UNIVERSITÀ DEGLI STUDI DI PARMA

Department of Pharmacy

PhD Program in
Biochemistry and Molecular Biology

XXVIII cycle

INVESTIGATION OF HEALTHY AND INFECTED NEWBORN SALIVA:

THE ROLE OF PROTEINS AND LIPIDS AS POTENTIAL INFLAMMATORY DIAGNOSTIC TOOLS

Coordinator:
Prof. Andrea Mozzarelli

Tutor:
Dr. Barbara Pioselli

PhD student:
GIOVANNI PIGHINI

2013-2015

INDEX

GENERAL INTRODUCTION **1**

BIBLIOGRAPHY 5

1- 48 HOURS AFTER THE BIRTH: WHAT'S IN THE SALIVA OF THE BABY? **8**

ABSTRACT 8

INTRODUCTION 9

Introduction to experimental platform *11*

Isoelectric focusing *13*

SDS-PAGE: second dimension *14*

Protein staining *15*

Protein identification *16*

Sample preparation for MALDI analysis *17*

MATERIALS AND METHODS 19

Reagents and tools *19*

Sample preparation *21*

2D-Electrophoresis *21*

Cup-loading *23*

Protein detection *23*

Protein digestion *24*

Mass spectrometry *24*

Western blot analysis *24*

RESULTS AND DISCUSSION 25

2D-electrophoresis: optimization of experimental set up *25*

2D-electrophoresis: healthy newborn saliva *26*

Protein identifications in healthy newborn saliva *28*

Isoelectric point *31*

Molecular weight *32*

Subcellular location *33*

Differential proteomics of healthy newborns: protein expression at 24 hours and 48 hours *34*
Healthy and infected newborns *37*
Alpha-1-antitrypsin in healthy newborns *39*
Conclusions 41
Bibliography 42
2- CHARACTERIZATION OF THE ACUTE PHASE PROTEIN ALPHA-1-ANTITRYPSIN USING LC-MS PLATFORMS: MARKER OF INFLAMMATION IN NEWBORN SALIVA AND "F" ALLELIC VARIANT.... 46
Abstract 46
Introduction 48
Serpins *49*
Alpha -1-antitrypsin: an acute phase protein involved in the onset of emphysema *49*
Function *51*
"F" variant *51*
Development of a LC-MS based method for the investigation of an inflammatory condition in the early stages of newborns life *52*
Materials and methods 54
Sample preparation: digestion of A1AT standard *55*
Sample preparation: digestion of melittin *56*
Spectrophotometric assay *56*
Liquid chromatography *56*
MS parameters *57*
Results and discussion 59
Alpha-1-antitrypsin: MS/MS spectrum *63*
Overcoming the shift of retention time *64*
LysC protease: a deeper investigation of the activity *67*
Melittin: a different substrate for LysC *68*
The role of serum albumin as "crowding agent" *71*
A1AT elution: crowding and chromatography effect *75*
Alpha-1-antitrypsin: proof of concept of quantitative analysis *77*

Alpha-1-antitrypsin: investigation of the "F" allelic variant *79*

CONCLUSIONS 85

BIBLIOGRAPHY 86

3- LIPIDS PROFILE INVESTIGATION OF HEALTHY AND INFECTED NEWBORN SALIVA: PHOSPHOLIPIDS AND EICOSANOIDS AS MARKER OF INFLAMMATION 89

ABSTRACT 89

INTRODUCTION 90

MATERIALS AND METHODS 93

Phospholipids: targeted mass spectrometry experiment setup *94*

Inflammatory mediators: targeted mass spectrometry experiment setup *96*

Liquid chromatography experiment setup *98*

Sample preparation *99*

RESULTS AND DISCUSSION 102

Phospholipids *102*

Lysoforms *106*

Inflammatory mediators *109*

Quantitative considerations about eicosanoids *111*

Healthy… or infected? *114*

CONCLUSIONS 117

Attachment A: potential identifications from Lipidmaps 119

Attachment B: potential identifications from Lipidmaps 119

BIBLIOGRAPHY 120

General introduction

Perinatal infections are a frequent occurrence that represents a risk factor for newborns life. Every year about 1.4 million newborns die in the world for infectious diseases [1]. An early diagnosis can increase the chance of a positive outcome for the treatment, morbidity and mortality [1,2].

The sources of infection for newborns have maternal or environmental origins: during pregnancy and at childbirth, infection is commonly a result of a chorioamnionitis or in other cases it can originates from diseases of maternal genital tract. After birth, the environment and the staff who care for the newborn become the main source of infection transmission [3].

The scientific and technological progress has greatly improved the chances of survival also for infants with a very low gestational age (23-24 weeks) [3]. Nevertheless, infectious diseases in the neonatal period are still associated with the possibility of permanent disabling health conditions as consequences of neurodevelopmental, motor, sensory and cognitive sequelae, but also chronic alterations of the lung, intestine and kidney functions can occur [3]. For this reason, neonatologists are constantly looking for markers able to identify the early onset of an infectious disease.

Several features should characterize an infection diagnostic assay:

i. it should have high predictive value and sensitivity and it should not be affected by non-infectious diseases.
ii. the biomarker concentration should be representative of the disease severity, and should be relevant for guiding the therapy.
iii. the reproducibility and the “early onset” are two fundamental features to permit a quick and efficient clinical intervention [4].

Several biochemical markers of sepsis were identified including, for example, the white blood cell count, acute phase proteins such as C-reactive protein (CRP) and procalcitonin (PCT), fibronectin, cytokines such as IL-1, IL-6, IL-2, IL-8, IL-10 and TNF, cell surface antigens such as CD11, neutrophil chemotaxis, adhesion molecules (ICAM), complement components and many others [4,5]. Nevertheless, the currently available

tests do not possess all the ideal biomarker features. Clinical observation remains crucial for the neonatologist as well as the identification of the pathogenic agent using cultures, although this is not always easy [6].

With the purpose of minimizing the risk of infection, several prenatal and postnatal protocols were developed. The main drawbacks associated to these protocols are the unsatisfactory sensitivity or specificity of the chosen biomarkers, the elevated costs and the need of invasive methods to collect relatively high volumes of biological fluids from newborns. For example, maternal indicators tests have low specificity, the amniotic fluid cultivation tests have high risk of contamination and the measurement of interleukins in the amniotic fluid and in maternal blood are not applicable as routine. For newborns, the most studied indicators of inflammation are CRP, PCT and interleukins, especially IL-6.

Blood levels of CRP increase after 6 hours from the onset of the inflammation and the highest value is reached between 18 and 24 hours. Nevertheless, this protein can also be produced as a result of non-infectious factors such as trauma, complicated delivery, activation of the inflammatory cascade after hemolysis [7,8]. Therefore, CRP might represent a reliable marker, but is indicative of an advanced inflammatory phase.

Another relevant protein is procalcitonin that is produced in the liver upon IL-6 and TNFα stimulation. This protein is not detectable in the blood of healthy subjects so it is considered a good early marker. In fact, differently from CRP the highest blood value is reached between 6 and 8 hours and remains high for approximately 24 hours [7,9,10]. The sensitivity reported in literature varies from 58% to 100%. On the other hand, the specificity varies between 50% and 100% and depends on the methods of analysis, the protocols, the number of samples and collection times [1,4]. Some issues are related to the fact that there is also a physiological increase of procalcitonin in the phase immediately post-partum. Furthermore, false negatives or positives were reported as a result of trauma, acute respiratory diseases and suffocation in the absence of bacterial infections.

IL-6 and IL-8 are sensitive and specific markers, whose levels increase even before the onset of the first symptoms (and suspected infection). They decrease after 12-24 hours, thus increasing the risk of false negatives [9,11]. A hypothetical solution to

increase their sensitivity is matching the detection of the IL-6 and -8 with a constant monitoring of CRP.

The detection of cellular markers such as CD11 is difficult to interpret due to the variability of the results. Furthermore, this analysis is very expansive and characterized by a low reproducibility and long times to obtain the results useful for clinical applications [7].

The development of new methods that use smaller volumes of biological samples to identify markers of inflammation and infection, the so-called micro-methods, are under evaluation. These new methods, already widely in use in neonatology, partially allow to overcome the problem of the volume of biological samples, but not the invasiveness.

Saliva is a biological fluid easy to collect, to store and to transport, so it could be very useful in the early diagnosis of infections [2]. Saliva collection is safe and non-invasive and overcomes the difficulties related to blood sampling. The analysis of saliva through proteomics is reproducible [2]. The issues related to the use of this biological fluid for diagnostic purposes are the times of analysis and consequently the clinical applicability, the methods standardization and the initial cost of the instrumentation [8]. In particular, the non-invasive analysis of saliva allows repeated sample collections. This is relevant in particular for newborns with very low weight.

In the previous work “Bioanalytical solutions to characterize human proteome of clinical interest: healthy newborns and their early infections” L. Tigli [12] performed targeted analyses focused on the development of an analytical method based on liquid chromatography coupled to mass spectrometry to qualitatively and quantitatively assay C-reactive protein, procalcitonin and calcitonin within salivary samples. Analytical MS-based methods present some undeniable advantages over the widely used immunochemical methods: high specificity, high throughput, capability to perform multiplex analyses (i. e. to monitor several analytes at the same time) and also to investigate structural modifications of proteins (e. g. post-translational modifications), fast method development and much lower cost per single analysis. This quantitative platform can be applied to the determination of any protein highlighted as potential

inflammatory marker. In fact, recent studies have shown the possibility to identify and quantify several salivary inflammatory proteins using proteomics techniques. In particular, saliva is able to describe a systemic condition thanks to the presence in it of molecules associated with hormonal states, metabolic and food influences, immunological conditions and cancer [2]. Selecting the main biomarkers of bacterial infection, representative sentinel molecules can be monitored to obtain an early diagnosis of an infectious disease (or inflammatory process). Once that the standardization of the saliva assay is obtained, the times of analysis will be comparable to blood ones, but with greater easiness and repeatability of material collection.

