F., Martos, J. L., Cornell, C. L., Carling, R. W., Kingsley, P. J., and Marnett, L. J. (2016) Multitargeting of selected prostanoid receptors provides agents with enhanced anti-inflammatory activity in macrophages. *FASEB J.* 30 (1), 394–404.

(7) Woodward, D. F., Wang, J. W., Ni, M., Bauer, A., Martos, J. L., Carling, R. W., and Poloso, N. J. (2017) In vivo studies validating multitargeting of prostanoid receptors for achieving superior anti-inflammatory effects. *FASEB J.* 31 (1), 368–375.

(8) Woodward, D. F., Martos, J. L., Carling, W. R., Jones, A. D., and Wang, J. W. (2012) Fatty acid amide hydrolase inhibitors for treating pain. U.S. Patent US 20120329843 A1.

(9) Mor, M., Rivara, S., Lodola, A., Plazzi, P. V., Tarzia, G., Duranti, A., Tontini, A., Piersanti, G., Kathuria, S., and Piomelli, D. (2004) Cyclohexylcarbamic acid 3′- or 4′-substituted biphenyl-3-yl esters as fatty acid amide hydrolase inhibitors: synthesis, quantitative structure-activity relationships, and molecular modeling studies. *J. Med. Chem.* 47 (21), 4998–5008.

(10) De Petrocellis, L., Melck, D., Ueda, N., Maurelli, S., Kurahashi, Y., Yamamoto, S., Marino, G., and Di Marzo, V. (1997) Novel inhibitors of brain, neuronal, and basophilic anandamide amidohydrolase. *Biochem. Biophys. Res. Commun.* 231 (1), 82–8.

(11) Chudyk, E. I., Dyguda-Kazimierowicz, E., Langner, K. M., Sokalski, W. A., Lodola, A., Mor, M., Sirirak, J., and Mulholland, A. J. (2013) Nonempirical Energetic Analysis of Reactivity and Covalent Inhibition of Fatty Acid Amide Hydrolase. *J. Phys. Chem. B* 117 (22), 6656–66.

(12) Palermo, G., Rothlisberger, U., Cavalli, A., and De Vivo, M. (2015) Computational insights into function and inhibition of fatty acid amide hydrolase. *Eur. J. Med. Chem.* 91, 15–26.

(13) Mileni, M., Garfunkle, J., DeMartino, J. K., Cravatt, B. F., Boger, D. L., and Stevens, R. C. (2009) Binding and inactivation mechanism of a humanized fatty acid amide hydrolase by alpha-ketoheterocycle inhibitors revealed from cocrystal structures. *J. Am. Chem. Soc.* 131 (30), 10497–506.

(14) Lodola, A., Sirirak, J., Fey, N., Rivara, S., Mor, M., and Mulholland, A. J. (2010) Structural Fluctuations in Enzyme-Catalyzed Reactions: Determinants of Reactivity in Fatty Acid Amide Hydrolase from Multivariate Statistical Analysis of Quantum Mechanics/Molecular Mechanics Paths. *J. Chem. Theory Comput.* 6 (9), 2948–60.

(15) Palermo, G., Campomanes, P., Cavalli, A., Rothlisberger, U., and De Vivo, M. (2015) Anandamide hydrolysis in FAAH reveals a dual strategy for efficient enzyme-assisted amide bond cleavage via nitrogen inversion. *J. Phys. Chem. B* 119 (3), 789–801.

(16) Palermo, G., Branduardi, D., Masetti, M., Lodola, A., Mor, M., Piomelli, D., Cavalli, A., and De Vivo, M. (2011) Covalent inhibitors of fatty acid amide hydrolase: a rationale for the activity of piperidine and piperazine aryl ureas. *J. Med. Chem.* 54 (19), 6612–23.

(17) Rauwerdink, A., and Kazlauskas, R. J. (2015) How the Same Core Catalytic Machinery Catalyzes 17 Different Reactions: the Serine-Histidine-Aspartate Catalytic Triad of α/β-Hydrolase Fold Enzymes. *ACS Catal.* 5 (10), 6153–6176.

(18) Wyffels, L., Muccioli, G. G., Kapanda, C. N., Labar, G., De Bruyne, S., De Vos, F., and Lambert, D. M. (2010) PET imaging of fatty acid amide hydrolase in the brain: synthesis and biological evaluation of an 11C-labelled URB597 analogue. *Nucl. Med. Biol.* 37 (5), 665–75.

(19) Karbarz, M. J., Luo, L., Chang, L., Tham, C. S., Palmer, J. A., Wilson, S. J., Wennerholm, M. L., Brown, S. M., Scott, B. P., Apodaca, R. L., Keith, J. M., Wu, J., Breitenbucher, J. G., Chaplan, S. R., and Webb, M. (2009) Biochemical and biological properties of 4-(3-phenyl-[1,2,4]thiadiazol-5-yl)-piperazine-1-carboxylic acid phenylamide, a mechanism-based inhibitor of fatty acid amide hydrolase. *Anesth. Analg.* 108 (1), 316–29.

(20) Ruzzini, A. C., Bhowmik, S., Ghosh, S., Yam, K. C., Bolin, J. T., and Eltis, L. D. (2013) 20013) A substrate-assisted mechanism of nucleophile activation in a Ser-His-Asp containing C-C bond hydrolase. *Biochemistry* 52 (42), 7428–38.

(21) Donvito, G., Piscitelli, F., Muldoon, P., Jackson, A., Vitale, R. M., D'Aniello, E., Giordano, C., Ignatowska-Jankowska, B. M., Mustafa, M. A., Guida, F., Petrie, G. N., Parker, L., Smoum, R., Sim-Selley, L., Maione, S., Lichtman, A. H., Damaj, M. I., Di Marzo, V., and Mechoulam, R. (2018) N-Oleoyl-glycine reduces nicotine reward and withdrawal in mice. *Neuropharmacology* No. 18, 30131.

(22) Ortar, G., Morera, E., De Petrocellis, L., Ligresti, A., Schiano Moriello, A., Morera, L., Nalli, M., Ragno, R., Pirolli, A., and Di Marzo, V. (2013) Biaryl tetrazolyl ureas as inhibitors of endocannabinoid metabolism: modulation at the N-portion and distal phenyl ring. *Eur. J. Med. Chem.* 63, 118–32.

(23) Chang, J. W., Niphakis, M. J., Lum, K. M., Cognetta, A. B., 3rd, Wang, C., Matthews, M. L., Niessen, S., Buczynski, M. W., Parsons, L. H., and Cravatt, B. F. (2012) Highly selective inhibitors of

monoacylglycerol lipase bearing a reactive group that is bioisosteric with endocannabinoid substrates. *Chem. Biol. 19* (5), 579−88.

(24) Chang, J. W., Cognetta, A. B., 3rd, Niphakis, M. J., and Cravatt, B. F. (2013) Proteome-wide reactivity profiling identifies diverse carbamate chemotypes tuned for serine hydrolase inhibition. *ACS Chem. Biol. 8* (7), 1590−9.

(25) Hassinen, T., and Peräkylä, M. (2001) New energy terms for reduced protein models implemented in an off-lattice force field. *J. Comput. Chem. 22* (12), 1229−1242.

(26) Clark, M., Cramer, R. D., and Van Opdenbosch, N. (1989) Validation of the general purpose tripos 5.2 force field. *J. Comput. Chem. 10* (8), 982−1012.

(27) Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem. 30* (16), 2785−2791.

(28) Trott, O., and Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem. 31* (2), 455−461.

(29) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem. 25* (13), 1605−1612.

(30) Grimme, S., Antony, J., Ehrlich, S., and Krieg, H. (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys. 132* (15), 154104.

(31) Riplinger, C., Pinski, P., Becker, U., Valeev, E. F., and Neese, F. (2016) Sparse maps–A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory. *J. Chem. Phys. 144* (2), 024109.

(32) Becke, A. D. (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys. 38* (6), 3098−3100.

(33) Lee, C., Yang, W., and Parr, R. G. (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys. 37* (2), 785−789.

(34) Neese, F. (2012) The ORCA program system. *WIREs Comput. Mol. Sci. 2* (1), 73−78.

(35) Neese, F. (2018) Software update: the ORCA program system, version 4.0. *WIREs Comput. Mol. Sci. 8*, e1327.

(36) Schäfer, A., Horn, H., and Ahlrichs, R. (1992) Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys. 97* (4), 2571−2577.

(37) Eichkorn, K., Treutler, O., Öhm, H., Häser, M., and Ahlrichs, R. (1995) Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett. 240* (4), 283−290.

(38) Eichkorn, K., Weigend, F., Treutler, O., and Ahlrichs, R. (1997) Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chem. Acc. 97* (1−4), 119−124.

(39) Weigend, F., and Ahlrichs, R. (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys. 7* (18), 3297−3305.

(40) Zheng, J., Xu, X., and Truhlar, D. G. (2011) Minimal augmentation of diffuse functions. *Theor. Chem. Acc. 128* (3), 295−305.

(41) The Ahlrichs (2d,2p) and (2df,2pd) polarization functions were obtained from the TurboMole basis set library under ftp.chemie.uni-karlsruhe.de/pub/basen.

(42) Neese, F., Wennmohs, F., Hansen, A., and Becker, U. (2009) Efficient, approximate and parallel Hartree-Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree-Fock exchange. *Chem. Phys. 356* (1−3), 98−109.

(43) Dunning, T. H. (1989) Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys. 90* (2), 1007−1023.

(44) Woon, D. E., and Dunning, T. H. (1993) Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys. 98* (2), 1358−1371.

(45) Weigend, F., Köhn, A., and Hättig, C. (2002) Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys. 116* (8), 3175−3183.

(46) Sinnecker, S., Rajendran, A., Klamt, A., Diedenhofen, M., and Neese, F. (2006) Calculation of Solvent Shifts on Electronic G-Tensors with the Conductor-like Screening Model (COSMO) and Its Self-Consistent Generalization to Real Solvents (Direct COSMO-RS). *J. Phys. Chem. A 110* (6), 2235−2245.

(47) Barone, V., and Cossi, M. (1998) Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A 102* (11), 1995−2001.

(48) Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S., Windus, T. L., Dupuis, M., and Montgomery, J. A. (1993) General atomic and molecular electronic structure system. *J. Comput. Chem. 14* (11), 1347−1363.

(49) Fox, T., and Kollman, P. A. (1998) Application of the RESP Methodology in the Parametrization of Organic Solvents. *J. Phys. Chem. B 102* (41), 8070−8079.

(50) Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol. 234* (3), 779−815.

(51) Case, D. A., Betz, R. M., Cerutti, D. S., Cheatham, T. E., Darden, T. A., Duke, R. E., Giese, T. J., Gohlke, H., Goetz, A. W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T. S., Le Grand, S., Li, P., Lin, C., Luchko, T., Luo, R., Madej, B., Mermelstein, D., Merz, K. M., Monard, G., Nguyen, H., Nguyen, H. T., Omelyan, I., Onufriev, A., Roe, D. R., Roitberg, A., Sagui, C., Simmerling, C. L., Botello-Smith, W. M., Swails, J, Walker, R. C., Wang, J., Wolf, R. M., Wu, X., Xiao, L., and Kollman, P. A. (2016) *AMBER 2016*, University of California, San Francisco.

(52) Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem. 25* (9), 1157−74.

(53) Dickson, C. J., Madej, B. D., Skjevik, A. A., Betz, R. M., Teigen, K., Gould, I. R., and Walker, R. C. (2014) Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput. 10* (2), 865−879.

(54) Götz, A. W., Williamson, M. J., Xu, D., Poole, D., Le Grand, S., and Walker, R. C. (2012) Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput. 8* (5), 1542−1555.

(55) Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S., and Walker, R. C. (2013) Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput. 9* (9), 3878−88.

(56) Le Grand, S., Götz, A. W., and Walker, R. C. (2013) SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput. Phys. Commun. 184* (2), 374−380.

(57) Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys. 23* (3), 327−341.

(58) Izaguirre, J. A., Catarello, D. P., Wozniak, J. M., and Skeel, R. D. (2001) Langevin stabilization of molecular dynamics. *J. Chem. Phys. 114* (5), 2090−2098.

(59) Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Di Nola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys. 81* (8), 3684−3690.

Multiple Reaction Monitoring Tandem Mass Spectrometry Approach for the Identification of Biological Fluids at Crime Scene Investigations

Anna Illiano,*,† Valentina Arpino,† Gabriella Pinto,† Andrea Berti,‡ Vincenzo Verdoliva,‡ Giuseppe Peluso,§ Piero Pucci,† and Angela Amoresano†

†Dipartimento di Scienze Chimiche, Università di Napoli Federico II, Napoli 80126, Italy
‡Carabinieri, Reparto Investigazioni Scientifiche (R.I.S.) di Roma, Viale di Tor di Quinto n. 151, Roma 00191, Italy
§Carabinieri, Sezione Investigazioni Scientifiche (S.I.S.) di Napoli, Corso Vittorio Emanuele n. 728, Napoli 80122, Italy

Ⓢ Supporting Information
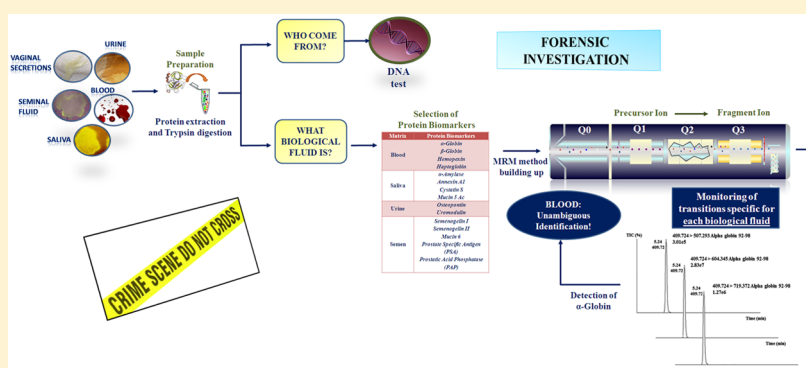
ABSTRACT: Knowledge of the nature of biofluids at a crime scene is just as important as DNA test to link the nature of the biofluid, the criminal act, and the dynamics of the crime. Identification of methods currently used for each biological fluid (blood, semen, saliva, urine) suffer from several limitations including instability of assayed biomolecules, and low selectivity and specificity; as an example of the latter issue, it is not possible to discriminate between alpha-amylase 1 (present in saliva) and alpha-amylase 2 (present in semen and vaginal secretion. In this context, the aim of the work has been to provide a predictive protein signature characteristic of each biofluid by the recognition of specific peptides unique for each protein in a single analysis. A panel of four protein biomarkers for blood, four for saliva, five for semen, and two for urine has been monitored has been monitored by using a single multiple reaction monitoring (MRM)-based method targeting concomitantly 46 different peptides. Then, The optimized method allows four biological matrices to be identified when present on their own or in 50:50 mixture with another biofluid. Finally, a valid strategy combining both DNA analysis and liquid chromatographic-tandem mass spectrometric multiple reaction monitoring (LC-MS-MRM) identification of biofluids on the same sample has been demonstrated to be particularly effective in forensic investigation of real trace evidence collected at a crime scene.