Bibliography

1) Mahbuba Meem, Joyanta K. Modak, Roman Mortuza and al. Biomarkers for diagnosis of neonatal infections: A systematic analysis of their potential as a point-of-care diagnostics. Glob Health. 2011 Dec; 1(2): 201–209.
2) Boja E, Hiltke T. Rivers R et al Evolution of Clinical Proteomics and its Role in Medicine J. Proteome Res., 2011, 10 (1), pp 66–84.
3) Buhimschi CS, Bhandari V, Hamar BD, et al. Proteomic profiling of the amniotic fluid to detect inflammation, infection, and neonatal sepsis. PLoS Med 2007;4:e18.
4) Ng PC. Diagnostic markers of infection in neonates. Arch Dis Child Fetal Neonatal Ed 2004; 89: F229-F235.
5) Bhandari V. Effective Biomarkers for Diagnosis of Neonatal Sepsis. J Ped Infect Dis (2014);
6) Pammi M, Flores A, Leeflang M and al. Molecular assays in the diagnosis of neonatal sepsis: A systematic review and meta-analysis. Pediatrics. 2011;128.
7) Hatherill M, Tibby SM, Sykes K and all. Diagnostic markers of infection: comparison of procalcitonin with C reactive protein and leucocyte count. Arch Dis Child 1999; 81:417–421.
8) Mussap M. Laboratory medicine in neonatal sepsis and inflammation. J Matern Fetal Neonatal Med 2012;25(Suppl 4):32-34.
9) Chiesa C, Pellegrini G, Panero A et al. C-reactive protein, interleukin-6, and procalcitonin in the immediate postnatal period: influence of illness severity, risk status, antenatal and perinatal complications, and infection. Clin Chem. 2003;49:60–68.
10) Isidor B, Caillaux G, Gilquin V and al. The use of procalcitonin in the diagnosis of late-onset infection in neonatal intensive care unit patients. Scand J Infect Dis. 2007;39:1063–1066.
11) Kocabaş E1, Sarikçioğlu A, Aksaray N and al. Role of procalcitonin, C-reactive protein, interleukin-6, interleukin-8 and tumor necrosis factor-alpha in the diagnosis of neonatal sepsis. Turk J Pediatr. 2007 Jan-Mar;49(1):7-20.

12) Tigli L. “Bioanalytical solutions to characterize human proteome of clinical interest: healthy newborns and their early infections”;

1-48 hours after the birth: what's in the saliva of the baby?

Abstract

We have carried out a discovery proteomics investigation aimed at identifying disease biomarkers present in saliva, and, more specifically, early biomarkers of inflammation. The proteomic characterization of saliva is possible due to the straightforward and non-invasive sample collection that allows repetitive analyses for pharmacokinetic studies. These advantages are particularly relevant in the case of newborn patients.

The salivary samples were collected from healthy and infected (n=1) newborns. Proteins were extracted through cycles of sonication, precipitated in ice-cold acetone, resuspended and resolved by 2D-electrophoresis. Image analysis of the gels was performed in order to identify differential protein expression in health and infected newborns. Selected spots were excised and digested with trypsin. The corresponding peptide mass fingerprinting was obtained by MALDI TOF/TOF mass spectrometry and submitted to database searches for the identification. We report the classification of proteins according to isoelectric point (pI), molecular weight (MW), localization and function. Furthermore, the protein alpha-1-antitrypsin, correlated with inflammation, was detected differently expressed in the infected newborn salivary sample and further investigated with immuno-based approach (western blot).

Introduction

Human saliva is a complex biological fluid composed of water (99%), electrolytes, nitrogenous products and macromolecules (proteins, nucleic acids) [1]. It derives predominantly from three major salivary glands (90%). A consistent part of the final fluid originates from the blood-saliva exchange through the capillary fenestrations that permit solutes and electrolytes to leave the blood vessels and reach the salivary glands.

Saliva contributes to the digestion of food and to the maintenance of oral hygiene. In particular, the three most relevant functions associated to saliva are to lubricate, to protect (due to the action of proteins such as cystatins, immunoglobulins and lysozymes) and to enable oral functions [2]. Nevertheless, the abundance of ions in saliva permits also to keep the acidity of the mouth within the range of pH 6.2–7.4. Both sympathetic and parasympathetic systems stimulate the production of saliva. In particular, the sympathetic stimulation leads to noradrenaline release onto the salivary acinar cells producing thick saliva, instead the parasympathetic stimulation leads to acetylcholine release onto the salivary acinar cells producing watery saliva.

In the last few years, saliva has been used not only to diagnose oral disease [3,4], but also for monitoring systemic conditions. For example, the recent use of saliva in the diagnosis of human immunodeficiency virus (HIV) [5], carcinoma, cardiac disease and autoimmune diseases has demonstrated the suitability of saliva as potential diagnostic tool. The significant fluid-exchange between blood and saliva which occurs during saliva production (this transfer can follow different pathways: diffusion, active transport, ultrafiltration, transudation from either the gingival crevicular fluid or the oral mucosa) allows the detection of any proteins and solutes present in the blood also in saliva. By comparing the main biological fluids, blood, urine and saliva, the latter is a more convenient biological fluid to monitor a condition of inflammation for the following reasons:

1. The non-invasive collection;
2. The amount of saliva that can be easily collected;
3. The immediate availability of the samples;

4. Low cost of storage;
5. Unlike blood, saliva does not clot so it is easier to handle.

The negative side is the lower concentration of proteins and solutes in saliva compared with the levels in blood. However exploiting high sensitive techniques, such as mass spectrometry, the low concentration of analytes in saliva is no longer a limitation.

Early-onset sepsis (EOS), defined as a septic state manifested in the first three days of life, is one of the major causes of neonatal mortality. Surviving infants are at increased risk for developing morbidities, including bronchopulmonary dysplasia, prolonged hospital stay, and neurodevelopment impairment. Therefore, the availability of diagnostic tools allowing the early detection of an ongoing infective state represents an urgency in neonatal clinical care units. In the present study protocol, saliva samples were collected with swabs and without using any substance to stimulate the salivation to avoid stress to the newborn. In particular saliva samples were collected during the first 48 hours of newborn life, according to an approved Ethic Committee procedure. Our aim was to characterize the healthy newborn salivary proteome using the 2D-electrophoresis and MALDI TOF/TOF mass spectrometry approach. In previous works, saliva characterization was focused on proteins expression in adults [6], general not-protein composition at the third day of life [7] and on the composition of the salivary proteome of preterm human newborn, but in this last case using sample collected not immediately after the delivery [8]. Even though saliva collection is considered a non-invasive procedure, samples availability was a considerable issue especially in the case of infected newborns. It needs to be acknowledged that it has been difficult to obtain authorization and inform consent from parents whose child was diagnosed with inflammation or infection. The present study thus reports for the first time the identification of newborn saliva proteins in the acidic soluble fraction of whole saliva, in the early hours after delivery. This condition has never been reported and represents a fundamental step to discriminate between the salivary proteome of healthy and infected newborns, in particular, to identify proteins that are differently expressed in inflammation. Applicability of this approach to non-invasive diagnosis of inflammation will also be discussed on the basis of the results

obtained upon comparison in between a consistent pool of saliva sample from healthy newborns and a few case study from infected newborns.

Introduction to experimental platform

Saliva samples were analysed with two-dimensional gel electrophoresis (2D-PAGE). This technique is composed by two steps: the first dimension consisting in isoelectric focusing (IEF), where proteins are loaded on IPG strips and separated according to their isoelectric point (pI), and the second dimension, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), where proteins with the same pI are separated according to their molecular weight (Figure 1-1).

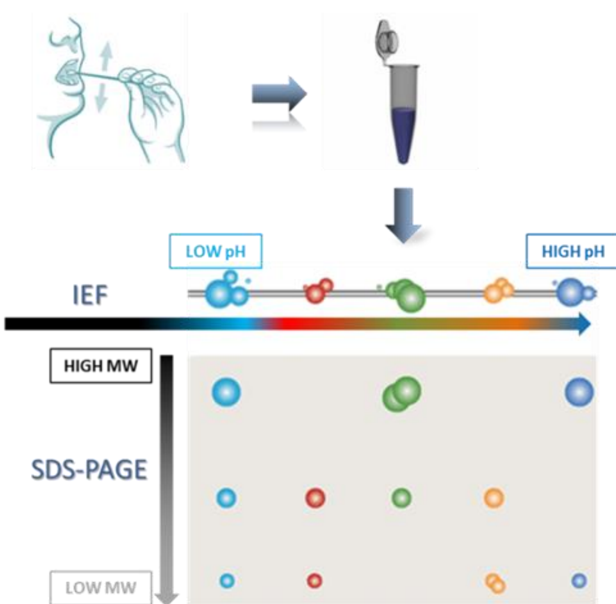


Figure 1-1: 2D PAGE analysis. After sample preparation proteins are loaded onto the IPG strip and separated according to their pI in the first dimension (IEF). Proteins are then separated in the second dimension (SDS-PAGE) according to their molecular weight. Modified from Bio-Rad.

This technique is a powerful tool to analyse and detect proteins from very complex biological mixtures and it also provides a method in order to estimate the number of the most abundant proteins in a sample [9]. To obtain an optimal resolution several parameters have to be considered. In particular, during the first dimension it is necessary to optimize the time and voltage for every step, considering the sample complexity, the presence of contaminants and the total amount of protein. The resolution is determined

by the slope of the pH gradient in the strip and the electric field strength applied. For this purpose, different sizes of IPG strips can be exploited to load different protein amounts. Furthermore, to deeper investigate a specific pH range, several IPG strips are commercially available: restricting the pH range maintaining constant the strip length is an optimal strategy to improve spots' focalization. Regarding the second dimension, the spot resolution can be increased varying the polyacrylamide gradient in the gel. In fact, by changing the acrylamide concentration the pores that originate during the polymerization have different size. This is relevant for an optimal protein separation according to the molecular weight [10].

The 2D-PAGE–MALDI-TOF/TOF and LC-MS are complementary techniques [11] as shown also in the previous PhD thesis within this collaboration project "Bioanalytical solutions to characterize human proteome of clinical interest: healthy newborns and their early infections" by L. Tigli. For example, to investigate the hydrophobic proteins, the 2D-PAGE is an unsuitable technique and this approach is also limited for high molecular weight proteins. On the other hand, it allows the investigation of post-translational modifications and protein isoforms, so multiple protein spots can be identified as the same protein. In the present work, we considered the 2D-PAGE/MALDI TOF analysis the most suitable technique to investigate a complex biological fluid such as saliva. This approach permit the comparison between samples collected at different conditions and this is fundamental not only to analyse the general complex composition of saliva, but also to select a specific protein. Once that this selection is performed, a targeted approach LC-MS based can be exploit to obtain an accurate and reproducible quantification. The sensitivity in this case is dependent on the analyser of the mass spectrometer: as shown by Tigli, LTQ-orbitrap has a very low sensitivity so, in a discovery phase, this approach is comparable to 2D-PAGE/MALDI TOF. Instead, an accurate proteins' quantification can be carry out exploiting a Qtrap mass spectrometer that guarantees the highest sensitivity working in Multiple Reaction Monitoring (MRM).

During sample preparation, some important features of proteins such as solubility, size, charge, and isoelectric point must be considered for the success of 2D-PAGE analysis.