In the last three decades the link between forensic investigations and scientific disciplines, such as Chemistry and Biology, has been enormously strengthened by the application of new tools for a detailed characterization of the crime scene starting from trace evidence. In particular, DNA testing is fundamental to identify the individuals involved in a crime often leading the investigators to define a unique genetic profile.[1]

However, the detection and identification of the type and origin of biofluids at a crime scene are just as important as DNA test to link the nature of the biofluid, the criminal act and the dynamics of the crime. For example, blood stains can indicate some form of physical altercation, whereas detection of semen or vaginal fluid can indicate the involvement of some form of sexual encounter or assault. Moreover, identification of the specific fluid or tissue from which the DNA sample was recovered is fundamental in correctly reconstructing the criminal event and the effective role of the donor.

The most common biofluids found at crime scenes are blood, semen, and saliva, but others such as vaginal fluid, urine, and sweat can also be found. Each of these fluids has one or more presumptive tests that are initially used to give some indication as to the identity of the substance; blood is quickly detected by means of luminol or benzidine,[2] whereas traces of semen, saliva, sweat, urine, and other biological fluids are detected by ultraviolet light or by other light sources at specific wavelengths.[3] Confirmatory tests, based on biochemical, spectroscopic, and

microscopy methods, and immunochemical techniques, are necessary to give a legal value to the trace evidence.[4−9]

However, these identification methods suffer from several limitations including instability of assayed biomolecules and low selectivity and specificity. As an example of the latter issue, it is not possible to discriminate between α-amylase 1 (present in saliva) and 2 (present in semen and vaginal secretion).[10] Moreover, biochemical assays are specific only for one biological matrix and several cascade tests might then be needed before the biological nature of a certain stain is uncovered. This is even more challenging when the sample is a mixture of different biological matrices thus greatly increasing the complexity, the cost and the time of the analytical procedures. Finally, some confirmatory tests very often cause sample loss and they are not compatible with downstream individual identification by DNA analysis.

For all these issues, forensic science is looking for a universal confirmatory test for the analysis of unknown stains which will be able to unambiguously identify the type and nature of any biofluids that might be present at a crime scene. The method should be applicable to mixtures and, more importantly, should preserve the samples for subsequent DNA analysis. Since biofluids have evolved to perform different functions, they contain different proteins, or different combinations of proteins, providing each biological matrix with a unique protein signature that can be used to distinguish among the various biofluids.[11−13] Recently, untargeted proteomics have been introduced for the determination of biological matrices in forensic science based on the identification of the most prominent proteins present in biofluids.[11,14] However, these methods can give poor or uncertain results when high amounts of non specific proteins occur in the samples, i.e. when low amounts of a particular fluid are present in combination with high amounts of other matrices or when the sample is contaminated by other tissues.

This Article reports a universal method alternative and complementary to "traditional" tests currently used in forensic investigation to identify an unknown stain of biofluids occurring at a crime scene. The method is based on a targeted proteomic approach that makes use of tandem mass spectrometry in multiple reaction monitoring (MRM) mode to selectively monitor a number of specific peptides belonging to proteins biomarkers of individual biological fluids. First, the most prominent and highly specific protein biomarkers for each biological matrix were identified by both conventional LC-MS/MS analyses and comparison with literature data. Specific peptide markers of each protein were then selected on the basis of their mass spectrometric behavior together with their specific precursor ion-product ion transitions as defined by their unique amino acid sequence. Then, a single MRM method was devised to detect the occurrence of the target peptides within the sample leading to the unambiguous discrimination among the different biological matrices in a single analysis because of its high sensitivity, selectivity and accuracy. The optimized method was tested and validated on specimens consisting of four biological matrices (blood, saliva, semen, and urine), or mixtures of matrices, spotted on different substrates like cloth, wood, plastic, plaster, and paper. Finally, a combined strategy allowing both DNA analysis and biofluids identification on the same sample was developed and demonstrated to be effective in forensic investigation by the analysis of real specimens collected at a crime scene.

MATERIALS AND METHODS

Control biofluids (blood, saliva, urine, semen) were provided by a private chemical laboratory. Guanidine, Tris(hydroxymethyl) aminomethane hydrochloride, dithiothreitol (DTT), ethylendiaminetetraacetate (EDTA), trypsin, iodoacetamide (IAM), ammonium bicarbonate (AMBIC), trichloroacetic acid (TCA) were purchased from Sigma-Aldrich. Bovine serum albumin (BSA), formic acid (HCOOH), methanol, chloroform, and acetonitrile (ACN) are from J.T. Baker. Bradford staining solution was from Bio-Rad. Pipette tips C18 (zip tip) and centrifugal filter units were purchased from Merck Millipore.

In-Solution Digestion of Proteins from Biological Matrices. Different protocols were performed to improve the number of identified proteins and thus to provide the higher sequence coverage for blood, urine, semen, and saliva. A protocol of protein precipitation by using chloroform/methanol/water[15] was carried out for each biofluids. After precipitation, protein concentration was determined by Bradford assay using BSA as standard.[16] Samples were submitted to reduction, alkylation and tryptic digestion. Samples were dissolved in denaturant buffer (urea 6 M, Tris 300 mM pH 8.0, EDTA 10 mM) containing DTT (10-fold molar excess on the Cys residues) at 37 °C for 2 h and then iodoacetamide (IAM) was added to perform carboamidomethylation using an excess of alkylating agent (5-fold molar excess on thiol residues). The mixture was then incubated in the dark at room temperature for 30 min. The product was purified by chloroform/methanol/water precipitation. Supernatants were removed and the pellets were dried. Digestion of proteins mixture was carried out in AMBIC 10 mM using trypsin at a 50:1 protein/enzyme mass ratio. The samples were incubated at 37 °C for 16 h and after acidification (10% HCOOH) they were dried. To eliminate any impurities the samples were suspended in 200 μL of AMBIC 100 mM, filtrated by centrifugal filter units (0.22 μm), and dried in a speed-vacuum concentrator. Finally, samples were suspended in 20 μL of HCOOH 1% and purified by reverse phase chromatography using ZipTip C18 cartridges (Millipore). Samples were evaporated and suspended in 10 μL of HCOOH 0.1% and analyzed by nanoLC-MS/MS and then by LC-MRM/MS as described below.

To improve the number of identified proteins from blood by LC-MS/MS, a step of depletion of abundant proteins was carried out using ProteoPrepR 20 Plasma Immunodepletion KitProteoprep 20 Plasma Immunodepletion Kit (Sigma-Aldrich, Milan, Italy) able to deplete 20 highly abundant proteins from human plasma or serum. Dried extracted proteins were subjected to reduction, carboamidomethylation, chloroform/methanol/water precipitation protocol and tryptic hydrolysis as described above.

Protein precipitation from urine was also performed using cold acetone followed by the same steps of tryptic digestion as described above. Cold acetone at −20 °C (400 μL) was added to urine (100 μL) and incubated at −20 °C for 2 h. After centrifugation at 12 000 rpm for 15 min the pellet was collected and dried under vacuum.

Protein precipitation from semen was also carried out using TCA. Semen was centrifuged (13 000 rpm, 30 min) to remove cellular material, and the proteins were precipitated (4 °C, 60 min) using an equal volume of cold TCA at final concentration of 10%. Following further centrifugation (10 000 rpm, 30 min), the supernatant was removed and the protein pellet washed with cold acetone (three times) to remove residual TCA.

Therefore, a unique precipitation/sample preparation procedure based on chloroform/methanol/water was adopted for all the samples because the origin or matrix of casework samples.

Test Specimens. Twenty test forensic samples were prepared by drying different biological fluids (100 μL) or mixture of them on various substrates: cloth, wood, plastic,

plaster and paper. Cloths were put directly in a plastic tube while the samples on the other substrates were recovered by a cotton-swab (Agilent Technologies, Palo Alto, CA). Samples were collected in plastic tubes and 2 mL of AMBIC 50 mM were added for 16 h under gentle agitation. Samples were then sonicated for 20 min and submitted to chloroform/methanol/water precipitation. After precipitation, reduction, alkylation and tryptic digestion were performed as previously described. The resulting peptide mixtures were analyzed by LC-MRM/MS as described below.

Analysis of Real Samples. Real samples consisting of traces of unknown biological fluids on cloths, paper and stubs collected at a crime scene (samples 1, 2, and 3) were provided by RIS Laboratory in Rome. Samples were treated as described above (see Test Specimens section) and submitted to tryptic digestion. Then, gDNA was extracted from the samples using the robotic platform Qiagen Biorobot EZ1 Advanced XL using the Qiagen EZ1 DNA Investigator kit.[17] The DNA was then washed and eluted in water. DNA degradation and quantification level was evaluated by multiplex RT-PCR using the Quantifiler Trio DNA Quantification Kit (Applied Biosystems, CA, USA)[18] on a 7500 Real-Time PCR System (Applied Biosystems). Results showed a total DNA amount of 0.081 ng/μL for sample 1, 0.71 ng/μL for sample 2, and 1.05 ng/μL for sample 3. Samples were also tested for possible DNA degradation showing a good quality with a degradation index of about 1.0 for all samples.

DNA profiling was obtained by multiple amplification of 17 polimorphic regions comprising the D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, TH01, D16S539, D2S1338, D19S433, ACTBP2 (SE33), D1S1656, D2S441, D10S1248, D12S391, D22S1045, and Amelogenin STR loci using the Investigator ES Splex SE Plus Kit Qiagen (Qiagen, Hilden, Germany).[18] Amplified DNA was analyzed by automated DNA sequencing on an Applied Biosystems 3500xL Genetic Analyzer.[19−21] Data were collected and elaborated using the 3500 Series Data Collection Software v. 2.0and the Gene Mapper ID-X Software v.1.4 (Life Technologies, Carlsbad, CA, USA).

LC-MS/MS Analysis. Peptide mixture were analyzed by LC-MS/MS on a 6520 Accurate-Mass Q-TOF LC/MS system (Agilent Technologies) equipped with a 1200 HPLC system and a chip cube (Agilent Technologies). After loading, the peptide mixture (1 μL) was concentrated and desalted at flow rate of 4 μL/min in a 40 nL enrichment column (Agilent Technologies chip) with 0.1% HCOOH as eluent. The sample was then fractionated on a C18 reverse phase capillary column (75 μm*43 mm in the Agilent Technologies chip) at flow rate of 400 nL/min, with a linear gradient of eluent B (0.1% HCOOH in 95% acetonitrile) in A (0.1% HCOOH in 2% acetonitrile) from 5% to 80% in 50 min. Peptides analysis was performed using data-dependent acquisition of one MS scan (mass range m/z 300−2400) followed by MS/MS scan of the five most abundant ions in each MS scan. MS/MS spectra were measured automatically when the MS signal was greater than the threshold of 50000 counts. Doubly- and triply- charged ions were preferably isolated and fragmented over singly charged ions. Data were acquired through Mass Hunter software (Agilent Technologies). The acquired data, containing MS and MS/MS spectra, were transformed in .mgf format and used for protein identification with a licensed version of Mascot Software (www.matrixscience.com).

Mascot search parameters included: NCBInr as database; trypsin as enzyme, allowed number of missed cleavage 3; *Homo-Sapiens* as taxonomy; carbamidomethyl, C as fixed modifications; oxidation of methionine (oxidation (M));, *Gln pyro-Glu*

(N-term Q) as variable modifications; 10 ppm MS tolerance, 0.6 Da MS/MS tolerance and peptide charge, from +2 to +3.

MRM Targeted Proteomic Approach. To build up a targeted MRM method, Skyline software (3.7, 64 bit version MacCoss Lab Software, University of Washington, USA) was used for the in silico selection of peptides with unique sequence for each selected protein. For each peptide, m/z precursor ion, m/z product ions and relative collision energy were provided by Skyline. Peptide mixture was analyzed by LC-MS/MS analysis using a Xevo TQ-S (Waters) equipped with an IonKey UPLC Microflow Source coupled to an UPLC Acquity System (Waters). For each run, 1 μL peptide mixture was injected and separated on a TS3 1.0 mm × 150 mm analytical RP column (Waters, Milford, MA, USA) at 45 °C with flow rate of 3 μL/min using 0.1% HCOOH in water (LC-MS grade) as eluent A and 0.1% HCOOH in ACN as eluent B. Peptides were eluted (starting 1 min after injection) with a linear gradient of eluent B in A from 7% to 95% in 55 min. The column was re-equilibrated at initial conditions for 4 min. The MRM mass spectrometric analyses were performed in positive ion mode using a MRM detection window of 0.5−1.6 min per peptide; the duty cycle was set to automatic and dwell times were minimal 5 ms. Cone voltage was set to 35 V.

■ RESULTS

Selection of Protein Biomarkers of Biological Matrices. Four biological matrices, blood, saliva, urine, and semen, representative of biofluids recovered at crime scenes, were analyzed to identify specific protein biomarkers candidates for each matrix. Different procedures for sample preparation were investigated (Material and Methods section). As an example, results obtained for blood with the only chloroform/methanol/water precipitation followed by tryptic hydrolysis, are shown in Table 1 where

Table 1. List of the Identified Proteins Following the Chloroform/Methanol/Water Precipitation Protocol Applied to Human Blood Sample

GI number	identification	MascotScore	number of peptides	sequence coverage (%)
gil3212456	chain A, crystal structure of human serum albumin	1186	35	24
gil 90108664	chain A, crystal structure of lipid-free human apolipoprotein A-I	602	18	22
gil2765421	immunoglobulin kappa heavy chain	192	4	34
gil28637	alpha-1 antitrypsin	188	5	5
gil386789	hemopexin precursor	178	7	10
gil28810	β-2-glycoprotein apolipoprotein H	160	6	1
gil4557871	serotransferrin precursor	146	6	5
gil38026	Zn-α2-glycoprotein	75	2	5
gil112910	α-2-HS-glycoprotein	71	2	4

GI number, the name of identified proteins, the number of peptides and the corresponding sequence coverage are reported.

The unsatisfactory number of identified proteins in the original assay severely limited the choice of specific protein biomarkers for blood candidates for the MRM method because the dominant abundance of unselective blood proteins. Thus, a different sample preparation procedure was then explored by a depletion step from the most abundant proteins to improve protein identification (Table 2). A significantly higher number of

Table 2. List of the Identified Proteins Following the Depletion Step Previous to the Chloroform/Methanol/Water Precipitation Protocol[a]

GI number	identification	MascotScore	number of peptides	sequence coverage (%)
gi|1431650	chain β of hemoglobin	156	6	20
gi|229751	chain α of hemoglobin	138	4	15
gi|3212456	chain A, crystal structure of human serum albumin	1995	48	34
gi| 194383506	serum transferrin	1012	34	29
gi|90108664	chain A, crystal structure of lipid-free human apolipoprotein A-I	562	22	27
gi|177870	α-2-macroglobulin	485	17	7
gi|13787109	α-1-antitrypsin	337	12	14

[a]The GI number, the number of peptides, the corresponding Mascot score, and sequence coverage are reported for human blood sample.

peptides were identified for each protein leading to a much larger sequence coverage.

Improved results were also obtained in saliva, semen, and urine analysis by using the chloroform/methanol/water precipitation protocol. Table S-1 reports the lists of the different proteins identified for the other matrices, including the GI number, the number of peptides and the Mascot score.

Proteins identified for each biofluids by LC-MS/MS were then compared with literature data[11] to eventually select the most specific protein biomarkers for each biological matrix. Following this procedure, proteins listed in Table S-2 were chosen as representative for each biological fluid.