For this reason, the optimal preparation is to be determined empirically adjusting sample treatment, concentration of the reagents and reaction times.

For the solubilization of proteins urea is the most commonly used chaotropic agent. Thiourea can be added for the solubilization of insoluble proteins. Urea and thiourea act on hydrogen bonds and are used to prevent unwanted aggregation or formation of secondary structures that affect protein mobility.

Detergents (SDS, CHAPS, Triton X-100, octyl-glucoside) are added to disrupt hydrophobic interactions and increase solubility of proteins at their pI. These agents can be nonionic or zwitterionic, allowing proteins to migrate according to their own charges.

Ampholytes are also added to the sample just before the first dimension. These molecules have a double function: the first is to contribute to maintain a stable pH gradient, allowing proteins to be separated according to their isoelectric points, the second one is to keep proteins in solution during the IEF. Salt is a remarkable contaminant of biological samples for 2D-PAGE analysis, it limits the voltage that can be applied without producing high current, increasing the time required for protein focusing. For this reason, only a concentration less than 40 mM of salt is tolerated. To overcome this issue, salt is removed during the initial high-current stage of IEF but proteins that require salt for solubility may precipitate once it is removed. In this case ampholytes, that are specific for pH range, act counterbalancing insufficient salt concentration in a sample. Before IEF, reduction and an alkylation step must be performed. These two reactions are necessary to carry out protein denaturation and block cysteine reactivity to avoid the formation of artifacts during the whole 2D-PAGE analysis.

Isoelectric focusing

A thin layer of acrylamide gel on the top of a rigid support is the main component of an immobilized pH gradient (IPG) strip. Specific molecules called immobilines, derived from acrylamide, have additional ionizable non-amphoteric function. Immobilines of various pKa create an immobilized pH gradient inside the acrylamide gel. IPG strips are commercially available in a variety of pH ranges and lengths.
Sample loading onto an IPG strip can be performed in two different ways:

1. Rehydration loading, where strip rehydration and sample loading occur simultaneously. This method is the most common: the general procedure is easy and it is suitable for loading not only large quantities of proteins but also diluted samples. There are two rehydration methods:

- Passive rehydration: no electric field is applied during the rehydration step;
- Active rehydration: a low voltage is applied (usually 20 – 120 V) to facilitate the entry of proteins into the strip.

2. Cup loading, which is especially required to improve proteins' focalization on alkaline pH gradients. This technique provides a gradual sample loading onto the IPG strip (previously rehydrated with a solution of urea, thiourea, chaps and reducing agent) using a specific cup at the anodic side.

The application of an electric potential across the gel makes one end of IPG strip more positive than the other one. The net charge of a denatured protein is the sum of all the negative and positive charges of the side chains of its residues and the amino- and carboxyl- termini, therefore at all pH values except their isoelectric point, proteins are positively or negatively charged. Thus, when an electric field is applied proteins that are positively charged move towards the negative end of the IPG gel and, on the other hand, the negatively charged ones move towards the positive end of the gel.

A standard IEF protocol is generally composed of:

1. a low-voltage step to remove salts;
2. a progressive gradient to a high voltage to mobilize the ions (usually 3-4 one–hour steps with a progressive voltage increase);
3. a high-voltage step to complete the isoelectric focusing: this step is performed through the application of a constant number of Volt-hours (Vh);
4. A low voltage step to keep proteins focalized at their pI.

SDS-PAGE: second dimension

During the second dimension, proteins are separated according to their molecular weight. The continuity between IPG strip and the polyacrylamide gel of the second

dimension is guaranteed by a layer of 0.5% agarose. The presence of SDS in polyacrylamide gel is fundamental to make the electrical charge of the proteins uniform in the sample. SDS binds to the vast majority of proteins at a constant ratio (1.4 g of SDS per 1 g of protein), thus the intrinsic electrical charge of the proteins in the sample does not influence their mobility in the gel since the anionic SDS-protein complexes carry a constant net negative charge per unit of mass. The most commonly used running buffer for SDS-PAGE is the Tris-Glycine Laemmli buffer.

Protein staining

Coomassie Blue and silver nitrate are the two most important staining methods to reveal proteins in-gel. These two methods are dramatically different in sensitivity: 0.1 ng of protein can be detected with silver stain whereas 100 ng is the minimal detectable amount for Coomassie Blue. However, the detection is not only affected by the amount of protein, but also by the characteristics of the amino acids involved in the binding with the stain, in particular:

- Coomassie blue: in an acid buffer the solfonic acid group of Comassie Blue gives an ionic interaction with basic residues (arginine, lysine and histidine), but also Van der Waals forces and hydrophobic interactions are involved. The formation of the complex stabilizes the negatively charged anionic form of the dye thus producing the shift in colour from reddish-brown to blue. The blue form has an absorption maximum at a wavelength of 595 nm. Recently an MS compatible version of the dye has been developed.
- Silver nitrate: this is the most sensitive staining technique. Silver ions, derived from silver nitrate solution, when interacting with proteins (in particular carboxylic acid groups, imidazole, sulfhydryl and amines moiety) are reduced to metallic silver, obtaining a brown-black colour over the proteins. Silver stains use either glutaraldehyde or formaldehyde as the enhancer (solution necessary to ensure the specificity, the efficiency of the binding and the following reduction to metallic silver). This stain is not MS compatible, therefore a complete stain removal (usually using a solution of potassium

ferricyanide and sodium thiosulfate) is necessary before MS analysis to avoid the formation of silver-adducts [12].

Protein identification

Bottom-up approach is exploited for protein identification and characterization. Spots in the 2D-PAGE gel are manually collected and individually exposed to proteolytic cleavage by trypsin. Each peptide mixture is subsequently analyzed with MALDI TOF/TOF mass spectrometer and the data obtained are submitted to database search to carry out protein identification. The quality of the mass spectrum is dependent on experimental variables, which affect sensitivity and reproducibility such as the choice of matrix, intrinsic properties of the analyte, matrix-analyte relative concentrations, method of sample spot preparation on the MALDI plate and laser characteristics.

MALDI matrixes (Table 1-1) are often composed of acidic molecules, which act as "proton source" to promote the ionization of the analyte. Matrix molecules are characterized by various conjugated double bonds. This chemical feature allows the conversion of the photon energy absorbed from the laser into the excitation energy of the solid system. Nevertheless they are functionalized with polar groups, allowing their use in aqueous solutions and promoting the co-crystallization with the analyte.

MATRIX	STRUCTURE	MW	SOLVENT
2,5-dihydroxybenzoic acid (DHB)	O OH OH HO	154,1	ACN, ACETONE
Alpha-cyano-4-hydroxycinnamic acid (HCCA)	N O HO OH	189,2	ACN/H_2O, ACETONE
3,5-dimethoxy-4-hydroxycinnamic acid (SA)	O H_3CO OH HO OCH_3	224,2	ACN/H_2O

Table 1-1: most common matrix for MALDI analysis.

Sample preparation for MALDI analysis

The homogeneity of the sample-matrix mixture is a critical factor to obtain optimal sample ion yields and the best preparation for a given sample is to be determined empirically. The main sample preparation techniques are:

i. DRIED-DROPLET: this is the most widely used preparation in which the sample is mixed with matrix in a molar ratio between 1:1000 to 1:10000. A volume of saturated solution of matrix is mixed 1:1 with the sample solution and the mixture is then spotted on MALDI plate and allowed to dry. The dried-droplet method tolerates the presence of salts and buffers very well. The dried-droplet method is usually a good choice for samples containing more than one protein or peptide component. A common problem in the dried droplet method is the aggregation of higher amounts of analyte/matrix crystals in a ring at the edge of the drop. Normally these crystals are not homogeneous and are irregularly distributed. Another problem that may be associated with this method is the segregation: when the solvent evaporates and the matrix crystallizes, the salts and part of the analyte are excluded from the matrix crystals.

ii. FAST-EVAPORATION: a layer of matrix prepared in a highly volatile solvent is applied on the plate. The fast evaporation of the solvent promotes the formation of a thin homogeneous film of crystals of matrix. In a second time a drop of sample is applied over the first layer. The fast-evaporation method provides polycrystalline surfaces with a roughness 10-100 times lower than the dried-droplet method, thus the more uniform distribution of the analyte improves the sensitivity, resolution and mass measurement accuracy. A common problem associated with this method is a low spot-to-spot reproducibility, due to the fast evaporation of the solvent.

iii. TWO LAYERS: this technique is a variation of the fast-evaporation method. A layer of matrix prepared in a highly volatile solvent is applied on the plate but, in this case, the second layer is composed of a mixture of matrix and analyte solution. The addition of matrix to the second step provides improved results, particularly for proteins and mixtures of peptides and proteins. In particular, differently to the fast-evaporation method, the two layers preparation provides enhanced sensitivity and overall spot-to-spot reproducibility.

iv. SANDWICH: the sample is applied on top of a fast-evaporated matrix-only layer as in the fast-evaporation method, followed by the deposition of a second layer of matrix usually in a non-volatile solvent.

v. QUICK AND DIRTY: a drop of analyte is applied over the MALDI plate, then a drop of matrix solution is added on the top of it and mixed with the analyte. It is a fast method relevant for the analysis of in-plate protein digestions. Nevertheless, the results are not as reproducible or consistent as those obtained with the dried-droplet method.

Materials and methods

Reagents and tools

Chemicals

Acetic acid

Agarose

Ammonium bicarbonate (ABC)

3-[(3-cholamidopropyl)dimethylammonium]-1-propanesulfonate (CHAPS)

Formic acid (FA)

Iodoacetamide (IAA)

Dithiothreitol (DTT)

Glycerol

Sodium dodecyl sulfate (SDS)

Thiourea

Trifluoroacetic acid (TFA)

Tris•HCl

Tris(2-carboxyethyl)phosphine hydrochloride (TCEP•HCl)

Urea

Water, LC-MS Ultra CHROMASOLV, tested for UHPLC-MS

Potassium ferricyanide

Sodium thiosulfate

All these reagents are from SIGMA-ALDRICH

2D-PAGE and western blot

Bio-Lyte® ampholytes pH 3.9 – 5.1, pH 4.7 – 5.9, pH 7-10 and pH 5.5 – 6.7

Mineral oil

2X Laemmli Sample Buffer

10X Tris/Glycine Buffer

Bio-Safe Coomassie Stain

Silver Stain Plus kit

7 cm Focusing Tray

Tributylphosphine (TBP)

Protean IEF Cell

Mini-PROTEAN TGX Precast Gels

Mini-PROTEAN Tetra Cell

11 cm Focusing Tray

ReadyStrip IPG Strips, 11 cm, pH 4 – 7

ReadyStrip IPG Strips, 7 cm, pH 7-10

PROTEAN i12 IEF System

Trypsin from porcine pancreas

Any kD™ Criterion™ TGX™ Gel 11 cm

Dodeca™ electrophoresis cells system

ChemiDoc MP System

Cup loading Tray for the Protean® IEF Cell.