Development of MRM Methods Specific for Each Biological Matrix and a Single MRM Method for All the Matrices. A number of unique peptides belonging to the target proteins characteristic of individual biological matrix were selected by in silico analysis using the Skyline software that provided the predicted best transitions and collision energy to generate maximal fragment intensities. The in silico data of the selected peptides were then compared with the experimental fragmentation spectra obtained by the LC-MS/MS analyses. Peptides defined by Skyline analysis and showing the best signal-to-noise ratio in the experimental fragmentation spectra were selected for developing MRM methods for each biological fluid. These methods contained all the precursor ion−product ions transitions and collision energy associated with the selected peptides from the defined target proteins constituting the protein signature of each individual biofluids in Table S-2.

As an example, two peptides from α-globin, three from β-globin and hemopexin, five from haptoglobin, and six from α-2-macroglobulinfor a total of 25 peptides and 106 transitions were selected to monitor the presence of blood and to build the MRM method for this biological matrix.

The total ion current chromatogram (A) and the MRM TIC analysis of the three peptides 132−143, 82−94, and 30−39 from β-globin (B) are shown in Figure 1. Panel C displays the MRM TIC chromatogram of the three transitions, m/z 409.72 to 507.29, m/z 409.72 to 604.34, and m/z 409.72 to 719.37, used to monitor the α-globin (92−98) peptide. The different transitions perfectly coeluted at a retention time of 5.24 min thus indicating that they belong to the same precursor ion.

Similar MRM methods were developed for the other biological matrices (Table S-2). For urine, as an example, Figure S-1 shows

the total ion current chromatogram (A) and the MRM TIC analysis of the two peptides (507−605) and (204−211) from uromodulin (B). Panel C displays the TIC chromatogram of the three transitions for the uromodulin (204−211) peptide. The different transitions perfectly coeluted at a retention time of 2.99 min thus indicating that they all originated from the same precursor ion.

Some of the MRM TIC chromatograms of the selected peptides for the specific detection of proteins from saliva and semen were reported in Supporting Information (Figures S-2 and S-3).

Once the optimized MRM methods were developed for each individual biological fluid, the next step involved the development of a single MRM method able to detect all the biological fluids in a single analysis. For each selected protein biomarker, all the selected peptide sequences, m/z precursor ion, m/z product ions and the optimized collision energy were tabulated in Table S-4.

As the mass spectrometer was able to handle a very large number of transitions per run, the single MRM method was built up by using a total of 46 peptides and 212 transitions.

Analysis of Test Specimens. For all the test specimen samples, the MRM chromatograms solely showed the mass transitions associated with the target peptides belonging to the specific proteins constituting the unique signature of the defined matrix. No transitions related to peptides from proteins characteristic of other biological fluids were detected for all the analyzed samples. As an example, Figure 2A shows the MRM TIC chromatogram of the three transitions, m/z 637.86 to 687.35, m/z 637.86 to 850.42, m/z 637.86 to 949.49, m/z 637.86 to 1048.56, and m/z 637.86 to 1161.64 used to monitor the β-globin (30−39) peptide form a blood sample spotted on plaster. Panel B showed the corresponding MRM chromatogram for the (1075−1086) peptide from mucin-6, a protein specific of semen, illustrating the absence of any nonspecific transition of interfering proteins. This result clearly demonstrates that the specific signature of blood could be easily monitored whereas no interference from cross transitions of peptides belonging to other biological matrices was detected.

Similar results were obtained on all the other spotted samples and are summarized in Table 3. Each biological matrix could be unambiguously identified by the developed MRM method even on different substrates. No cross transitions were observed whatsoever with the exception of urine where peptides from PSA and PAP proteins, a specific signature of semen, was also detected. However, the presence of seminal proteins in urine is somehow expected and provides further information on the origin of the sample that was obviously from a male donor.

To test the MRM optimized method capability to unambiguously define the occurrence of individual biofluids when occurring in a 50:50 ratio mixture, blood−semen, blood−urine, and semen−urine were spotted onto different substrates. The negative controls, for a total of 5 sample, consisted of the exclusive use of substrates: cloth, wood, plastic, plaster, and paper. As shown in Table S-5, each individual component of the binary mixture was unambiguously identified by the specific transitions of the unique peptide signature; whenever urine was present, specific peptides from semen proteins were also detected.

Analysis of Real Crime Scene Samples. The developed procedure was then applied to real samples collected at a crime scene by the RIS Laboratory in Rome. Samples were prepared according to the procedure described above and following enzymatic hydrolysis, DNA was extracted from each sample according to the normal procedure performed at the RIS
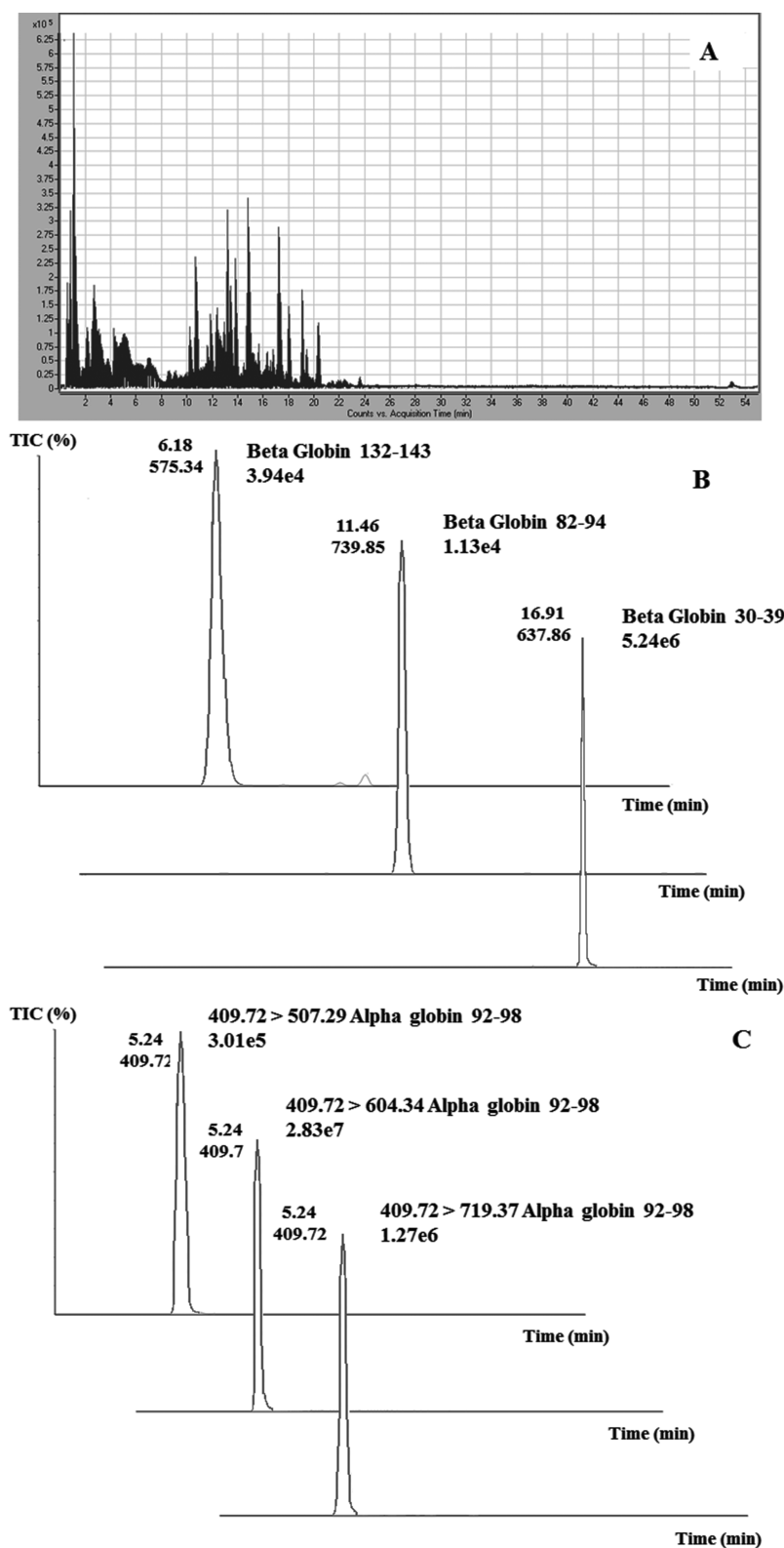
Figure 1. Identification of all the protein of blood fluid by LC-MS/MS analysis and MRM/MS data of the defined target proteins constituting the protein signature of blood. Panel A: Total ion current (TIC) as a function of time (min.). LC-MS-MS analysis allowed the identification of protein content on blood tryptic peptide mixture. Panel B: BPI Chromatograms for β-globin 132–143, 82–94, 30–39 peptides eluted at 6.18, 11.46, and 16.91 min, respectively. Panel C: MRM TIC Chromatograms extracted for α-globin (92–98) peptide. The monitored transitions for this peptide: m/z 409.72 to 719.37, m/z 409.72 to 604.34, and m/z 409.72 to 507.29 are perfectly coeluted at 5.24 min as an unambiguous identification of blood detection.

Laboratory. DNA extracts were quantified, the degradation level was evaluated and the DNA profiling obtained as described in Materials and Methods section. The resulting genotypic profile

of sample 1 is shown in Figure S-4 and was identified as a male genotype profile called "Profile A". The quality of the data exceeded the minimum quality requirements established by
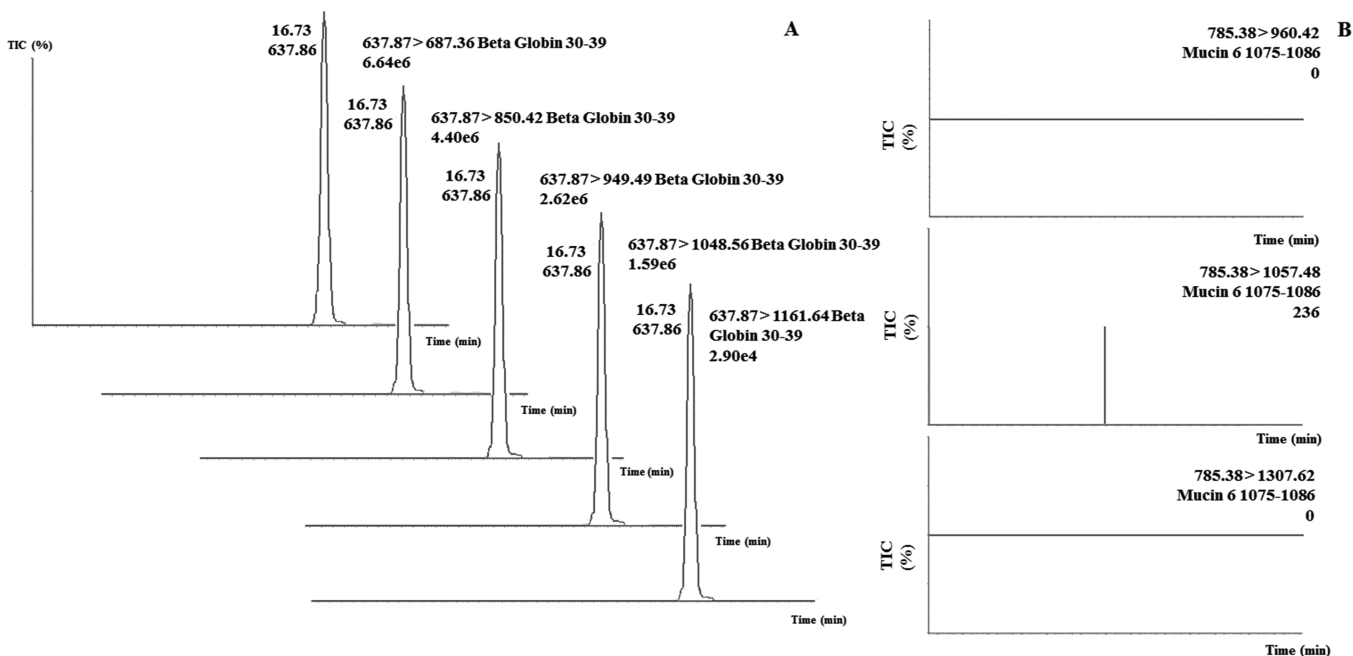
Figure 2. MRM/MS analysis of test specimens: blood spotted on plaster. Panel A: MRM TIC Chromatograms extracted for β-globin (30−39) peptide. The best five monitored transitions m/z 637.87 to 687.36, m/z 637.87 to 850.42, m/z 637.87 to 949.49, m/z 637.87 to 1048.56, and m/z 637.87 to 1161.64 are reported in figure eluted at 16.73 min. Panel B: MRM TIC chromatogram extracted for (1075−1086) peptide from mucin-6 a protein specific of semen, illustrating the absence of any nonspecific transition (m/z 785.38 to 960.42, m/z 785.38 to 1057.48, m/z 785.38 to 1307.62).

internal validation procedures for this Lab for comparative procedures to be used for personal identification purposes.

Following DNA analysis, the remainder of the samples containing the mixture of tryptic peptides was submitted to the developed MRM analytical procedure. Figure 3 shows the TIC analysis of the three peptides 30−39, 82−94, and 132−143 monitored for the human β-globin and two peptides 92−98 and 127−138 monitored for the human α-globin (A). Panel B displays the MRM TIC chromatogram corresponding to four transitions, m/z 637.87 to 687.35, m/z 637.87 to 850.42, m/z 637.87 to 949.48, and m/z 637.87 to 1048.55, used to monitor the β-globin (30−39) peptide. As shown in Table 4, the MRM analysis was able to unambiguously define individual protein component occurring in the different samples by their unique peptide signature even after DNA extraction and analysis. Therefore, this combined procedure led to both DNA profiling analysis and the unambiguous identification of blood indicating the biological fluid the DNA was collected from.

■ DISCUSSION

Knowledge of the nature of biofluid discovered at a crime scene can influence the outcome of a case. Identification of a suspect's DNA on a victim is quite different if it comes from saliva, suggesting intimate contact, wither consensual or forced, or from blood or seminal fluid, indicating a physical struggle or a sexual assault. However, this is not always an easy task, since many biofluid stains are either invisible to the naked eye or similar in appearance to other fluids or substances. Even when the identity of a stain may seem obvious to a forensic investigator, absolute confirmation is necessary to give a legal value to the trace evidence to either prove or disprove a fact in a lawsuit. This is especially important with the possible occurrence of mixtures as single stain could contain multiple biofluids.

Most biochemical and immunologic tests which have been used for presumptive or confirmatory methods suffer from several limitations leading to sample destruction resulting in sample loss for subsequent DNA analysis or to incompatibility with downstream individual identification assays, that is, DNA profiling. Moreover, most of the current methods are designed to detect a single biofluid forcing the investigators to decide which test to perform in the presence of a limited amount of sample.

Contrary to DNA, proteins were rarely considered and scarcely used as sources of useful biological traces in crime scene investigations. Proteins tend to be less stable than DNA and are easily degraded making their identification by immunological methods unfeasible. Moreover, the amount of available samples cannot be amplified by "PCR-like" procedures thus requiring the use of analytical techniques with extremely high sensitivity.[22]

More recently, proteomics strategies have been applied to the identification of biofluids that overcome most of the previous difficulties. In proteomic approaches, proteins are identified by their peptide fragments making the degradation no longer a problem. Moreover, contrary to DNA, proteins are tissue-specific providing a unique signature to identify biological tissues and fluids. Finally, the tiny amount of trace evidence usually recovered at a crime scene is matched by the extraordinary sensitivity of modern mass spectrometers. Untargeted proteomic approaches demonstrated to be effective in the definition of biofluids in traces from crime scenes by identification of specific proteins belonging to individual biological sample.[5] However, severe limitations in sensitivity, in the analysis of mixed traces and, more important, in preserving the samples for subsequent DNA analysis still exist. A recent study has greatly overcome this drawback by MALD IMS imaging directly on a blood fingermark leaving the ridge detail completely preserved.[10] Despite the fact that the proposed method is as destructive as the other approaches, the great compatibility between DNA extraction and MS strategy protocols allows to limit the destruction of the entire trace evidence. Indeed, a combined strategy to provide both DNA analysis profiling by RIS laboratory and identification of the

Table 3. List of Proteins Detected in Each Test Specimen Prepared by Drying Different Biological Fluids on Various Substrates: Cloth, Wood, Plastic, Plaster, and Paper[a]

matrixes	substrates	blood				urine		saliva				semen				
		α hemoglobin	β hemoglobin	hemopexin	haptoglobin	uromodulin	osteopontin	α amylase	Muc 5 Ac	annexin A1	custatin S	prostate specific antigen	prostatic acid phosphatase	Muc 6	semenogelin I	semenogelin II
blood	cloth	X	X	X	X											
	wood	X	X	X	X											
	plastic	X	X	X	X											
	plaster	X	X	X	X											
	paper	X	X	X	X											
urine	wood					X	X									
	plastic					X	X									
saliva	cloth							X	X	X	X					
	wood							X	X	X	X					
	plastic							X	X	X	X					
	plaster							X	X	X	X					
semen	cloth											X	X	X	X	
	wood											X	X	X		
	plastic											X	X	X	X	

[a]Matrix spotted on substrates. X is reported when the unique peptides belonging to the target proteins characteristic of individual biological matrix were identified by MRM analysis.
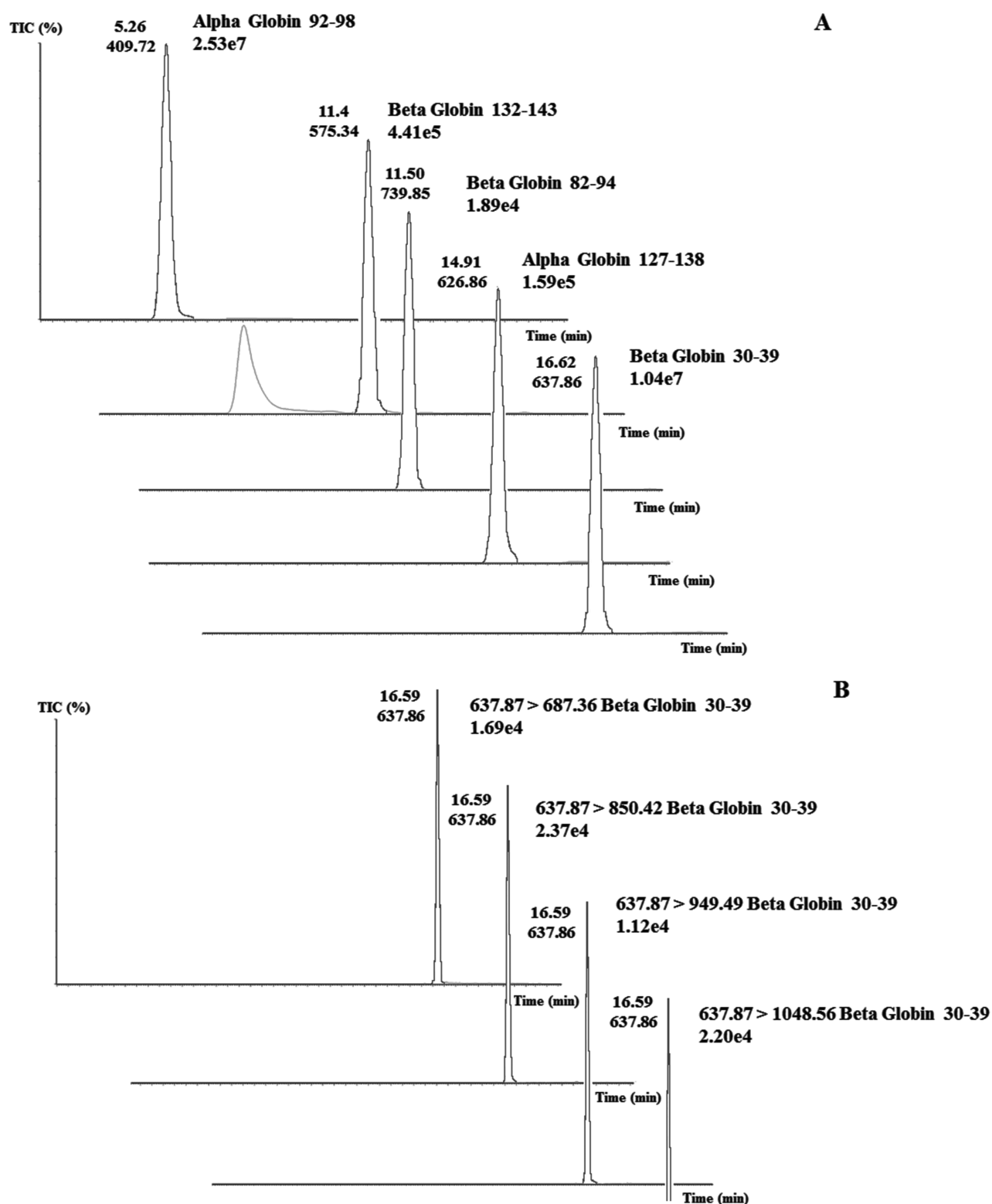
Figure 3. MRM/MS analysis of real crime scene samples after DNA extraction. Panel A: BPI Chromatograms for α-globin 92−98 and 127−138 peptides eluted at 5.26 and 14.91 min respectively and BPI Chromatograms for β-globin 132−143, 82−94, AND 30−39 peptides eluted at 11.40, 11.50, AND 16.62 min, respectively. Panel B: MRM TIC chromatogram corresponding to four transitions, *m/z* 637.87 to 687.35, *m/z* 637.87 to 850.42, *m/z* 637.87 to 949.48, and *m/z* 637.87 to 1048.55 perfectly coeluted at 16.59 min, used to monitor the β-globin (30−39) peptide.

corresponding biological fluids on the same sample collected at real crime scenes has been developed. Samples were deproteinized by treatment with trypsin instead of the commonly used nonspecific proteases, that is, Pronase, proteinase K, etc. DNA was then extracted from the digested samples and analyzed

according to the usual procedure, yielding optimal DNA profiles. Moreover, the proposed *bottom-up* approach makes this method valid also on old traces that have undergone degradation. Another advantage from the current approach is the unambiguous identi-fication of all target proteins allowing us to unequivocally

Table 4. Identification of the Biomarker Proteins by MRM Analysis in Three Samples Obtained from RIS Laboratories after DNA Extraction and Analysis[a]

		peptide	sample 1	sample 2	sample 3
blood	α Hb	K.VGAHAGEYGAEALER.M [16, 30]	x		
		R.VDPVNFK.L [92, 98]	x		
		K.FLASVSTVLTSK.Y [127, 138]	x		x
	β Hb	R.LLVVYPWTQR.F [30, 39]	x		x
		K.GTFATLSELHCDK.L [82, 94]	x		
		K.VVAGVANALAHK.Y [132, 143]	x		
	hemopexin	R.ELISER.W [82, 87]	x		
		K.VDGALCMEK.S [401, 409]	x		
	haptoglobin	K.DIAPTLTLYVGK.K [156, 167]	x		
		R.VMPICLPSK.D [202, 210]	x		
saliva	α amylase	R.WVDIALECER.Y [35, 44]			x
		K.SSDYFGNGR.V [258, 266]			x
	annexin 1	K.GVDEATIIDILTK.R [58, 70]			x
		R.SEIDMNDIK.A [303, 311]			x
		K.ILVALCGGN.- [337, 345]			x
	MUC 5AC	K.GVQLSDWR.D [1335, 1342]			x
		R.AQAQPGVPLR.E [2999, 3008]			x
	cystatin S	K.QLCSFEIYEVPWEDR.M [115, 129]			
		R.MSLVNSR.C [130, 136]			x
semen	MUC 6	K.VTNEFVSEEGK.F [182, 192]		x	
		R.ETDPCSMSQLNK.V [573, 584]		x	
		R.GVLLWGWR.S [646, 653]		x	
		K.VYHLPYYEACVR.D [1075, 1086]		x	
	SEM II	K.GHYQNVVDVR.E [217, 226]		x	
		K.DIFTTQDELLVYNK.N [251, 264]			
		K.ISYQSSSTEER.H [345, 355]		x	
		K.QDLLSHEQK.G [534, 542]		x	
	PSA	R.IVGGWECEK.H [20, 28]		x	
		K.HSQPWQVLVASR.G [29, 40]		x	
		R.LSEPAELTDAVK.V [121, 132]		x	
		K.FMLCAGR.W [190, 196]		x	
	PAP	R.SPIDTFPTDPIK.E [47, 58]		x	
		K.DFIATLGK.L [185, 192]		x	
		R.ELSELSLLSLYGIHK.Q [236, 250]		x	
	SEM 1	R.LWVHGLSK.E [165, 172]		x	
		K.VQTSLCPAHQDK.L [233, 244]		x	
		K.DVSQSSIYSQTEEK.A [307, 320]			
		K.GESGQSTNR.E [404, 412]		x	
urine	osteopontin	K.QNLLAPQTLPSK.S [51, 62]			
		R.GDSVVYGLR.S [145, 153]			
		R.ISHELDSASSEV.- [287, 298]			
	uromodulin	R.STEYGEGYACDTDLR.G [185, 199]			
		K.VFMYLSDSR.C [356, 364]			
		K.INFACSYPLDMK.V [420, 431]			
		R.VGGTGMFTVR.M [449, 458]			
		R.VLNLGPITR.K [597, 605]			

[a]X is reported when the unique peptides belonging to the target proteins characteristic of individual biological matrix were identified by MRM analysis.

discriminate among the different fluids, even in mixture by a single MRM run of less than 1 h.

Indeed, four proteins, α-globin, β-globin, haptoglobin, and hemopexin have been selected for blood signature, four for saliva (α-amylase 1, mucin-5, annexin A1, and cystatin S), five for semen (semenogelin I and II, PSA, PAP, and mucin-6) and two proteins determined for urine (uromodulin and osteopontin). For each protein biomarker, a number of unique tryptic peptides (proteotypics) has been selected with the use of bioinformatics tools and the best precursor ion-fragments transitions employed to set up a single MRM method able to identify any of the four

biofluids in unknown stains. The MRM LC-MS/MS analysis recognizes the specific peptides by their unique transitions. The identified proteotypic peptides have been correlated to each set of selected protein biomarkers that in turn unambiguously defined the biological fluids under investigation.

The MRM method optimized on pure samples of the four matrices, blood, urine, semen and saliva, is resulted to have good specificity and selectivity for each biofluid with no cross-contamination observed whatsoever. A single contamination was detected in the urine samples spotted on different surfaces that showed the occurrence of small amount of seminal fluid.

However, this presence was somehow expected and contributed to indicate a male donor. Then in a proof of concept, prepared test samples consisting of different fluids spotted on various substrates have been analyzed by using the single optimized MRM method.

The current results demonstrated that the developed strategy based on the MRM LC-MS/MS method could be a useful substrate in forensic science due to its capability of providing a universal approach for the identification of unknown stains recovered at crime scenes with high selectivity and specificity. Moreover, this approach can be carried out on the same samples used for DNA profiling revealing the nature of the tissue or fluid the DNA had been recovered from.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.7b04742.

List of proteins identified for each biological matrix by different pretreatment by LC-MS/MS analysis; list of proteins selected as representative for each biological matrix by matching LC-MS/MS analysis and literature data; MRM method reporting selected peptide sequence, m/z precursor, m/z product ions and collision energy (V) optimized for a single biological fluid and the unique MRM method able to detect the type of fluid from unknown biological evidence on the crime scene; specificity test of MRM method when biological fluids were mixed on cloth; total ion current chromatogram from LC-MSMS analysis and the MRM chromatograms analysis of recorded transition for selected peptides for highly specific and selective identification of each biological fluid; and DNA profiling obtained by multiple amplification of 17 polymorphic regions (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: anna.illiano@unina.it.

ORCID Ⓞ

Anna Illiano: 0000-0003-1491-8966

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Roewer, L. *Invest. Genet.* **2013**, *4*, 22.

(2) Tobe, S. S.; Watson, N.; Daeid, N. N. *J. Forensic Sci.* **2007**, *52*, 102−109.

(3) Auvdel, M. J. *J. Forensic Sci.* **1987**, *32*, 326−345.

(4) Webb, J. L.; Creamer, J. I.; Quickenden, T. I. *Luminescence* **2006**, *21*, 214−220.

(5) Virkler, K.; Lednev, I. K. *Forensic Sci. Int.* **2009**, *188*, 1−17.

(6) Cocks, J.; du Toit-Prinsloo, L.; Steffens, F.; Saayman, G. *Forensic Sci. Int.* **2015**, *249*, 225−232.

(7) Horjan, I.; Barbaric, L.; Mrsic, G. *Journal of forensic and legal medicine* **2016**, *38*, 101−105.

(8) Muro, C. K.; Doty, K. C.; de Souza Fernandes, L.; Lednev, I. K. *Forensic Chemistry* **2016**, *1*, 31−38.

(9) Muehlethaler, C.; Leona, M.; Lombardi, J. R. *Anal. Chem.* **2016**, *88*, 152−169.

(10) Pang, B.; Cheung, B. K. *J. Forensic Sci.* **2008**, *53*, 1117−1122.

(11) Van Steendam, K.; De Ceuleneer, M.; Dhaenens, M.; Van Hoofstat, D.; Deforce, D. *International Journal of Legal Medicine* **2013**, *127*, 287−298.

(12) Legg, K. M.; Powell, R.; Reisdorph, N.; Reisdorph, R.; Danielson, P. B. *Electrophoresis* **2017**, *38*, 833−845.

(13) Deininger, L.; Patel, E.; Clench, M. R.; Sears, V.; Sammon, C.; Francese, S. *Proteomics* **2016**, *16*, 1707−1717.

(14) Legg, K. M.; Powell, R.; Reisdorph, N.; Reisdorph, R.; Danielson, P. B. *Electrophoresis* **2014**, *35*, 3069−3078.

(15) Wessel, D.; Flügge, U. *Anal. Biochem.* **1984**, *138*, 141−143.

(16) Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248−254.

(17) Anslinger, K.; Bayer, B.; Rolf, B.; Keil, W.; Eisenmenger, W. *Leg. Med.* **2005**, *7*, 164−168.

(18) Liu, J. Y. *Forensic Sci. Int.: Genet.* **2014**, *13*, 10−19.