PVDF 0.2 µm membrane

TRANS-BLOT TURBO

2D-PAGE and western blot consumables and devices are all from BIO-RAD

Skim Milk Powder, Sigma-Aldrich

Tween 20, Sigma-Aldrich

Sodium chloride, Sigam-Aldrich

Goat antibody to alpha-1-antitrypsin, Abcam

3,3’,5,5’-Tetramethylbenzidine (TMB)

Solvents

Acetone

Ethanol

Acetonitrile

All the solvents are from SIGMA-ALDRICH.

Mass spectrometry

4800 MALDI-TOF/TOF, AB Sciex

α-cyano-4-hydroxycinnamic acid, LaserBio Labs

2,5-Dihydroxybenzoic acid, Sigma-Aldrich

PepMix4, LaserBio Labs, Sigma-Aldrich

ProteoMass Peptide MALDI-MS Calibration Kit, Sigma-Aldrich

Software

4000 Series Explorer, AB SCIEX

GPS Explorer, AB Sciex

Image Lab, BIO-RAD

PDQuest, BIO-RAD

www.matrixscience.com (online software)

Sample preparation

The extraction of proteins from swabs embedded of saliva was performed using a solution containing 40 mM TRIS, pH 7.4, and 1% SDS to keep proteins in solution. Swabs were sonicated for 3 minutes and vortexed, in ice. Low temperature ice avoids cleavage of proteins by proteases (some preliminary tests of extraction were performed adding protease inhibitors to the sample extraction buffer, but results were comparable to those obtained on ice). Proteins precipitation was carried out by adding ice-cold acetone to the sample in a 5:1 ratio and the precipitate was later resuspended in the solubilization buffer composed by 7 M urea, 2 M thiourea and 4% CHAPS. SDS-PAGE analysis was performed for each extracted sample and Coomassie Blue staining was used for proteins detection. Proteins were quantified using the Image Lab software (Bio-Rad®).

The same procedure, performed on blank swabs, did not detect any proteins indicating that swabs are not a source of protein contamination.

2D-Electrophoresis

Before the loading onto the IPG strip, proteins were treated with 65 mM DTT to reduce disulfide bonds for 45 minutes, and then Bio-Lyte ampholytes were added at 0.02% final concentration.

HEALTHY NEWBORNS (n=6): Two distinct pools were created, merging respectively the extraction solutions of samples collected from six healthy newborns within 24 hours after delivery (day 1) and between 24 and 48 hours after delivery (day 2). Pools were created using the same amount of protein from each sample and a total protein quantity of 40 µg was loaded onto each IPG strip (pH 4-7) using the passive rehydration technique. Two technical replicates were created both for day 1 and day 2 pools.

INFECTED NEWBORN (n=1): the extraction solutions of day 1 and day 2 were mixed and 40 µg of protein were loaded onto the IPG strip (pH 4-7) using the passive rehydration technique. Two technical replicates were performed.

PROTEAN® i12TM IEF Biorad® system was used to carry out the first dimension of 2D-PAGE analysis. Passive rehydration loading was performed before the mobilization and focusing steps as described in Table 1-2.

Step	Volt	Time
1	0	16 h
2	300	1 h
3	600	1 h
4	1000	1 h
5	4000	1 h
6	8000	1 h
7	8000 until 26000 V/h	

Table 1-2: protocol of focalization

After IEF, strips were rehydrated first with 1% DTT and then with 4% IAA in equilibration buffer (6 M urea, 30% glycerol, 1.5 M TRIS pH 8.8, 0.2% SDS) each time for 15 minutes. The second dimension was carried out with Any kD™ Criterion™ TGX™ Gel 11 cm in Dodeca™ electrophoresis cell system Biorad®. The running buffer used for the SDS-PAGE step was TRIS/Glycine.

Cup-loading

Cup loading method was performed for proteins analyzed in the pH range 7-10 (7 cm strip) with the following protocol:

- 12 hours of strip passive rehydration using 195 µl of rehydration buffer and 5 µl of TBP;
- 40 µg of proteins were loaded after reduction with TBP (40x diluted) and ampholytes (pH 7-10) were added (dilution 100x);
- Loading on the cup-loading tray.
- Execution of the IEF protocol (Table 1-3)
- Execution of the second dimension (as reported in paragraph "2D electrophoresis")

Step	Volt	Time
1	200	16 h
2	300	1 h
3	600	1 h
4	1000	1 h
5	4000	1 h
6	8000	1 h
7	8000 until 29000 V/h	

Table 1-3: cup loading IEF protocol

Protein detection

Bio-Safe Coomassie Stain (Biorad®) was used for a preliminary staining to reveal the most abundant proteins. Then the proteins were destained using a solution composed by 50% ethanol, 40% water, 10% acetic acid. Afterwards the Silver Stain Plus Kit (Biorad®) was used as more sensitive protein staining. Gel images were acquired with ChemiDoc™ XRS system (Biorad®).

Protein digestion

Spots of interest were manually excised. The silver destaining step was performed with a solution 1:1 of 30 mM potassium ferricyanide and 100 mM sodium thiosulfate. Spots were then dehydrated with acetonitrile and rehydrated with a trypsin-containing solution (20 µg/ml) for 30 minutes at room temperature. Peptides were extracted overnight with ammonium bicarbonate, 0.1% TFA for 45 minutes and an additional cycle of extraction was performed with acetonitrile-0.1% TFA in a 1:1 ratio for 45 minutes. The extracted solution was vacuum dried in a speedvac and the peptides were finally resupended in acetonitrile-0.1% TFA in a 1:1 ratio.

Mass spectrometry

Mass spectrometry analysis was performed using a 4800 Plus MALDI TOF/TOF™. Dried-droplet sample spot preparations were performed using alpha cyano-4-hydroxycinnamic acid (10 mg/ml in acetonitrile-0.1% TFA mix 70:30) as the matrix. Protein identification was carried out submitting the peptide mass fingerprinting spectra to Mascot database search. Also DHB (10 mg/ml in acetonitrile-0.1% TFA mix 70:30) in dried-droplet preparation was tested to compare this matrix with HCCA.

Western blot analysis

The protein transfer from the gel to the PVDF membrane was carried out by TRANS-BLOT TURBO (Biorad®) using the following protocol: 30 minutes, up to 1.0 A, 25 V. Western blot analysis was performed using a 0.2 µm membrane in PVDF (Biorad®). After the transfer, the membrane was treated for one minutes with TBS-T (20 mM Tris pH 7.5, 150 mM NaCl, 0.1% tween 20). Then the membrane was covered with a solution 3% of skim milk in TBS-T overnight. Three washing steps with TBS-T were then performed and finally membrane was incubated with the goat antibody to alpha-1-antitrypsin (ratio 1:5000) conjugated with horseradish peroxidase (HRP) in a 3% skim milk solution (incubation for two hours). The membrane was then washed with TBS-T. Protein detection was performed using 3,3’,5,5’-Tetramethylbenzidine (TMB) substrate.

Results and discussion

2D-electrophoresis: optimization of experimental set up

The optimization of 2D-PAGE analysis steps is required to deeper investigate saliva samples. Both during sample preparation and isoelectrofocusing the best preparation is to be empirically determined to increase protein solubility and, consequently, to obtain an optimal focalization according to the isoelectric point and the molecular weight. Due to the low total amount of proteins extracted from each cotton swab, the volume of resuspension buffer (urea, thiourea and chaps) added after acetone precipitation was optimized at 100 µl, thus obtaining a clear and, at the same time, concentrated solution. While the other steps of sample preparation are standard, it is crucial to adapt the protocol of focalization according to the sample features. Passive rehydration for 16 hours was applied for optimal sample loading into the IPG strip for analyses in the pH range 4-7, whereas both cup loading (Figure 1-2) and active rehydration loading were necessary for analyses in the pH range 7-10, confirming that in saliva the acidic proteins are considerably more abundant compared to the basic ones. For this reason, in the present work we focused the attention in the 4-7 pH range.

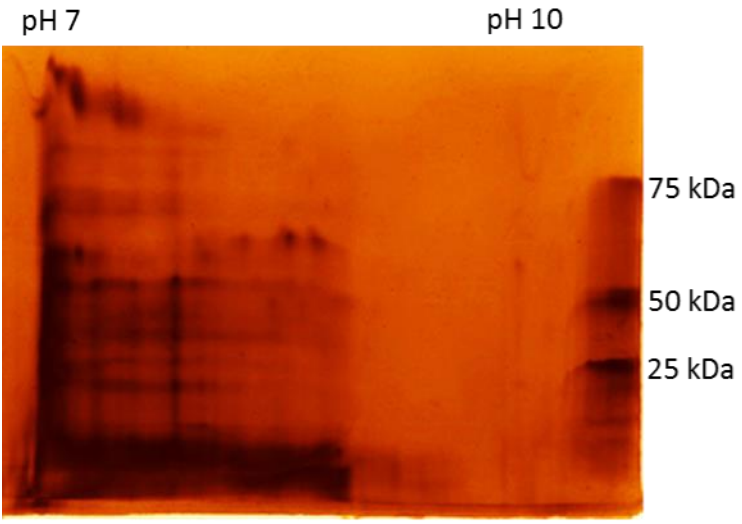


Figure 1-2: 7-10 pH gel using cup-loading. Silver staining.

Salts and electrolytes represents two of the most significant contaminants in saliva samples. Therefore, the optimization of the initial low-voltage steps is crucial in order to clean out these molecules and improve protein focalization. In particular, for this purpose, a 300 V step and subsequently a 600 V step were applied obtaining an almost complete salt removal. The following high voltage steps are necessary for protein mobilization towards their point of focalization. This movement is affected by several intrinsic features of the sample. For this reason, we determined empirically the most suitable voltage to apply. We varied the step 7 of the protocol reported in "Materials and Methods", increasing the voltage from 26000 to 30000 V/hours and we observed which voltage guaranteed the best protein focalization.

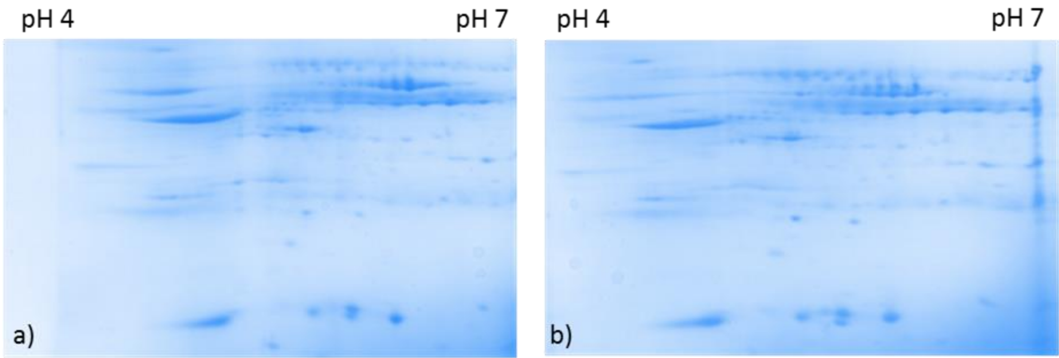


Figure 1-3: a) 11 cm gel, pH 4-7 obtained with 26000 V/hours; b) 11 cm gel, pH 4-7 obtained with 30000 V/hours; Coomassie blue staining.

Protein staining using Coomassie Blue did not show significant focalization differences between 26000 V/hours (Figure 1-3) and 30000 V/hours (Figure 1-3) gels. However, at pH values close to 7, spots appear to be better defined at 26000 V/hours gel, especially for high molecular weight proteins. Therefore, the 2D-PAGE analyses of both the pools of healthy newborns and the single infected newborn were performed using 26000 V/hours, which is also useful for decreasing the time of the analysis.

2D-electrophoresis: healthy newborn saliva

The clinical data of newborns whose samples were pooled to perform 2D-electrophoresis analysis of day 1 and day 2 are reported in Table 1-4.

NEWBORN	RACE	SEX	WEIGHT (g)	AGE (weeks)	DELIVERY	SAMPLING DAY 1	SAMPLING DAY 2
1061/2013	CAUCASIAN	F	3140	37	CAESAREAN	12 HOURS	33 HOURS
1062/2013	CAUCASIAN	M	2790	38	SPONTANEOUS	11 HOURS	32 HOURS
1056/2013	CAUCASIAN	M	4210	41	CAESAREAN	18 HOURS	39 HOURS
1588/2013	CAUCASIAN	M	3640	40	SPONTANEOUS	14 HOURS	38 HOURS
1645/2013	CAUCASIAN	F	4270	40	SPONTANEOUS	15 HOURS	39 HOURS
2201/2013	CAUCASIAN	M	2620	38	CAESAREAN	7 HOURS	32 HOURS

Table 1-4: clinical data of healthy newborns whose proteins were pooled to perform 2D-PAGE analysis.

For each sample, protein concentration was determined in order to create a final pool containing the same amount of protein from each newborn. Two technical replicates were performed for the pool obtained from 6 samples collected from healthy newborns within 24 hours after delivery (day 1), (Figure 1-4: A and B) and other two technical replicates were performed for the pool obtained from 6 samples collected from the same newborns between 24 and 48 hours after delivery (day 2) (Figure 1-4: C and D).

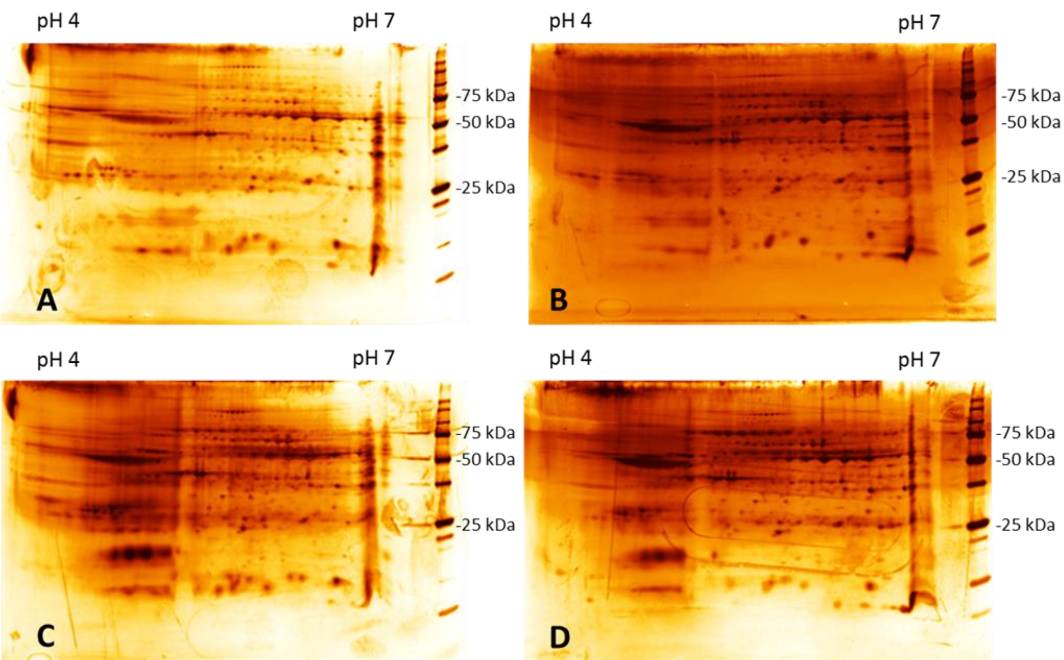


Figure 1-4: A and B are the technical replicates day1; C and D are the technical replicates day2; Silver staining.

First, optimization of sample staining was investigated. If the sample is composed of "standard" proteins, silver staining is more sensitive than Coomassie blue. Gel "A" in Figure 1-4, corresponding to the first technical replicate of day 1, was at first stained with Coomassie blue and, in a second time, with silver nitrate. Image analysis was performed and 192 spots were detected in silver staining while 138 were detected for Commassie blue staining, confirming that for the characteristics of these samples silver staining is the most suitable stain to highlight the highest number of proteins. Then, our goal was to define the standard composition of a healthy newborn salivary proteome, and the differences in number of spots and composition between day 1 and day 2 of saliva collection. In particular, after the creation of a master gel and its image analysis, 192 spots were detected on "day 1" gel and 195 on "day 2" gel using silver nitrate staining. So, no significant differences were observed in the total number of proteins comparing day 1 and day 2.

The sensitivity using the silver nitrate staining is approximately 0.1 ng and this is relevant to detect all the components of the protein pattern in newborn saliva. On the other hand, it is necessary to consider that this low amount of protein, even if detectable, is unlikely to produce reliable identifications by MS analysis. In fact, after tryptic digestion and MS analysis, the number of generated peptides in such extreme cases is usually insufficient to be confident for an identification. However, all the 2D-PAGE spots detected, starting from 40 µg of total protein amount, were collected and digested and 34 identifications were obtained. Mascot database search for each peptide mass fingerprint was carried out with a mass tolerance <10 ppm, according to MALDI calibration also performed to a value of 10 ppm of maximum outlier error.

Protein identifications in healthy newborn saliva

MALDI-TOF analysis was performed and the peptide mass fingerprint was obtained for each digested spot, giving the reliable identity of proteins upon defined requirements. In particular, the two most important criteria that we considered to define protein identity were: a number of identified peptides greater than three (two for low molecular weight proteins) and agreement between experimental and theoretical molecular weight

and isoelectric point. With the purpose to improve the MALDI spectrum quality and consequently the confidence in the identifications, several spot preparations were performed using different matrices. Serum albumin was taken as a reference protein to compare HCCA (10 mg/ml) and DHB (10 mg/ml) in dried-droplet preparation. These two matrices strongly differ in the quality of crystals formed onto the MALDI plate. This is crucial for the peptide ionization, and, therefore, for the quality of the spectrum and for the confidence of the protein identification. The serum albumin sequence coverage using HCCA was 72% with 50 matched peptides, whereas these values dramatically decrease to 36% and 21 respectively when using DHB. The same approach was also applied to another relevant protein, apolipoprotein A1. Also in this case the results obtained using HCCA were better than those obtained with DHB. In particular, 52% of sequence coverage and 20 matched peptides were observed with HCCA versus 25% of sequence coverage and 9 matched peptides using DHB. Therefore, after these preliminary experiments 10 mg/ml HCCA in acetonitrile-0.1% TFA mix 70:30 appears to be the most suitable matrix to allow the ionization of these analytes. Obtaining an optimal MALDI spectrum (regarding the signal intensity, number of m/z and the resolution) is the key to improve the protein sequence coverage and consequently the confidence in the identification.

Table 1-5 reports the identifications carried out using 10 mg/ml HCCA in acetonitrile-0.1% TFA mix 70:30.

PROTEIN	N. IDENTIFIED PEPTIDES	pI	MW kDa
Protein S100 A14	4	5.1	11
Actin Cytoplasmic 1	5	5.1	41
Ras related protein M-ras	5	8.82	23
Stratifin	10	4.68	27
Protein S100 A11	5	6.5	11
Transitional endoplasmic reticulum	15	5.1	89
Alpha lactalbumin	8	4.8	16
Thyroid hormone receptor associated protein 3	7	10	108

Alpha-1-acid glycoprotein 2	3	5.12	21
Leukocyte elastase inhibitor	6	5.9	42
78 kDa glucose-regulated protein	12	5	70
Protein S100 A8	21	6.5	10
Annexin A1	12	6.6	40
Prolactin-inducible protein	3	5.4	13
Protein S100-A9	5	5.7	13
Ig alpha-1 chain C	17	6.1	37
Small integral membrane protein	2	9.2	8
Actin Cytoplasmic 2	8	5.3	41
Keratin type 2 6C	6	6.3	57
Glutathione S-transferase	16	5.4	23
Keratin type 1 cytoskeletal 13	25	5	44
Interleukin-1 receptor antagonist	12	5.8	20
Heat shock protein beta 1	6	5.9	22
Ras related protein rab-15	2	5.5	24
FABP	2	6.6	15
Polymeric immunoglobulin receptor	4	5.6	83
Peroxiredoxin 6	5	6	25
Cystatin B	2	6.9	11
Serum albumin	50	5.9	67
Apolipoprotein A1	20	5.6	30
Serpin B3	10	6.3	44
Tubulin-folding cofactor B	2	5	27
Protein BEX1	2	6.3	15
Serotransferrin	11	6.7	75

Table 1-5: identified proteins in the acidic soluble fraction of whole saliva of healthy newborns, collected in the first 48 hours after the delivery.

The gel-based analysis performed on healthy newborn salivary samples led to the identification of a number of proteins linked to epidermal tissue functions and, more

interestingly, some proteins involved in inflammatory processes. In particular, we have identified: annexin A1 [13], heat shock proteins, S100 protein [14] (Table 1-5). 34 proteins were identified analysing a very low total amount of protein (40 μg) from healthy newborns, some of these proteins were already identified by other research groups such as the S100 proteins by Castagnola [8] or annexin A1 in adult by Neyraud [15]. Moreover, a study in 2005 [16] reported that the expression of glutathione S-transferase could increase in the presence of LPS, a component of membranes of gram-negative bacteria. Thus, a high expression of this protein might be indicative of microbial infection. This evidence suggests that newborns are "constitutively" in a stressed condition in the early 48 hours after birth.