(19) Pascali, J. P.; Bortolotti, F.; Tagliaro, F. *Electrophoresis* **2012**, *33*, 117−126.

(20) Moreno, L. I.; McCord, B. *Handbook of Capillary and Microchip Electrophoresis and Associated Microtechniques*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2007, 733−756.

(21) Landers, J. P. *Handbook of Capillary and Microchip Electrophoresis and Associated Microtechniques*; CRC Press, 2007.

(22) Sethi, S.; Brietzke, E. *Int. J. Neuropsychopharmacol.* **2016**, *19*, pyv096.
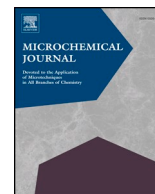
Identification of proteinaceous binders in paintings: A targeted proteomic approach for cultural heritage

Roberto Vinciguerra[a,1], Anna Illiano[a,1], Addolorata De Chiaro[a], Andrea Carpentieri[a], Anna Lluveras-Tenorio[b], Ilaria Bonaduce[b], Gennaro Marino[a,c], Piero Pucci[a], Angela Amoresano[a,*], Leila Birolo[a,d,*]

[a] Department of Chemical Sciences, University of Naples, Federico II, Complesso Universitario di Monte S. Angelo, 80126 Naples, Italy
[b] Department of Chemistry and Industrial Chemistry, University di Pisa, 56124 Pisa, Italy
[c] University "Suor Orsola Benincasa" Naples, Italy
[d] Task Force "Metodologie Analitiche per la Salvaguardia dei Beni Culturali", University of Naples Federico II, I-80126 Naples, Italy

A B S T R A C T

Identification of proteins in paintings and polychrome objects is a challenge, which requires the development of tailored analytical approaches. In the present study, a targeted proteomics approach was developed for discriminating among the three most common proteinaceous materials used as paint binders, i.e. milk, egg, and animal glue. In this study a specific database of peptides was created based on tandem MS analyses of tryptic digests of several paint samples collected from a variety of art objects of different ages and conservation conditions. Specific peptide markers of each protein were then selected and monitored by LC-MSMS in Multiple Reaction Monitoring (MRM) ion mode, together with their specific precursor ion-product ion transitions, as defined by their unique amino acid sequence. The developed method enabled a sensitive and reliable detection of the target peptides in a selection of case studies, leading to the unambiguous identification of the proteins used as paint binders. The method showed greatly increased sensitivity compared to currently available strategies.

1. Introduction

The inherently multidisciplinary nature of proteomics has recently led to its application to assorted areas that include the field of cultural heritage [1] for the identification of protein-based materials in artworks and polychromies, archaeological remains, and paleontological objects [2].

The most up-to-date approaches for the characterization of proteinaceous materials in artworks and polychromies, and in general when dealing with ancient proteins, rely on mass spectrometry technologies that are able to identify proteins from the MS analysis of their digested peptides [2,3] (and references therein). Typically, MS-based proteomics pipelines aimed at identifying proteins in ancient samples follow the classical bottom up approach, where samples are enzymatically digested to peptides, which are either analyzed by simple MS, or separated by liquid chromatography and eventually analyzed by tandem mass spectrometry. Peptides and thereof proteins are identified by

correlating experimental spectra to virtual ones in a protein sequence database [4]. In the classical full scan operating mode (typical of LC-MSMS analysis in discovery proteomics or shotgun proteomics, otherwise named global profiling proteomics or untargeted proteomics), the mass spectrometer continuously repeats the full scan mass spectrum and selects and fragments the n (typically $n = 1$–10) most abundant ions (information-dependent acquisition, IDA).

So far, the applications to the cultural heritage field are untargeted, discovery proteomics based experiments. The fascinating potential of this operating mode lies in the possibility of uncovering proteins without any "a priori" knowledge or hypothesis. It is therefore compatible with both well and poorly characterized systems in cultural heritage, the only basic requirements being the actual presence of proteins in the sample and the presence of the protein sequences or homologous sequences in a properly selected database [2,3].

Since the first report of LC-MSMS in 2006 [5], despite the tremendous success of these approaches in a wide range of works of art

* Corresponding authors at: Department of Chemical Sciences, University of Naples Federico II, Complesso Universitario di Monte S. Angelo, Via Cinthia, 80126 Naples, Italy.

E-mail addresses: angamor@unina.it (A. Amoresano), birolo@unina.it (L. Birolo).

[1] These authors equally contributed to the paper.

and archaeological remains [5–24] standard based proteomics strategies keep being developed to address the field-specific analytical challenges in samples from cultural heritage, by improving the steps in the standard proteomics protocols, from the solubilization step on [20,21,25–28]. So far, the greatest attention was addressed to the preparation of the sample, correctly considering the unusual nature/state of the sample itself as the biggest challenge in the analysis.

However, as far as the mass spectrometric analysis itself is concerned, some of the limits in the detection in a LC-MSMS analysis can also arise from the automated peak selection and the linked instrument's bias towards repeatedly selecting and fragmenting the species with the most intense MS signals. The complexity of these unusual samples in works of art and archaeological remains, in terms of chemical composition, the unavoidable presence of contaminating protein-based materials (even coming from the environmental dust), as well as the high levels of molecular damage found in ancient samples, can overwhelm even the most modern mass spectrometer [3,21,25], sometime making the identification of protein components in these samples extremely challenging. This because the mass spectrometer might miss signals coming from important, relevant proteins, but instead might "waste its time" on signals from uninformative contaminants, for instance. Searching for significative signals in a multitude of untargeted signals might then become like looking for a needle in a haystack with the naked eye.

There are cases, however, such as the identification of the protein-based binders in paintings, where, in principle, the search could be basically simplified to a discrimination between the most commonly used materials, i.e., animal glues, egg, and milk casein. Therefore, the analysis can be "targeted" towards a search for the protein components of these materials, within a limited range of possibilities, in an alternative, focused approach that by selectively searching for a restricted set of marker proteins, can eventually guarantee a higher sensitivity and reduce the interference from contaminations as lowest as possible.

In these cases, a significant advance in the solution of the analytical challenge can be the multiple reaction monitoring (MRM) ion mode, an alternative mass spectrometric strategy, the actual gold standard in mass spectrometric technique for selectivity and high signal-to-noise ratio that selectively and non-redundantly search for specific ions of one or, more desirably, a few peptides that will unequivocally constitute protein markers to be identified as unique signatures for each binder. The mass spectrometer will therefore focus on selected signals, getting rid of the "noise" of the other ions from the background, thus sensibly increasing sensitivity. Quoting the commentary on Nature Methods on 2013 [29], MRM "can be thought of as the mass spectrometrist's ELISA. It targets proteins using a predetermined assay with high sensitivity and selectivity [30]: molecular ions of a target peptide are selected in the first mass analyzer, fragmented in the collision cell and one or several of the fragment ions uniquely derived from the target peptide are measured by the second analyzer. MRM is referred to as targeted approach as only predetermined ions are measured, and mass spectrometer acts as a mass filter selectively monitoring a specific analyte molecular ion and one or several fragment ions generated from the analyte. MRM has been first developed and applied for decades in the pharmaceutical industry to quantify small molecules [31] and evolved to be applied in the field of proteomics, to detect low abundant proteins in complex matrices, to verify biomarker candidates, and has proved to be successful in clinical settings [32–42].

Mass spectrometry-based strategies that search for specific peptides as markers in the analyses of protein-based binders have been reported [20,43–46]. However, in none of them, to the best of our knowledge, has ever been used to discriminate between collagen-, milk- and egg-based binders. We propose this technique as a probe to specifically identify peptide markers of proteinaceous materials in paintings. Herein, we developed and applied a method based on a targeted proteomic approach that makes use of tandem mass spectrometry in multiple reaction monitoring (MRM) ion mode to selectively monitor a number of specific peptides belonging to proteins markers of individual proteinaceous binders in paintings. This alternative strategy, by selectively and non-redundantly searching for specific signatures of the proteinaceous binders, could represent a significative advance in the solution of the analytical challenge set by ancient samples. Differently from previously reported methods [20,43,44,46] that also search for specific peptides but in untargeted, typical shotgun runs, where relative abundance of the single peptide ions can still strongly affect the results, MRM by its two-level mass filtering instrumental operating mode results in an increase of selectivity, and an incredibly high signal-to- noise ratio for the target analytes, getting rid of most of the nuisances arising from abundant contaminating proteins in the samples that can unavoidably affect, for instance, mass fingerprinting methods but also standard LC-MSMS approach.

In this work, the most prominent and specific protein markers for protein-based binder were chosen by screening a collection of LC-MSMS analyses of a panel of paintings/test samples that had been carried out during the years. Samples had different origin, had been prepared with different procedures, with different pigments, in order to be as general as possible in the selection of the target peptides, and provide a versatile tool that could be applied to any painting sample. Specific peptide markers of each protein were then selected on the basis of their frequency of detection in the set of samples and of their mass spectrometric behavior together with their specific precursor ion-product ion transitions as defined by their unique amino acid sequence. Then, a single MRM method was constructed to detect all the target peptides in a single analysis with high sensitivity, selectivity and accuracy, leading to the unambiguous discrimination among the different proteinaceous binders. The method was tested on a panel of samples of known composition and then successfully applied to a set of paint samples of unknown composition.

Results presented here demonstrate the unique ability of MRM method to selectively assess the nature of proteinaceous binders in paintings, and its successful application to samples where standard LC-MSMS analysis had failed, or to disclose previously undetected components.

2. Materials and methods

2.1. Reagents

Ammonium hydrogen carbonate (AMBIC), Ethylenediaminetetraacetic acid (EDTA); Tri(hydroxymethyl)aminomethane (TRIS), and TPCK-treated trypsin were from Sigma; recombinant Peptide N-Glycosidase F (PNGaseF) was from Roche. Formic acid and Acetonitrile (ACN) were purchased from Baker. Deionized water was obtained from Millipore cartridge equipment.

2.2. Painting samples

The list and the characteristics of paint replicas and historical samples used to build the database of selected peptides are reported in Table S1. Historical samples were from collections that have been published elsewhere [26,47–50], and criteria for sampling and description of the single cases are therein given.

Samples were motley analyzed in the course of the years, and were generally treated in heterogeneous phase with trypsin with variations in respect to the minimally [21] invasive protocol as reported in Table S1, and the resulting peptide mixtures were analyzed by LC-MSMS, as detailed in the therein indicated references [21,25,26,47–50]. For those samples that were used in the setup of the database that have not been published previously, digestion with trypsin, LC-MSMS analysis and data analysis were carried out as described in detail in the Supplementary information.

In order to build up the database, the proteins and the peptides most frequently detected in standard global profiling experiments were screened regardless the treatment procedures they had undergone and

the mass spectrometer that was used in the specific analysis. A relational database was designed on PivotTables (https://support.office.com) to simplify the data processing.

2.3. LC-MSMS analysis in multiple reaction monitoring ion mode

In MRM mode, two steps of mass filtering are used on a triple quad mass spectrometer. The technique is more sensitive than full scan MSMS because both Q1 and Q3 are parked on a single *m/z* ion, with both mass analyzers dwelling on a single ion.

In a typical MRM workflow, the targets must be selected. MRM-based assay usually starts with the selection of a signal providing optimal signal intensity and discriminating the target peptide from the other species. The predicted best transitions and collision energy to generate the maximal fragment intensities are then provided by in silico analysis by using Skyline software [51]. These data are then used for MRM method construction. The detailed information of the herein used MRM method is reported as follows.

Peptide mixtures were analyzed by LC-MSMS analysis using a Xevo TQ-S (Waters) with an UPLC Microflow Source coupled to an UPLC Acquity System (Waters) using an IonKey (Advion) device. For each run, 1 μl peptide mixture was injected and separated on a TS3 1.0 mm × 150 mm analytical RP column (Waters, Milford, MA, USA) at 60 °C with flow rate of 3 μl/min. Peptides were eluted (starting 1 min after injection) with a linear gradient of eluent B (0.1% Formic acid in 100% ACN) in A (0.1% Formic acid in water (LC–MS grade)) from 7% to 95% in 55 min. The column was re-equilibrated at initial conditions for 4 min with eluent A. MRM mass spectrometry analyses were performed in positive ion mode. Skyline Software drew the choice for the selection of virtual best transitions and collision energy calculated to generate maximal fragmentation intensities for each analyte. The developed MRM method was used with MRM detection window set to 0.5–1.6 min per peptide, the duty cycle was set to automatic and dwell times were minimal 5 ms. Cone voltage was set to 35 V. The mass spectral MRM parameters are reported in Table S2.

3. Results

3.1. Database setup and selection of the binder specific proteins

Results from LC-MSMS analyses of a collection of 105 samples from artistic and archaeological samples (Table S1), which had showed to contain milk, animal glue and egg, were screened to identify specific protein biomarker candidates for each material. Milk proteins were found in 43 samples, animal glue in 29 samples and egg in 43 samples. Table S3 shows the details of the identifications for all the samples that were inserted in of the database that had not been published elsewhere.

Fig. 1 reports the summary of the results relative to the samples that contained milk. Whenever milk was detected, alpha S1 casein was always identified, while beta casein was identified in 32 samples out of 43, and alpha S2 casein observed in 30 samples (Details in Table S4). Similarly, in the animal glue containing samples, as expected, two collagen proteins were detected in almost all samples, collagen alpha-1(I) identified in all samples and collagen alpha-2(I) in 28 out of 29 samples (Fig. S1, and details in Table S5).

The case of egg binder was more complex since artists could alternatively use either albumen or yolk or whole egg. Moreover, pure albumen can be quite easily obtained from a whole egg whereas yolk is very likely contaminated by egg white. Therefore, the results obtained with the 43 egg containing samples were evaluated by dividing them into two subgroups of samples, i.e. those containing only albumen (27 samples) and those containing yolk plus albumen (16 samples). No sample was considered as a pure yolk. As expected (Fig. S2, and details in Table S6), chicken ovalbumin, the main protein of egg white, was identified in all samples containing only albumen (27 out of 27) and in 38 out of the total 43 samples including those containing also egg yolk.
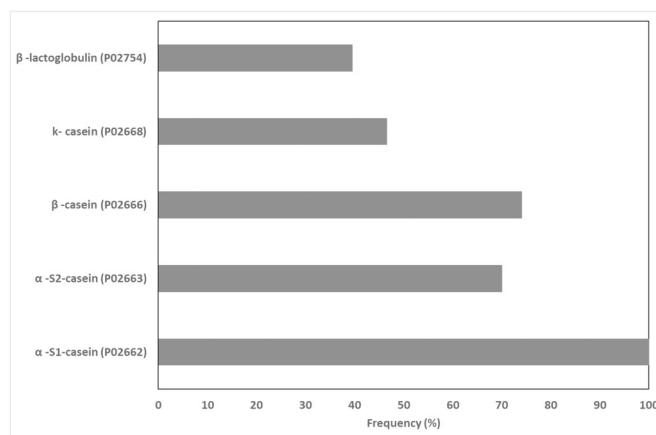

Fig. 1. Graphical representation of proteins occurrence in 43 milk containing samples. The occurrence has been reported as percentage of the analyzed samples. The UniProt code is reported in brackets. Details of the proteins occurrence are shown in Table S4.