The identified proteins were sorted according to isoelectric point, molecular weight and subcellular localization.

Isoelectric point

The pH ranges 4-7 of the IPG strip were partitioned in four sub-intervals (Figure 1-5). Approximately, 50% of the identified proteins exhibit a pI values between 5 and 6. An interesting observation is that 9% of proteins exhibit a pI higher than 7. In principle, the thyroid hormone receptor associated protein 3, small integral membrane protein and ras related protein M-ras should not be detectable in this pH range. We tentatively explain the presence of these proteins due to post-translational modifications that might cause a remarkable shift of pI, as observed for strongly glycosylated proteins and phosphorylated proteins [17].

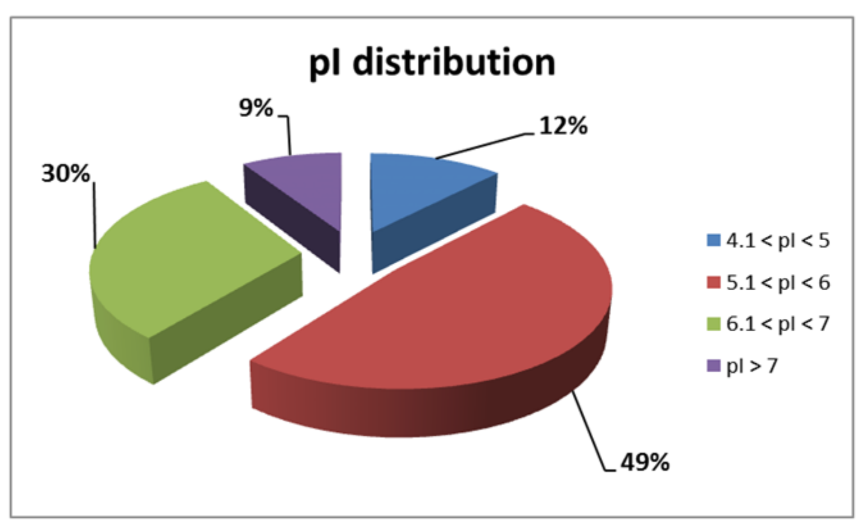


Figure 1-5: pI distribution of the identified proteins.

Molecular weight

All the molecular weight ranges were covered, but the highest number of identified proteins is between 10 and 50 kDa. Only 3% of proteins exhibits a MW higher than 100 kDa (Figure 1-6). The explanation might be attributable to the high dimension of these proteins that does not allow the penetration into the acrylamide pores of IPG strips. It is also relevant that some high molecular weight spots were not identified. Considering the staining intensity, these spots revealed a suitable amount of protein to reach an identification, but MALDI spectrum was poor of signals. The explanation might be attributable to the protein structure. In particular, some glycoproteins are physiologically linked with lipids and carbohydrates, and these modifications affect their susceptibility to the proteolytic cleavage [18]. Mucins, for example, are a very common class of strongly glycosylated salivary proteins. Considering also isoelectric point and molecular weight, these not-identified intense spots might be own attributable to mucins.

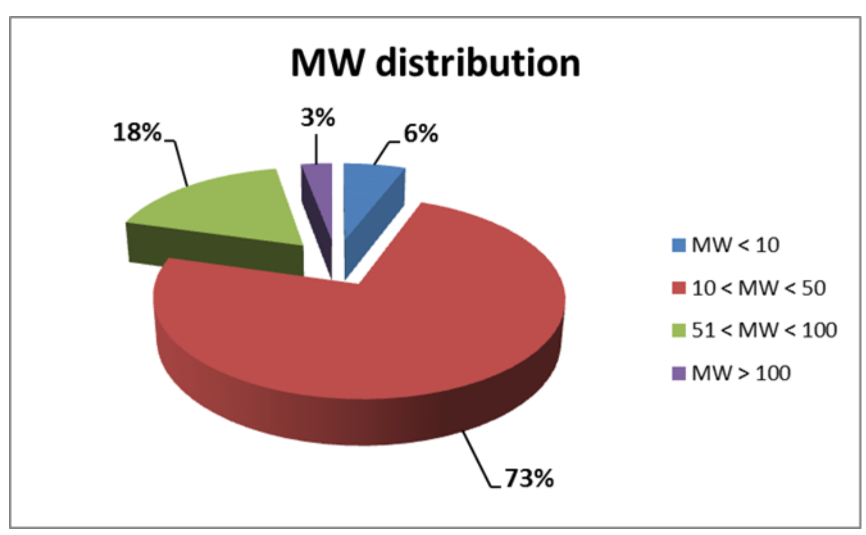


Figure 1-6: molecular weight distribution of the identified proteins.

Subcellular location

Proteins that have high hydrophobic amino acid content on the surface, such as membrane proteins, have low solubility in aqueous solvents. However, urea, thiourea and CHAPS guarantee a high solubility and the detection also of these proteins. Approximately, 50% of the identified proteins are secreted or cytoplasmic. For our purpose, both these two subcellular locations are relevant. In fact, the "blood exchange" may potentially enrich the secreted group of inflammatory proteins and, on the other hand, the cytoplasmic proteins (but also the membrane proteins) might be representative of tissue damage (Figure 1-7).

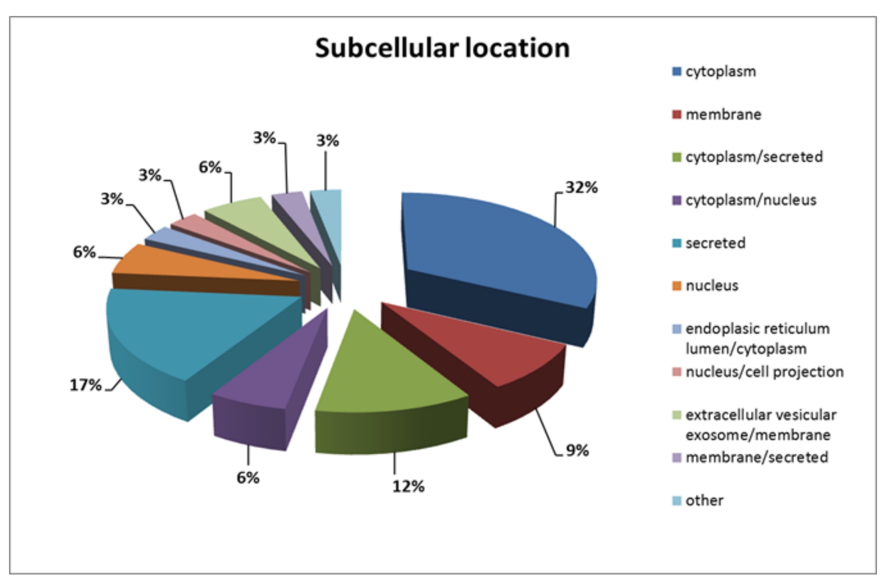


Figure 1-7: subcellular location of the identified proteins.

Differential proteomics of healthy newborns: protein expression at 24 hours and 48 hours

The comparison of proteins present at 24 and 48 hours from delivery was carried out exploiting the image analysis. In particular, qualitative and quantitative differences in protein expression were observed comparing different saliva samples. At first, we aimed to discover which proteins appeared to be differently expressed in day 1 compared to day 2. For this purpose, two master gels were created, one with all the spots of the day 1 technical replicates and the other one with all the spots of the day 2 technical replicates. The comparison of the master gels provided information about both overexpressed proteins and proteins only detectable in day 1 or day 2. The identified proteins which qualitatively or quantitatively vary, according to the day of collection are reported in Table 1-6.

PROTEIN	VARIATION
Ras related protein rab-15	ONLY DAY 2
Alpha-1-acid glycoprotein	TWO FOLD UP DAY 1
Transitional endoplasmic reticulum	ONLY DAY 1
78 kDa glucose regulated protein	ONLY DAY 1
Thyroid hormone receptor-associated protein	ONLY DAY 1
Ras related protein M-ras	ONLY DAY 1
Tubulin- folding cofactor B	ONLY DAY 1
Protein BEX 1	TWO FOLD UP DAY 2
Alpha lactalbumin	TWO FOLD UP DAY 2
Protein S100-A9	ONLY DAY 2
Protein S100-A14	ONLY DAY 2
Stratifin	TWO FOLD UP DAY 2
Prolactin inducible protein	TWO FOLD UP DAY 2
Small integral membrane protein	TWO FOLD UP DAY 2
Leukocyte elastase inhibitor	TWO FOLD UP DAY 1
Annexin A1	ONLY DAY 2

Table 1-6: proteins differently expressed between day1 and day2 of saliva collection.

Figure 1-8: charts with differently expressed proteins between day1 and day2 of saliva collection. “Only day 1” or “Only day 2” = 100%; “Two fold up day 1” or “Two fold up day 2” = 50%.

Observing the Table 1-6 it is interesting to consider the over-expression of the alpha-1-acid glycoprotein in day 1. The detection of this acute phase protein [19] confirms the stressed condition of newborns especially during the first 24 hours after delivery. Two of the four identified S100 [14] proteins (A9 and A14) were detected only in day 2. These proteins are involved in the regulation of protein phosphorylation, transcription factors, Ca^{2+} homeostasis, the dynamics of cytoskeleton constituents, enzyme activities, cell growth and differentiation, and especially in the inflammatory response. Also the presence of the protein annexin A1 only in day 2 is remarkable: this protein suppresses phospholipase A2, blocks eicosanoid production and inhibits various inflammatory events. In other words, the expression of this protein in day 2 might be considered part of the feedback mechanism to turn off the inflammatory condition caused by delivery.

Healthy and infected newborns

The saliva sample of the potentially infected newborn (Table 1-7) was compared with the results obtained from the analysis of saliva samples collected from healthy newborns. 40 µg of proteins were loaded on the gel for the infected newborn. In this case, day 1 and day 2 were mixed to obtain a single sample analyzed in two technical replicates. Clinically, this infant was classified as “potentially infected” because positive for the streptococcus bacteria at delivery.

NEWBORN	RACE	SEX	WEIGHT (g)	AGE (weeks)	DELIVERY
482	CAUCASIAN	F	3740	41	SPONTANEOUS

Table 1-7: clinical data of the infected newborn.