Similarly, ovotransferrin was observed in 21 out of 27 albumen only samples and in 8 of the samples containing yolk plus albumen, thus accounting for 29 out of 43 total samples. Vitellogenin-1 and vitellogenin-2 were the most frequently identified proteins from yolk, observed in 12 and 14, out of 16 samples containing egg yolk.

From the results obtained, alpha S1 casein, alpha S2 casein and beta casein were selected for the subsequent selection of the peptides to be used for the identification of milk, collagen alpha-1(I) collagen alpha-2(I) for animal glue, ovalbumin and ovotranferrin as generic representative of egg containing binders and vitellogenin-1 and vitellogenin-2 specifically for yolk containing binders.

3.2. Selection of the target peptides

Unique peptides belonging to the target proteins characteristic of each proteinaceous binder were thus selected to develop the MRM assay. Specific peptides were chosen on the basis of their frequency of observation and quality of fragmentation spectra in the LC-MSMS analyses. As an example, Fig. 2 reports the frequency of observation of the alpha S1 casein peptides in the milk containing samples. The peptide YLGYLEQLLR encompassing the 106–115 region of the alpha S1 casein was observed in 77% of cases while the peptide FFVAPFPEVFGK, encompassing the 38–49 sequence was identified in 74% of cases and the peptide HQGLPQEVLNENLLR, position 23–37, in 53%. These three peptides were recognised as the most recurrent in milk binder and were thus selected as suitable targets for the development of the MRM method. Similarly, four peptides from alpha S2 casein and two peptides from beta casein were selected as the most frequently detected in the LC-MSMS analyses (Fig. S3 and S4, in the Supplementary material).

These peptides were also considered to investigate whether they could be used to discriminate among the different milk sources i.e. either bovine or ovine milk. The biological origin of proteins in samples from cultural heritage can be ascertained by identifying species-specific peptides, and is an attractive and still challenging problem [6,17,19,20,47] that can in principle be tackled by MRM approach. Two of the selected peptides from alpha S1 casein and three peptides from alpha S2 casein were demonstrated to be able to discriminate bovine and ovine milk [47]. In particular, FFVAPFPEVFGK and HQGLPQEVLNENLLR are specific of bovine alpha S1 casein, and FALPQYLK, AMKPWIQPK and NAVPITPTLNR of bovine alpha S2 casein, respectively. Similarly, the orthologues peptides FVVAPFPEVFR and NAGPFTPTVNR that are specific to goat alpha S1 casein and goat alpha S2 casein, respectively, were also inserted in the MRM method (Table S2).
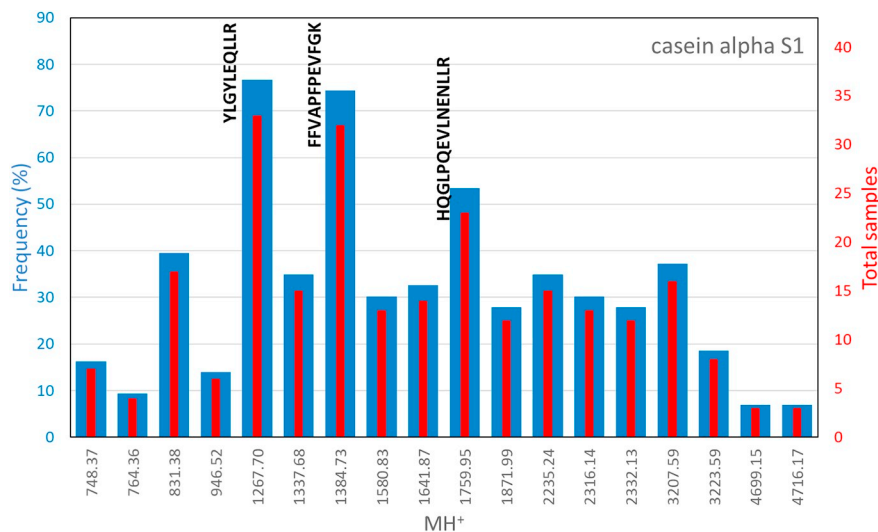
Fig. 2. Frequency of alpha S1 casein detected peptides in milk containing samples. On the main vertical axis the percentage is reported as calculated in respect to the number of samples in the milk containing samples database The secondary vertical axis reports the actual number of observations. In the statistical analysis, detected peptides have been considered only once, regardless possible deamidation detection, while oxidated peptides were considered separately. The sequence of the selected peptides is also reported.

Table 1

Selected peptides: peptides selected upon the analysis of the occurrence in group of samples. The number of times the peptide has been observed and the corresponding percentage in the group of samples is reported. Position in the protein sequence is given as flanking apexes. For peptides from vitellogenin 1 and vitellogenin 2, the percentage in the set of yolk and albumen samples is also shown.

Sample group (number of samples)	Protein (UniProt entry)	Peptide	MH$^+$	No. of observations	%	% in the yolk and albumen samples
Milk containing samples (43)	Alpha S1 casein (P02662)	[106]YLGYLEQLLR[115]	1267.70	33	77	
		[38]FFVAPFPEVFGK[49]	1384.73	32	74	
		[23]HQGLPQEVLNENLLR[37]	1759.95	23	53	
	Alpha S2 casein (P02663)	[189]FALPQYLK[196]	979.56	24	56	
		[130]NAVPITPTLNR[140]	1195.68	22	51	
		[204]AMKPWIQPK[212]	1098.61	15	35	
		[96]ALNEINQFYQK[106]	1367.7	19	44	
	Beta casein (P02666)	[192]AVPYPQR[198]	830.45	30	69	
		[199]DMPIQAFLLYQEPVLGPVR[217]	2186.17	19	44	
		[199]DMPIQAFLLYQEPVLGPVR[217] + Oxi (M)	2202.16	23	39	
Animal glue containing samples (29)	Collagen alpha-1(I) (P02453)	[1062]SGDRGETGPAGPAGPIGPVGAR[1083]	1975.99	23	79	
		[1084]GPAGPQGPR[1092]	836.44	17	59	
		[958]GVVGLPGQR[966] + Hydroxy (P)	898.51	21	72	
	Collagen alpha-2(I) (P02465)	[326]GIPGPVGAAGATGAR[340] + Hydroxy (P)	1267.67	16	55	
		[572]GIPGEFGLPGPAGAR[586] + 2 Hydroxy (P)	1427.73	19	65.5	
		[1066]IGQPGAVGPAGIR[1078]	1192.68	19	65.5	
Albumen containing samples (43)	Ovalbumin (P01012)	[265]LTEWTSSNVMEER[277]	1581.72	23	53	
		[128]GGLEPINFQTAADQAR[143]	1687.87	32	74	
		[324]ISQAVHAAHAEINEAGR[340]	1773.90	30	70	
		[144]ELINSWVESQTNGIIR[159]	1858.97	32	74	
	Ovotransferrin (P02789)	[595]ANVMDYR[601]	868.40	21	49	
		[120]GTEFTVNDLQGK[131]	1308.64	23	53	
		[289]AQSDFGVDTK[298]	1067.50	23	53	
		[275]DDNKVEDIWSFLSK[288]	1695.82	26	60	
		[155]GAIEWEGIESGSVEQAVAK[173]	1959.96	26	60	
		[141]SAGWNIPIGTLLHR[154]	1534.85	25	58	
Yolk containing samples (15)	Vitellogenin-1 (P87498)	[831]LTELLNSNVR[840]	1158.65	4	9	25
		[472]SNIEEVLLALK[482]	1228.71	4	9	25
		[315]LQDLVETTYEQLPSDAPAK[333]	2118.06	4	9	25
		[813]VAGNVQAQITPSPR[826]	1437.78	6	14	37.5
	Vitellogenin-2 (P02845)	[1515]MVVALTSPR[1523]	973.55	10	23	62.5
		[456]EALQPIHDLADEAISR[471]	1777.91	10	23	62.5
		[260]QQLTLVEVR[268]	1085.63	13	30	81
		[642]VGATGEIFVVNSPR[655]	1445.77	12	28	75
		[226]QSDSGTLITDVSSR[239]	1465.71	11	26	69
		[240]QVYQISPFNEPTGVAVMEAR[259]	2236.11	11	26	69
		[919]NIGELGVEKR[928]	1114.62	7	16	44
		[1543]LPLSLPVGPR[1552]	1048.65	8	19	50

The analysis of peptides from animal glue samples lead to the selection of three peptides for the definition of collagen alpha-1(I) and three peptides for collagen alpha-2(I). Non-hydroxylated or unambiguously singly hydroxylated peptides were preferred, since hydroxylation in different position within the same peptide leads to different ions upon fragmentation; however, a peptide of collagen alpha-2(I) with multiple hydroxylation sites was also considered as a test, and the transitions for the different possibilities were separately considered (Tables 1 and S2).

Peptides identified in egg binder showed a larger distribution of observation frequency, making the selection of the marker peptides more complex. A higher number of peptides was therefore selected for ovoalbumin, ovotransferrin, vitellogenin-1 and vitellogenin-2 (four, six, four and eight peptides respectively), in order to maximise our probability of detecting egg in the MRM analysis of samples of unknown composition.

Table 1 reports the ultimate list of the selected peptides with the frequency of observation in the set of paintings/test samples analyzed. To design the MRM method, the selected peptides were thus analyzed in silico using the Skyline software to determine the predicted best transitions and collision energy to generate maximal fragmentation intensities. Table S2 reports the full list of defined peptides for each single target protein of the various proteinaceous binders, with the specific transitions and collision energy that have been used to develop the MRM method.

3.3. MRM analysis

A single MRM method able to detect in a single run the presence of egg, animal glue and milk proteins in samples from cultural heritage was thus developed. This method contains all the precursor ion-daughter ions transitions associated with the selected target peptides constituting the protein signature of each individual proteinaceous binder (see Table S2).

As a preliminary check of the developed assay, some samples with known composition, (containing either milk, animal glue or egg as single proteinaceous component, as well as some combinations of them), were used as test and were analyzed by using the newly set-up MRM method. As an example, Fig. 3 shows the MRM TIC chromatograms for the transitions selected for the peptide 106–115 of bovine alpha S1 casein. All the signals were recorded at the same retention time, thus confirming that the transitions were originated by the same peptide. Table S7 reports the transitions detected in the MRM analyses of the test samples with known composition. The method showed the
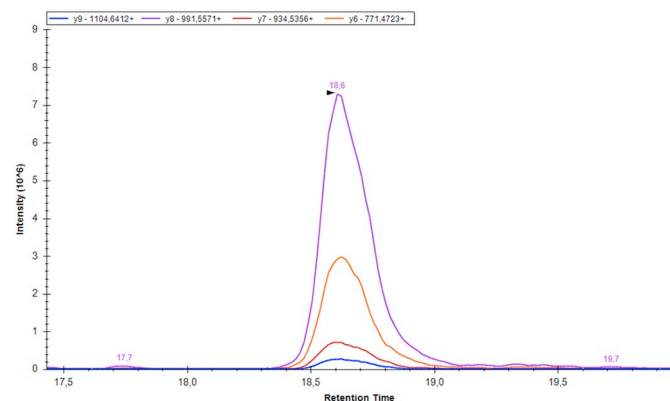

Fig. 3. Test specimen MRM analysis. MRM TIC chromatograms for the best four transitions (634.35 > 1104.64, 634.35 > 991.56, 634.35 > 934.53, 634.35 > 771.47) related to ^{106}YLGYLEQLLR^{115} peptide belonging to alpha S1 casein from *Bos taurus*. The perfect co-elution of the monitored transitions for the selected peptide, at 18.6 confirmed the presence of milk in the analyzed sample and the high selectivity and specificity of the developed method.

transitions associated to the target peptides leading to the unambiguous identification of the specific protein and thus to the assessment of the individual proteinaceous binder. Moreover, in the test sample E17 four peptides from collagen proteins and in the test sample G9 three peptides from vitellogenin-2 could be clearly detected indicating the presence of contaminating material besides the expected binders, contamination possibly occurring during the preparation of the painting tests themselves.

3.4. Analysis of historical samples

The developed MRM approach was thus applied to a sample from the paintings of the Monumental Cemetery of Pisa that had previously been examined by a standard proteomic approach [48]. Several transitions that were ascribed to peptides from caseins and collagens could be detected, thus confirming the occurrence of milk and animal glue (Table S8). Moreover, most interestingly, MRM transitions of peptides from ovalbumin, ovotransferrin and vitellogenin 2 could also be clearly observed indicating the presence of egg within the sample (Fig. 4). It should be underlined that no egg proteins were identified in previous experiments carried out by the standard LC-MSMS procedure [48], thus confirming the higher sensitivity achieved with the MRM assay. Furthermore, when the MRM analysis was carried out on a second sample from the Monumental Cemetery painting that had been previously submitted to a deglycosylating pretreatment with N-Glycosidase F, an even higher number of peptides from egg proteins were detected (Table S8), accordingly with the observation that such a step improves accessibility to protease of the polypeptidic chain of egg proteins [26]. It is worth mentioning, however, that even after the deglycosylation step, no egg proteins were detected in untargeted LC-MSMS analysis carried out on another aliquot of the same sample, possibly because of the overwhelming presence of milk and animal glue proteins. The presence of egg in these samples is not surprising, as it has been previously shown, by means of GC/MS based approaches, that egg was often present in samples collected from the Monumental Cemetery paintings, as it was used as paint binder in other selected areas of the mural paintings [52–54].

Paint samples from the polychromies of the giant Buddha statues from the Baymian valley in Afghanistan that had already been analyzed by standard LC-MSMS procedure [47] were also subject to the developed MRM assay. The detection of several mass transitions associated to the specific peptides from caseins confirmed the presence of milk in the binder, as already assessed by previous analyses [47] (Table S8). The greatly improved performances of MRM procedure were demonstrated by the analysis of some samples from the Buddha statues for which untargeted LC-MSMS analysis provided no results. Fig. S3 shows the MRM analysis of the fragments from the Western Buddha 22-1 [47] clearly displaying the occurrence of several transitions associated to two peptides from bovine alpha S1 and one from alpha S2 caseins, thus confirming the presence of milk also in this sample. Similarly, the MRM approach was able to detect and identify the binder in other three Buddha samples that were negative in the standard LC-MSMS analyses [47] (Table S8). Moreover, since we have previously demonstrated that a mixture of bovine and ovine milks was used as binder in the decoration of the Buddha statues [47], we took advantage of this observation to test the capability of our MRM methods to distinguish highly similar protein sequences even in ancient samples, and eventually provide a tool that could be useful in the challenging problem of identifying organism species. We therefore introduced in the MRM method the transitions for two peptides from alpha S1 and alpha S2 casein from *Capra hircus*, that are homologous to two peptides in the corresponding alpha S1 and alpha S2 casein from *Bos taurus* (Table S2). The concomitant detection of the peptides from ovine and bovine caseins in the MRM analyses confirmed that a mixture of bovine and ovine milks was used as binder [47]. Accordingly, the MRM chromatograms of the sample 214-2 [47] (Fig. 5) clearly displayed the specific
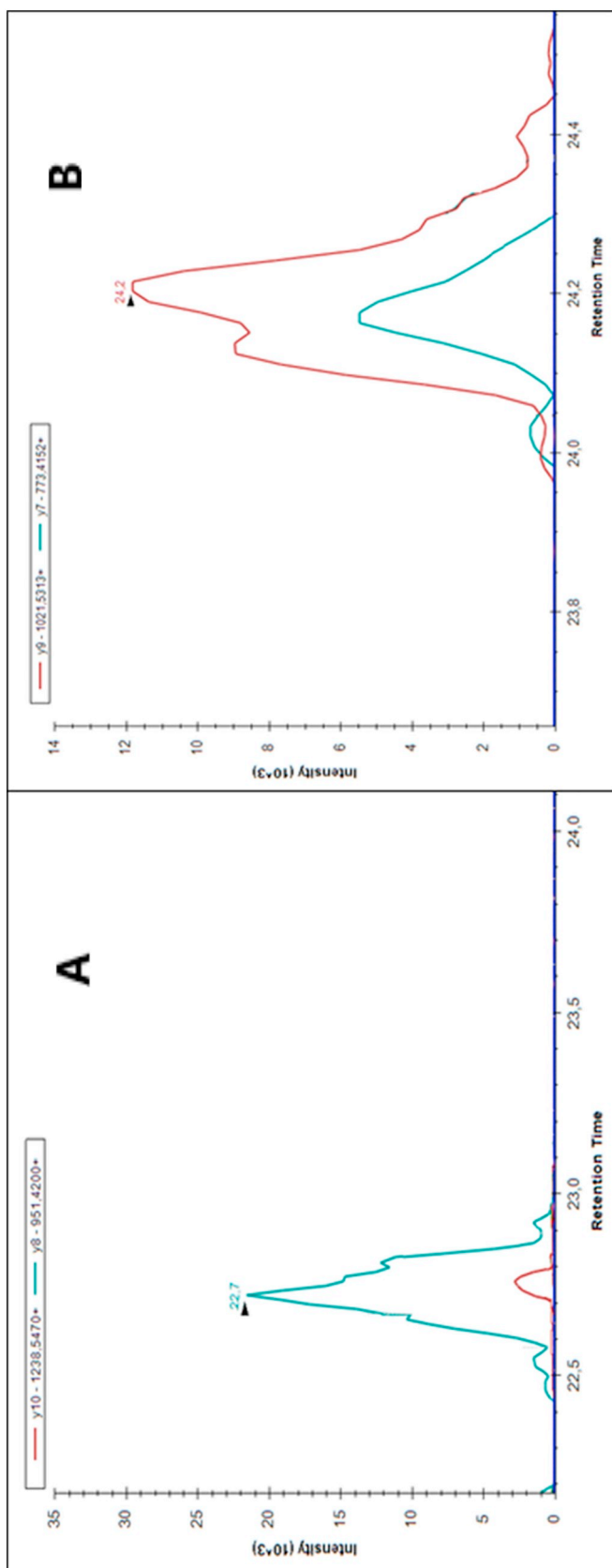
Fig. 4. Monumental Cemetery sample. Panel A: the MRM analysis on Monumental Cemetery sample showed the presence of unidentified egg's proteins from previous LC-MSMS analyses. This data demonstrates the higher sensitivity of the MRM approach and in particular in Panel A representatives MRM TIC chromatograms of monitored transitions 791.36 > 1238.59 and 791.36 > 951.42 for the [265]LTEWTSSNVMEER[277] peptide of Ovalbumin are shown. In Panel B, representatives MRM TIC chromatograms of monitored transitions 654.83 > 1021.53 and 654.83 > 773.41 for [120]GTEFTVNDLQGK[131] peptide of Ovotransferrin are reported.
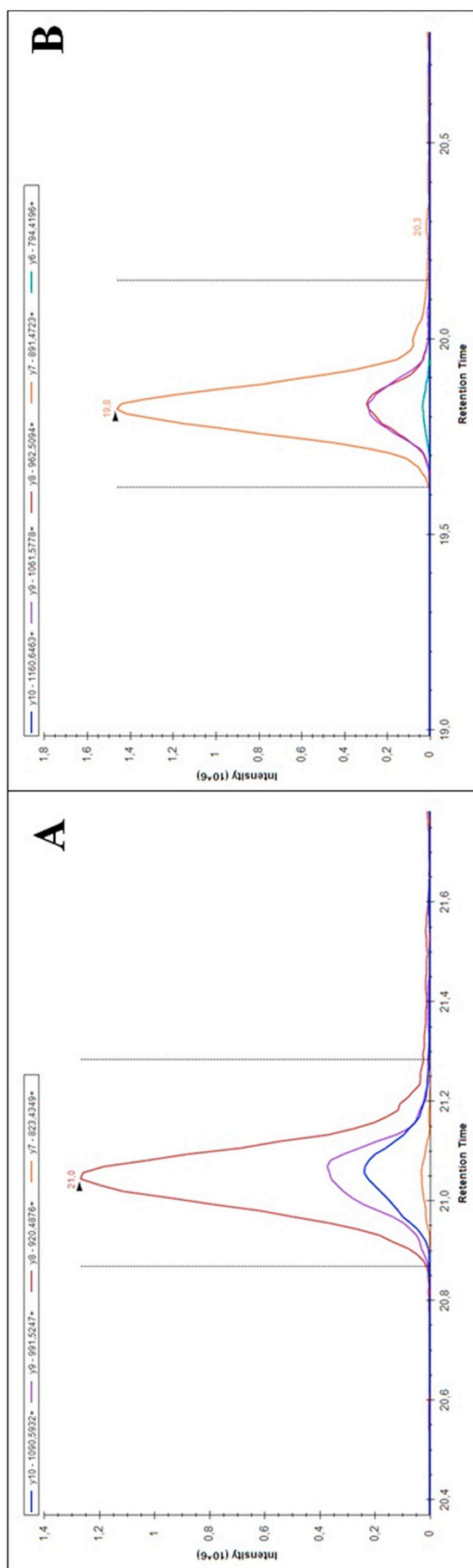
Fig. 5. Buddha sample 214 Clay. Panel A shows representatives MRM TIC chromatograms related to 692.86 > 1237,66, 692.86 > 1090.59, 692.86 > 991.52, 692.86 > 920.49, 692.86 > 823.43 transitions monitored for [38]FFVAPFPEVFGK[49] peptide belonging to alpha S1 casein of *Bos taurus*. Panel B shows representatives MRM TIC chromatograms related to the five monitored transitions (654.35 > 1160.64, 654.35 > 1061.57, 654.35 > 962.50, 654.35 > 891.47, 654.35 > 794.42) for [38]FVVAPFPEVFR[48] peptide for alpha S1 casein of *Capra hircus*. These reported data unambiguously confirmed the presence of bovine and ovine milk in the Buddha 241 Clay sample.

mass transitions for the FFVAPFPEVFGK (Fig. 5A) and FVVAPFPEVFR (Fig. 5B) peptides from bovine and ovine caseins, respectively.

These results clearly demonstrated that the targeted proteomic approach was able to identify proteins that were not detected with a standard proteomics approach, that is capable of identifying proteins in complex mixtures, and distinguish homologous peptides.

4. Discussion

The field of diagnosis in cultural heritage is continuously in search for more and more reliable and alternative methods to identify the materials used in a work of art while preserving as much as possible the artistic object itself [55]. While fully noninvasive techniques will always be preferred as first choice, the development of very sensitive methodologies will greatly reduce the impact of minimal invasive procedures. Identification of proteinaceous binders in paintings was successfully addressed by proteomics approaches providing informative results on precious and irreplaceable samples ([2] and references therein). This paper reports an evolution of this process turning from an untargeted to a targeted proteomics strategy in which only specific peptides representative of selected, relevant proteins, are searched to discriminate between the three most common typologies of protein-based binders in paintings.

Once defined the list of specific proteins most commonly detected in each binder-milk, animal glue and egg based- the selection of diagnostic peptides and the corresponding set of suitable mass transitions needed to develop the MRM method constitutes the critical step. Predictive models based on experimental data [56] have proven to be extremely powerful in the selection of proteotypic peptides in targeted proteomics, and several software packages exist that facilitate this step of the assay development [57]. However, among the factors that could affect the effective detection of peptides in this specific application, the peculiar, solid and aged nature of the samples was considered, and a more empirical approach to the selection of target peptides was herein used.

As a large cohort of painting samples had already been examined during the years in our laboratory by standard LC-MSMS analyses, we screened the results in search for peptides most frequently detected in these untargeted experiments for the three classes of protein-based binders. Discriminative peptides of caseins and collagen that could be considered as representatives of milk and animal glue binders were easily selected. Moreover, an appropriate selection of casein peptides was also used as test to distinguish the animal origin of milk, a specific request in case of milk binders.

Selection of peptides representative of egg binders was more challenging for a number of reasons. First, as stated above, artists used either albumen or yolk or the whole egg and often yolk was contaminated by albumen. Moreover, it has been demonstrated that the extensive glycosylation of egg proteins results in a poor production of suitable peptides and that the confidence of identification can be improved by the introduction of a deglycosylation step before trypsin digestion [26]. The molecular signature of egg binders was eventually defined by selecting ovoalbumin and ovotransferrin as generic representative of the presence of egg and vitellogenin 1 and vitellogenin 2 as specific markers of yolk binders.

The method was validated by analyzing model samples of known composition and was then used to characterize samples from ancient paintings and polychromies: the Monumental Cemetery in Pisa and the Giant Buddha statues in the Baymian valley in Afghanistan. These samples had already been analyzed by standard LC-MSMS procedure, and MRM analyses confirmed previous positive results, and provided successful identification in cases where untargeted procedures had failed, demonstrating the higher sensitivity and selectivity of this targeted approach. As an example, the overabundant milk and animal glue used in the restorations of the mural paintings of the Monumental Cemetery in Pisa prevented the identification of egg in standard LC-

MSMS analyses [48], while the MRM analysis unambiguously revealed the presence of egg-derived peptides, possibly revealing the original binder, covered then by overabundant layers of materials used in restorations.

The higher sensitivity of the MRM procedure also allowed the identification of traces of milk proteins in some of the samples from the Giant Buddha statues for which the untargeted approach had failed [47]. Moreover, the specific case of the Giant Buddha statues was the test case to investigate the potentiality of MRM to distinguish highly similar protein sequences even in the case of ancient samples. This positive result can be considered as the preliminary test paving the way to the exploitation of the MRM assay in the challenging field of species identification. The biological origin of proteins by identification of species-specific peptides is an extremely challenging problem in several fields of the cultural heritage analysis [6,17,19,20,47] that could thus be faced by MRM approach, as we demonstrated by the specific case of distinguishing the animal origin of the milk.

In conclusion, we demonstrated that a targeted, MRM analysis can be successful in samples such as those from ancient objects. Searching for specific signals in a multitude of untargeted signals might be like looking for a needle in a haystack with the naked eye, while mass spectrometer working in a targeted mode could be considered as the use of a magnet in this search.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.microc.2018.09.021.

References

[1] S.T. en Have, Y. Ahmad, A.I. Lamond, Proteomics– current novelties and future directions, J. Anal. Bioanal. Tech. 01 (2011) 1–6, https://doi.org/10.4172/2155-9872.S3-001.

[2] S. Dallongeville, N. Garnier, C. Rolando, C. Tokarski, Proteins in art, archaeology, and paleontology: from detection to identification, Chem. Rev. 116 (2015) 2–79, https://doi.org/10.1021/acs.chemrev.5b00037.

[3] J. Hendy, F. Welker, B. Demarchi, C. Speller, C. Warinner, M.J. Collins, A guide to ancient protein studies, Nat. Ecol. Evol. 2 (2018) 791–799, https://doi.org/10.1038/s41559-018-0510-x.

[4] Y. Zhang, B.R. Fonslow, B. Shan, M.C. Baek, J.R. Yates, Protein analysis by shotgun/bottom-up proteomics, Chem. Rev. 113 (2013) 2343–2394, https://doi.org/10.1021/cr3003533.

[5] C. Tokarski, E. Martin, C. Rolando, M. Post-Traductionnelles, A. Cedex, Identification of proteins in renaissance paintings by proteomics, Anal. Chem. 78 (2006) 1494–1502.

[6] S. Dallongeville, M. Richter, S. Schäfer, M. Kühlenthal, N. Garnier, C. Rolando, C. Tokarski, Proteomics applied to the authentication of fish glue: application to a 17th century artwork sample, Analyst 138 (2013) 5357, https://doi.org/10.1039/c3an00786c.

[7] P. Villa, L. Pollarolo, I. Degano, L. Birolo, M. Pasero, C. Biagioni, K. Douka, R. Vinciguerra, J.J. Lucejko, L. Wadley, A milk and ochre paint mixture used 49,000 years ago at Sibudu, South Africa, PLoS One 10 (2015) e0131273, , https://doi.org/10.1371/journal.pone.0131273.

[8] N. Bleicher, C. Kelstrup, J.V. Olsen, E. Cappellini, Molecular evidence of use of hide glue in 4th millennium BC Europe, J. Archaeol. Sci. 63 (2015) 65–71, https://doi.org/10.1016/j.jas.2015.08.012.

[9] M. Buckley, N.D. Melton, J. Montgomery, Proteomics analysis of ancient food vessel stitching reveals > 4000-year-old milk protein, Rapid Commun. Mass Spectrom. 27

(2013) 531–538, https://doi.org/10.1002/rcm.6481.

[10] E. Cappellini, L.J. Jensen, D. Szklarczyk, A. Ginolhac, R. a R. Da Fonseca, T.W. Stafford, S.R. Holen, M.J. Collins, L. Orlando, E. Willerslev, M.T.P. Gilbert, J.V. Olsen, Proteomic analysis of a Pleistocene mammoth femur reveals more than one hundred ancient bone proteins, J. Proteome Res. 11 (2012) 917–926, https://doi.org/10.1021/pr200721u.

[11] M. Mackie, P. Rüther, D. Samodova, F. Di Gianvincenzo, C. Granzotto, D. Lyon, D.A. Peggie, H. Howard, L. Harrison, L.J. Jensen, J.V. Olsen, E. Cappellini, Palaeoproteomic profiling of conservation layers on a 14th century Italian wall painting, Angew. Chem. Int. Ed. (2018), https://doi.org/10.1002/anie.201713020.

[12] A. Shevchenko, Y. Yang, A. Knaust, H. Thomas, H. Jiang, E. Lu, C. Wang, A. Shevchenko, Proteomics Identifies the Composition and Manufacturing Recipe of the 2500-year Old Sourdough Bread From Subeixi Cemetery in China, (2014), https://doi.org/10.1016/j.jprot.2013.11.016.

[13] C. Warinner, J. Hendy, C. Speller, E. Cappellini, R. Fischer, C. Trachsel, J. Arneborg, N. Lynnerup, O.E. Craig, D.M. Swallow, A. Fotakis, R.J. Christensen, J.V. Olsen, A. Liebert, N. Montalva, S. Fiddyment, S. Charlton, M. Mackie, A. Canci, A. Bouwman, F. Rühli, M.T.P. Gilbert, M.J. Collins, Direct evidence of milk consumption from ancient human dental calculus, Sci. Rep. 4 (2014) 7104, , https://doi.org/10.1038/srep07104.