The analysis of the protein patterns showed significant differences in composition and in the number of spots. Above all, two intense spots were detected only in the potentially infected newborn (Figure 1-9);

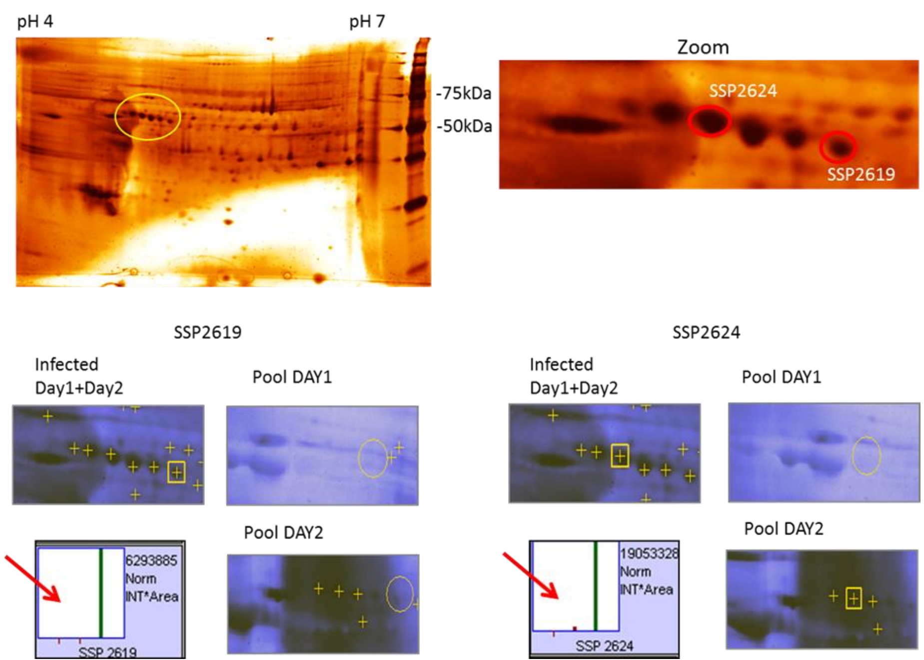


Figure 1-9: image analysis results comparing healthy and infected newborn saliva. SSP2619 and SSP2624 were detected only in the infected sample (green chart) corresponding to the protein alpha-1-antitrypsin.

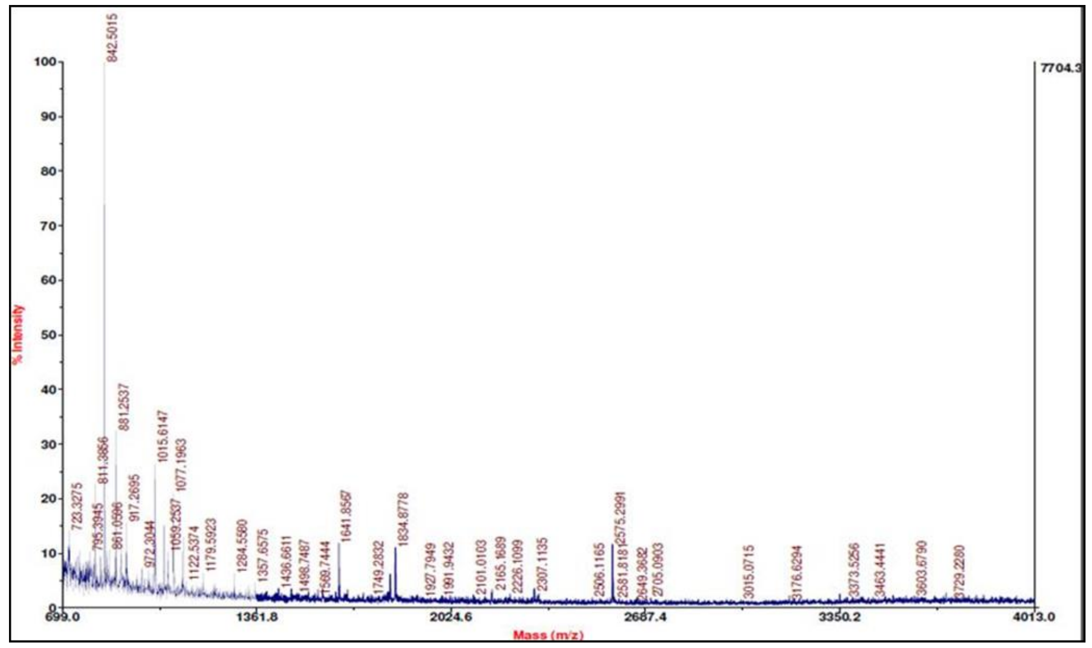


Figure 1-10: MALDI spectrum of alpha-1-antitrypsin identification.

Based on the image analysis, these two spots were collected and identified as the protein alpha-1-antitrypsin (A1AT) (Figure 1-10). This is a positive acute phase protein

that acts with a negative feedback on inflammatory response. As all the positive acute phase proteins, A1AT serum levels increase during inflammation. This feature is especially relevant for diagnostic applications. The main function of A1AT is to inhibit neutrophil elastase but, when exploited in therapy, A1AT has also been shown to reduce neutrophil infiltration into kidneys during ischemia, to reduce neutrophil infiltration during acute and chronic inflammation and to directly inhibit the activity of caspase-3 (a protease implicated in the apoptosis) [20]. Furthermore, as reported in the second chapter, A1AT is also relevant for its several allelic variants, which could represent a risk factor for emphysema. Considering the blood-saliva exchange, which takes place during saliva production, the proteins that follow this pathway might be representative to characterize a systemic inflammatory condition. For this purpose, it is necessary to investigate that the identified A1AT is not due to a sample contamination. Indeed, A1AT is a common protein in infants and is detectable not only in blood, but also in maternal colostrum. If a blood contamination is relatively easy to avoid just analysing the salivary swabs, a contamination of maternal origin is considerably more complicated to discern. Fortunately, the image analysis provides an answer. If we consider alpha lactalbumin, a constitutive milk protein, we identified it both in infected and healthy saliva samples and no significant amount difference was observed comparing the infected newborn and the pool of healthy newborns. Thus, this indicates that the increased level of alpha-1-antitrypsin was not due to a contamination, but to an inflammatory condition.

Alpha-1-antitrypsin in healthy newborns

It is known that alpha-1-antitrypsin is constitutively expressed both in adults [21, 22] and infants [23]. For this reason a deeper investigation of the differential expression of this protein between healthy and infected newborns was required. For this purpose, we considered the immune-based approach that is a sensitive and specific technique. Western blot analysis confirmed the presence of alpha-1-antytrypsin also within healthy newborn saliva (Figure 1-11).

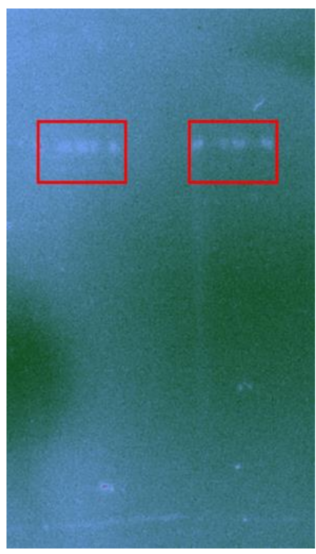


Figure 1-11: results of western blot analysis. Highlighted the two bands corresponding to alpha-1-antitrypsin in healthy newborn saliva.

A1AT quantification in healthy newborns saliva represents the starting point in order to estimate its differential expression during inflammation. For this purpose, different amounts of A1AT standard were analysed using the western blot approach obtaining a calibration curve that allowed to estimate a A1AT concentration of 35 ng/µl for the two healthy newborns.

Conclusions

In the present work, we performed the evaluation of newborn saliva as a potential diagnostic tool to diagnose a septic condition in the early hours after delivery. In particular, the standard protein pattern in healthy newborn saliva was characterized indicating that the first 48 hours after delivery represent a stressed phase in newborn life, supported by the detection of proteins related to inflammatory states such as heat shock proteins, annexin A1, alpha-1-acid glycoprotein and the S100 proteins. The identified proteins are considerably heterogeneous for isoelectric point, molecular weight and subcellular location. Image analysis was exploited to understand the differentially expressed proteins between the two different days of collection. 16 out of the 34 identified proteins appear to have a different concentration in day 1 compared to day 2. Alpha-1-antitrypsin was detected differentially expressed comparing healthy and infected newborn saliva. However, it should point out that alpha-1-antitrypsin, differently for example from C-reactive protein, does not represent a clinically recognized marker of inflammation. For this reason, the present study should be considered as a starting point for following investigations where the detection of A1AT salivary levels will be coupled with the detection of other inflammatory signals. The LC-MS approach might be exploited in a high throughput manner, instead of western blot analysis, to quantify A1AT in newborn saliva. This is extremely important when one considers clinical applications involving the screening of hundreds of thousands of patients. In fact, it is possible to generate multiple combinations of protein patterns with a powerful diagnostic impact.

Bibliography

1) Humphrey, S. P., & Williamson, R. T. (2001). A review of saliva: normal composition, flow, and function. The Journal of prosthetic dentistry, 85(2), 162-169.
2) De Almeida, P. D. V., Gregio, A. M., Machado, M. A., De Lima, A. A., & Azevedo, L. R. (2008). Saliva composition and functions: a comprehensive review. J Contemp Dent Pract, 9(3), 72-80.
3) Al Kawas, S., Rahim, Z. H., & Ferguson, D. B. (2012). Potential uses of human salivary protein and peptide analysis in the diagnosis of disease. Archives of oral biology, 57(1), 1-9.
4) Deepa, T., & Thirrunavukkarasu, N. (2010). Saliva as a potential diagnostic tool. Indian journal of medical sciences, 64(7), 293.
5) Majid S, Shafi M. (2014) Saliva as a diagnostic tool: a review. J Pharm Biol Sci.;9:9–14
6) Huq, N. L., Cross, K. J., Ung, M., Myroforidis, H., Veith, P. D., Chen, D., ... & Reynolds, E. C. (2007). A review of the salivary proteome and peptidome and saliva-derived peptide therapeutics. International Journal of Peptide Research and Therapeutics, 13(4), 547-564.
7) Ben-Aryeh, H., Lapid, S., Szargel, R., Benderly, A., & Gutman, D. (1984). Composition of whole unstimulated saliva of human infants. Archives of oral biology, 29(5), 357-362.
8) Castagnola, M., Inzitari, R., Fanali, C., Iavarone, F., Vitali, A., Desiderio, C., ... & Messana, I. (2011). The surprising composition of the salivary proteome of preterm human newborn. Molecular & Cellular Proteomics, 10(1), M110-003467.
9) O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. Journal of biological chemistry, 250(10), 4007-4021.
10) Berkelman, T., Stenstedt, T., Bjellqvist, B., Laird, N., McDowell, M., Olsson, I., & Westermeier, R. 2D Electrophoresis Using Immobilized pH Gradients: Principles and Methods, Amersham Biosciences, Piscataway, NJ, 1998. Ključne reći.