[14] F. Welker, M.J. Collins, J.A. Thomas, M. Wadsley, S. Brace, E. Cappellini, S.T. Turvey, M. Reguero, J.N. Gelfo, A. Kramarz, J. Burger, J. Thomas-Oates, D.A. Ashford, P.D. Ashton, K. Rowsell, D.M. Porter, B. Kessler, R. Fischer, C. Baessmann, S. Kaspar, J.V. Olsen, P. Kiley, J.A. Elliott, C.D. Kelstrup, V. Mullin, M. Hofreiter, E. Willerslev, J.-J. Hublin, L. Orlando, I. Barnes, R.D.E. MacPhee, Ancient proteins resolve the evolutionary history of Darwin's South American ungulates, Nature 522 (2015) 81–84, https://doi.org/10.1038/nature14249.

[15] F. Welker, M. Hajdinjak, S. Talamo, K. Jaouen, M. Dannemann, F. David, M. Julien, M. Meyer, J. Kelso, I. Barnes, S. Brace, P. Kamminga, R. Fischer, B.M. Kessler, J.R. Stewart, S. Pääbo, M.J. Collins, J.-J. Hublin, Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne, Proc. Natl. Acad. Sci. 113 (2016) 11162–11167, https://doi.org/10.1073/pnas.1605834113.

[16] G. Zilberstein, U. Maor, E. Baskin, A. D'Amato, P.G. Righetti, Unearthing Bulgakov's trace proteome from the Master i Margarita manuscript, J. Proteome 152 (2017) 102–108, https://doi.org/10.1016/j.jprot.2016.10.019.

[17] S. Dallongeville, N. Garnier, D.B. Casasola, M. Bonifay, C. Rolando, C. Tokarski, Dealing with the identification of protein species in ancient amphorae, Anal. Bioanal. Chem. 399 (2011) 3053–3063, https://doi.org/10.1007/s00216-010-4218-2.

[18] C. Solazzo, W.W. Fitzhugh, C. Rolando, C. Tokarski, Identification of protein remains in archaeological potsherds by proteomics, Anal. Chem. 80 (2008) 4590–4597, https://doi.org/10.1021/ac800515v.

[19] S. Dallongeville, M. Koperska, N. Garnier, G. Reille-Taillefert, C. Rolando, C. Tokarski, Identification of animal glue species in artworks using proteomics: application to a 18th century gilt sample, Anal. Chem. 83 (2011) 9431–9437, https://doi.org/10.1021/ac201978j.

[20] A. Chambery, A. Di Maro, C. Sanges, V. Severino, M. Tarantino, A. Lamberti, A. Parente, P. Arcari, Improved procedure for protein binder analysis in mural painting by LC-ESI/Q-q-TOF mass spectrometry: detection of different milk species by casein proteotypic peptides, Anal. Bioanal. Chem. 395 (2009) 2281–2291, https://doi.org/10.1007/s00216-009-3183-0.

[21] G. Leo, L. Cartechini, P. Pucci, A. Sgamellotti, G. Marino, L. Birolo, Proteomic strategies for the identification of proteinaceous binders in paintings, Anal. Bioanal. Chem. 395 (2009) 2269–2280.

[22] W. Fremout, M. Dhaenens, S. Saverwyns, J. Sanyova, P. Vandenabeele, D. Deforce, L. Moens, Development of a dedicated peptide tandem mass spectral library for conservation science, Anal. Chim. Acta 728 (2012) 39–48, https://doi.org/10.1016/j.aca.2012.03.037.

[23] W. Fremout, M. Dhaenens, S. Saverwyns, J. Sanyova, P. Vandenabeele, D. Deforce, L. Moens, Tryptic peptide analysis of protein binders in works of art by liquid chromatography–tandem mass spectrometry, Anal. Chim. Acta 658 (2010) 156–162, https://doi.org/10.1016/j.aca.2009.11.010.

[24] H. Rao, Y. Yang, I. Abuduresule, W. Li, X. Hu, C. Wang, Proteomic identification of adhesive on a bone sculpture-inlaid wooden artifact from the Xiaohe Cemetery, Xinjiang, China, J. Archaeol. Sci. 53 (2015) 148–155, https://doi.org/10.1016/j.jas.2014.10.010.

[25] R. Vinciguerra, A. De Chiaro, P. Pucci, G. Marino, L. Birolo, Proteomic strategies for cultural heritage: form bones to paintings, Microchem. J. 126 (2015) 341–348, https://doi.org/10.1016/j.microc.2015.12.024.

[26] R. Vinciguerra, E. Galano, F. Vallone, G. Greco, A. Vergara, I. Bonaduce, G. Marino, P. Pucci, A. Amoresano, L. Birolo, Deglycosylation step to improve the identification of egg proteins in art samples, Anal. Chem. 87 (2015) 10178–10182, https://doi.org/10.1021/acs.analchem.5b02423.

[27] H. Rao, B. Li, Y. Yang, Q. Ma, C. Wang, Proteomic identification of organic additives in the mortars of ancient Chinese wooden buildings, Anal. Methods 7 (2015) 143–149, https://doi.org/10.1039/c4ay01766h.

[28] S. Orsini, A. Yadav, M. Dilillo, L.A. McDonnell, I. Bonaduce, Characterization of Degraded Proteins in Paintings Using Bottom-up Proteomic Approaches: New Strategies for Protein Digestion and Analysis of Data, (2018), https://doi.org/10.1021/acs.analchem.8b00281.

[29] P. Picotti, B. Bodenmiller, R. Aebersold, Proteomics meets the scientific method, Nat. Methods 10 (2013) 24–27, https://doi.org/10.1038/nmeth.2291.

[30] V. Vidova, Z. Spacil, A review on mass spectrometry-based quantitative proteomics: targeted and data independent acquisition, Anal. Chim. Acta 964 (2017) 7–23, https://doi.org/10.1016/j.aca.2017.01.059.

[31] J.D. Baty, P.R. Robinson, Single and multiple ion recording techniques for the analysis of diphenylhydantoin and its major metabolite in plasma, Biol. Mass Spectrom. 4 (1977) 36–41, https://doi.org/10.1002/bms.1200040104.

[32] C.E. Parker, C.H. Borchers, Mass spectrometry based biomarker discovery, verification, and validation—quality assurance and control of protein biomarker assays, Mol. Oncol. 8 (2014) 840–858, https://doi.org/10.1016/j.molonc.2014.03.006.

[33] L.R. Ruhaak, C.B. Lebrilla, Applications of multiple reaction monitoring to clinical glycomics, Chromatographia 78 (2015) 335–342, https://doi.org/10.1007/s10337-014-2783-9.

[34] A. Illiano, V. Arpino, G. Pinto, A. Berti, V. Verdoliva, G. Peluso, P. Pucci, A. Amoresano, Multiple reaction monitoring tandem mass spectrometry approach for the identification of biological fluids at crime scene investigations, Anal. Chem. 90 (2018) 5627–5636, https://doi.org/10.1021/acs.analchem.7b04742.

[35] F.L. Craciun, V. Bijol, A.K. Ajay, P. Rao, R.K. Kumar, J. Hutchinson, O. Hofmann, N. Joshi, J.P. Luyendyk, U. Kusebauch, C.L. Moss, A. Srivastava, J. Himmelfarb, S.S. Waikar, R.L. Moritz, V.S. Vaidya, RNA sequencing identifies novel translational biomarkers of kidney fibrosis, J. Am. Soc. Nephrol. 27 (2016) 1702–1713, https://doi.org/10.1681/ASN.2015020225.

[36] M.A. Gillette, S.A. Carr, Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry, Nat. Methods 10 (2013) 28–34, https://doi.org/10.1038/nmeth.2309.

[37] R. Hüttenhain, M. Soste, N. Selevsek, H. Röst, A. Sethi, C. Carapito, T. Farrah, E.W. Deutsch, U. Kusebauch, R.L. Moritz, E. Niméus-Malmström, O. Rinner, R. Aebersold, Body fluids using targeted proteomics, 4 (2013), https://doi.org/10.1126/scitranslmed.3003989.

[38] U. Kusebauch, D.S. Campbell, E.W. Deutsch, L. Hood, R. Aebersold, R.L. Moritz, C.S. Chu, D.A. Spicer, M.-Y. Brusniak, J. Slagel, Z. Sun, J. Stevens, B. Grimes, D. Shteynberg, M.R. Hoopmann, P. Blattmann, A.V. Ratushny, O. Rinner, P. Picotti, C. Carapito, C.-Y. Huang, M. Kapousouz, H. Lam, T. Tran, E. Demir, J.D. Aitchison, C. Sander, Human SRMAtlas: a resource of targeted assays to quantify the complete human proteome, Cell 166 (2016) 766–778, https://doi.org/10.1016/j.cell.2016.06.041.

[39] S. Surinova, M. Choi, S. Tao, P.J. Schüffler, C.-Y. Chang, T. Clough, K. Vysloužil, M. Khoylou, J. Srovnal, Y. Liu, M. Matondo, R. Hüttenhain, H. Weisser, J.M. Buhmann, M. Hajdúch, H. Brenner, O. Vitek, R. Aebersold, Prediction of colorectal cancer diagnosis based on circulating plasma proteins, EMBO Mol. Med. 7 (2015) 1166–1178, https://doi.org/10.15252/emmm.201404873.

[40] H.A. Ebhardt, A. Root, C. Sander, R. Aebersold, Applications of targeted proteomics in systems biology and translational medicine, Proteomics 15 (2015) 3193–3208, https://doi.org/10.1002/pmic.201500004.

[41] P. Picotti, M. Clement-Ziza, H. Lam, D.S. Campbell, A. Schmidt, E.W. Deutsch, H. Röst, Z. Sun, O. Rinner, L. Reiter, Q. Shen, J.J. Michaelson, A. Frei, S. Alberti, U. Kusebauch, B. Wollscheid, R. Moritz, A. Beyer, R. Aebersold, P.D. Program, A complete mass spectrometric map for the analysis of the yeast proteome and its application to quantitative trait analysis, Nature 494 (2013) 266–270, https://doi.org/10.1038/nature11835.

[42] C. Karlsson, L. Malmström, R. Aebersold, J. Malmström, Proteome-wide Selected Reaction Monitoring Assays for the Human Pathogen *Streptococcus pyogenes*, (2012), https://doi.org/10.1038/ncomms2297.

[43] R. Hynek, S. Kuckova, J. Hradilova, M. Kodicek, Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry as a tool for fast identification of protein binders in color layers of paintings, Rapid Commun. Mass Spectrom. 18 (2004) 1896–1900, https://doi.org/10.1002/rcm.1570.

[44] W. Fremout, S. Kuckova, M. Crhova, J. Sanyova, S. Saverwyns, R. Hynek, M. Kodicek, P. Vandenabeele, L. Moens, Classification of protein binders in artist's paints by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry: an evaluation of principal component analysis (PCA) and soft independent modelling of class analogy (SIMCA), Rapid Commun. Mass Spectrom. 25 (2011) 1631–1640, https://doi.org/10.1002/rcm.5027.

[45] I.D. van der Werf, C.D. Calvano, F. Palmisano, L. Sabbatini, A simple protocol for Matrix Assisted Laser Desorption Ionization-time of flight-mass spectrometry (MALDI-TOF-MS) analysis of lipids and proteins in single microsamples of paintings, Anal. Chim. Acta 718 (2012) 1–10, https://doi.org/10.1016/J.ACA.2011.12.056.

[46] C.D. Calvano, I.D. van der Werf, F. Palmisano, L. Sabbatini, Fingerprinting of egg and oil binders in painted artworks by matrix-assisted laser desorption ionization time-of-flight mass spectrometry analysis of lipid oxidation by-products, Anal. Bioanal. Chem. 400 (2011) 2229–2240, https://doi.org/10.1007/s00216-011-4919-1.

[47] A. Lluveras-Tenorio, R. Vinciguerra, E. Galano, C. Blaensdorf, E. Emmerling, M.P. Colombini, L. Birolo, I. Bonaduce, GC/MS and proteomics to unravel the painting history of the lost Giant Buddhas of Bāmiyān (Afghanistan), PLoS One 12 (2017), https://doi.org/10.1371/journal.pone.0172990.

[48] G. Leo, I. Bonaduce, A. Andreotti, G. Marino, P. Pucci, M.P. Colombini, L. Birolo, Deamidation at asparagine and glutamine as a major modification upon deterioration/aging of proteinaceous binders in mural paintings, Anal. Chem. 83 (2011) 2056–2064, https://doi.org/10.1021/ac1027275.

[49] M. Gambino, F. Cappitelli, C. Cattò, A. Carpen, P. Principi, L. Ghezzi, I. Bonaduce, E. Galano, P. Pucci, L. Birolo, F. Villa, F. Forlani, A simple and reliable methodology to detect egg white in art samples, J. Biosci. 38 (2013) 397–408, https://doi.org/10.1007/s12038-013-9321-z.

[50] L. Birolo, A. Tomeo, M. Trifuoggi, F. Auriemma, L. Paduano, A. Amoresano, R. Vinciguerra, C. De Rosa, L. Ferrara, A. Giarra, A. Luchini, C. De Maio, G. Greco, A. Vergara, A hypothesis on different technological solutions for outdoor and indoor Roman wall paintings, Archaeol. Anthropol. Sci. 9 (2017) 591–602, https://doi.

org/10.1007/s12520-016-0408-y.

[51] B. MacLean, D.M. Tomazela, N. Shulman, M. Chambers, G.L. Finney, B. Frewen, R. Kern, D.L. Tabb, D.C. Liebler, M.J. MacCoss, Skyline: an open source document editor for creating and analyzing targeted proteomics experiments, Bioinformatics 26 (2010) 966–968, https://doi.org/10.1093/bioinformatics/btq054.

[52] I. Bonaduce, M.P. Colombini, Gas chromatography/mass spectrometry for the characterization of organic materials in frescoes of the Momumental Cemetery of Pisa (Italy), Rapid Commun. Mass Spectrom. 17 (2003) 2523–2527, https://doi.org/10.1002/rcm.1222.

[53] U. Baldini, C. Baracchini, I. Bonaduce, A. Caleca, G. Caponi, M. Colombini, E. Luppichini, M. Spampinato, Una storia complicata: gli affreschi del Camposanto Monumentale di Pisa, in: V. Arcadia Ricerche (Ed.), Sci. e Beni Cult. 2005, pp. 17–27.

[54] A. Andreotti, I. Bonaduce, M. Perla Colombini, C. Baracchini, A. Caleca, A. Paolucci,

Saving the Medieval paintings by the master painter of The Triumph of Death in Pisa, 15th Trienn. Conf. New Delhi, 22–26 Sept. 2008 Prepr. (ICOM Comm. Conserv.), 2008, pp. 825–832.

[55] M.P. Colombini, A. Andreotti, I. Bonaduce, F. Modugno, E. Ribechini, Analytical strategies for characterizing organic paint media using gas chromatography/mass spectrometry, Acc. Chem. Res. 43 (2010) 715–727, https://doi.org/10.1021/ar900185f.

[56] P. Mallick, M. Schirle, S.S. Chen, M.R. Flory, H. Lee, D. Martin, J. Ranish, B. Raught, R. Schmitt, T. Werner, B. Kuster, R. Aebersold, Computational prediction of proteotypic peptides for quantitative proteomics, Nat. Biotechnol. 25 (2007) 125–131, https://doi.org/10.1038/nbt1275.

[57] C.M. Colangelo, L. Chung, C. Bruce, K.-H. Cheung, Review of software tools for design and analysis of large scale MRM proteomic datasets, Methods 61 (2013) 287–298, https://doi.org/10.1016/j.ymeth.2013.05.004.