11) Chandramouli, K., & Qian, P. Y. (2009). Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. Human Genomics and Proteomics, 1(1), 239204.

12) Imai, B. S., & Mische, S. M. (1999). Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. Electrophoresis, 20, 601-605.

13) Perretti, M., & Dalli, J. (2009). Exploiting the Annexin A1 pathway for the development of novel anti-inflammatory therapeutics. British journal of pharmacology, 158(4), 936-946.

14) Donato, R., Cannon, B. R., Sorci, G., Riuzzi, F., Hsu, K., Weber, D. J., & Geczy, C. L. (2013). Functions of S100 proteins. Current molecular medicine, 13(1), 24.

15) Neyraud, E., Sayd, T., Morzel, M., & Dransfield, E. (2006). Proteomic analysis of human whole and parotid salivas following stimulation by different tastes. Journal of proteome research, 5(9), 2474-2480.

16) Xue, B., Wu, Y., Yin, Z., Zhang, H., Sun, S., Yi, T., & Luo, L. (2005). Regulation of lipopolysaccharide-induced inflammatory response by glutathione S-transferase P1 in RAW264. 7 cells. FEBS letters, 579(19), 4081-4087.

17) Locke, D., Koreen, I. V., & Harris, A. L. (2006). Isoelectric points and post-translational modifications of connexin26 and connexin32. The FASEB journal, 20(8), 1221-1223.

18) Van der Post, S., Subramani, D. B., Bäckström, M., Johansson, M. E., Vester-Christensen, M. B., Mandel & Hansson, G. C. (2013). Site-specific O-glycosylation on the MUC2 mucin protein inhibits cleavage by the Porphyromonas gingivalis secreted cysteine protease (RgpB). Journal of Biological Chemistry, 288(20), 14636-14646.

19) Fournier, T., Medjoubi-N, N., & Porquet, D. (2000). Alpha-1-acid glycoprotein. Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology, 1482(1), 157-171.

20) Al-Omari, M., Korenbaum, E., Ballmaier, M., Lehmann, U., Jonigk, D., Manstein, D. J., ... & Janciauskiene, S. (2011). Acute-phase protein α1-antitrypsin inhibits

neutrophil calpain I and induces random migration. Molecular Medicine, 17(9-10), 865.

21) Janciauskiene, S. M., Bals, R., Koczulla, R., Vogelmeier, C., Köhnlein, T., & Welte, T. (2011). The discovery of α1-antitrypsin and its role in health and disease. Respiratory medicine, 105(8), 1129-1139.

22) Pederson, E. D., Stanke, S. R., Whitener, S. J., Sebastiani, P. T., Lamberts, B. L., & Turner, D. W. (1995). Salivary levels of α 2-macroglobulin, α 1-antitrypsin, C-reactive protein, cathepsin G and elastase in humans with or without destructive periodontal disease. Archives of oral biology, 40(12), 1151-1155.

23) Munjal, Y. P., & Sharma, S. K. (2012). API textbook of Medicine. JP Medical Ltd.

2- Characterization of the acute phase protein alpha-1-antitrypsin using LC-MS platforms: marker of inflammation in newborn saliva and “F” allelic variant

Abstract

Alpha-1-antitrypsin (A1AT), a protein associated with inflammation, has been detected within newborns human saliva. A 2D gel-based proteomic investigation performed on saliva samples collected according to an approved study protocol during the first 48 hours after birth from healthy newborns and potentially infected newborns, highlights that the proteomic pattern shows significant differences. It is known that saliva is a complex biological fluid composed of many substances including proteins. The use of saliva for diagnostic purposes is common. However, to our knowledge, the identification of systemic inflammatory proteins within newborn saliva has never been reported so far. Poor information is often available on newborn saliva protein content and highly sensitive analytical platform are required to deal with very small amount of proteins. A non-invasive diagnostic LC-MS-based assay that identifies and quantifies inflammatory proteins within saliva could be particularly useful for newborns. Alpha-1-antitrypsin (A1AT) is associated with proteome of newborns with risk of inflammation. This protein is a protease inhibitor belonging to the serpin superfamily and a component of the “acute phase proteins”. It protects tissues from enzymes of inflammatory cells, especially neutrophil elastase, and is responsible for negative feedback on the inflammatory response. Furthermore, the “F” allelic variant of A1AT resulting from a single amino acid substitution of arginine by cysteine in the position 223 has a lower affinity for neutrophil elastase. For this reason, this specific variant might represent a risk factor for the development of lung and liver diseases, in particular when coupled with Z allele (PiFZ heterozygote phenotype). Therefore, the development of a targeted LC-MS based method to monitor A1AT is relevant to establish a condition of inflammation in newborns. The same LC-MS method is also useful to investigate the presence of the F allelic variant in biological samples and to predict potentially pulmonary diseases. Several tools were

exploited to characterize the “F” variant (liquid chromatography, ion mobility, MS/MS analysis and MRM). Furthermore, in the present work we also performed a deep investigation of the matrix effect observed for the detection of the A1AT.

Introduction

In the previous work, we characterized the salivary proteome of healthy and infected newborns during the first 48 hours after the delivery. Then, comparing healthy newborns salivary proteome and infected newborns salivary proteome, an acute phase protein was detected differently expressed in the infected saliva sample. Acute-phase proteins (APPs) are a class of proteins whose plasma concentration change during an inflammatory condition. In response to trigger event, local inflammatory cells (neutrophil granulocytes and macrophages) secrete several cytokines: interleukins IL1, IL6, IL8 and TNFα. In liver, these cytokines promote the synthesis of “positive” acute-phase proteins [1] whose plasma concentrations increase during the inflammation, but there are also “negative” acute-phase proteins which concentration decrease [2] (Figure 2-1).

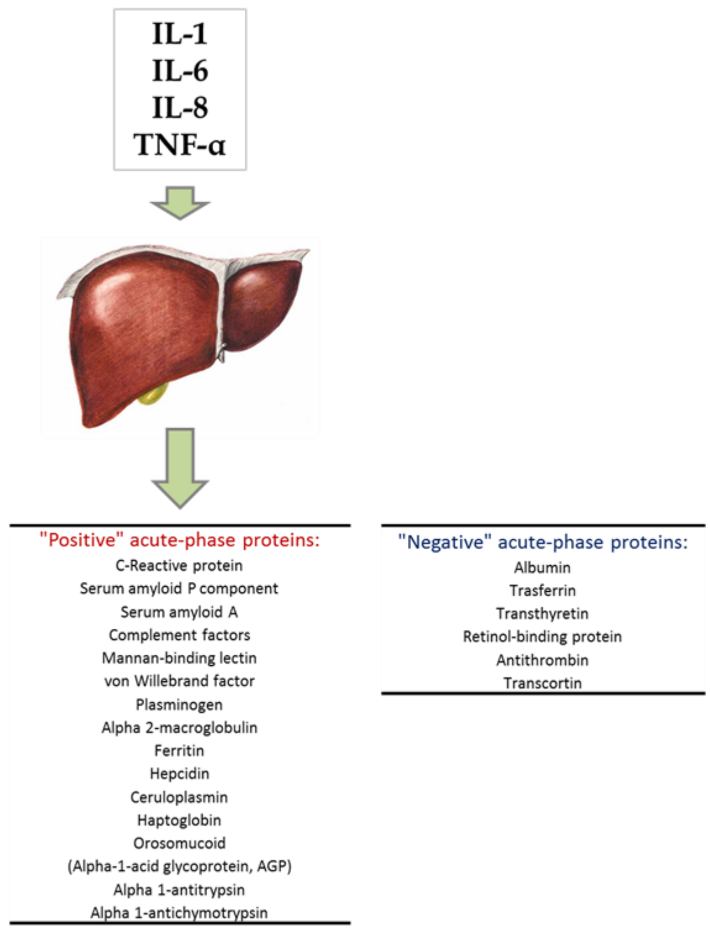


Figure 2-1: positive and negative acute phase proteins.

Positive acute-phase proteins are classified in two subgroups according to physiological functions. The first subgroup usually acts directly to destroy the triggering event, for example the C-reactive protein, mannose-binding protein, complement factors, ferritin, ceruloplasmin, serum amyloid A and haptoglobin. The second subgroup instead acts with a negative feedback on the inflammatory response turning off the inflammation and repairing the damaged tissues. In particular, some of these proteins belong to the "serpin" family [3].

Serpins

The word "SERPIN" is an acronym derives from serine protease inhibitor. Initially, in 1980, the provisional name was "ovalbumin-antithrombin superfamily" because of the sequence similarity between two proteins, chicken ovalbumin and human antithrombin III. Nevertheless, a sequence homology between antithrombin III and another human protease inhibitor, the alpha-1-antitrypsin, were also discovered. Unlike other serine protease inhibitor family, serpins have a long and flexible "reactive centre loops" (RCL), typically composed of 20-24 residues. The native serpin is composed of approximately 400 amino acids that fold into an N-terminal helical domain and a C-terminal b-barrel domain. The RCL and five central β-sheets (β-sheet A) are the two most important structures for the function of serpins. In particular, the incorporation of RCL into β-sheet makes serpin structure hyperstable. This conformational change can occur upon proteolytic cleavage and, especially, the rapidity of RCL incorporation obtaining a stable conformation distinguished the inhibitory (alpha-1-antitrypsin) from the non-inhibitory serpins [4].

Alpha -1-antitrypsin: an acute phase protein involved in the onset of emphysema

An accurate quantification of acute phase proteins in saliva could represent a non-invasive diagnostic tool to characterize an inflammatory condition at its early stages. In the present work, we aim to characterize the alpha-1-antitrypsin that we detected differently expressed in infected newborn saliva, as inflammatory signal in infants.

Moreover, several LC-MS platforms will be exploit to describe the "F" allelic variant of this protein and to discriminate it from the normal one in biological samples. Alpha-1-antitrypsin (A1AT) is a single-chain polypeptide consisting of 394 [1,5] (418 amino acid with a 24 amino acid signal peptide) amino acids in the mature form and exhibits a number of glycoforms. In fact, the 15% of the molecular weight (about 52 kDa), is composed of carbohydrates. These carbohydrates are linked to the proteins by arginine residues at the positions 46, 83 and 247 [5]. In particular, the complex with the negatively-charged sialic acid contribute to generate A1AT heterogeneity. This condition is well observed in 2D-electrophoresis analysis [6,7] (Figure 2-2).

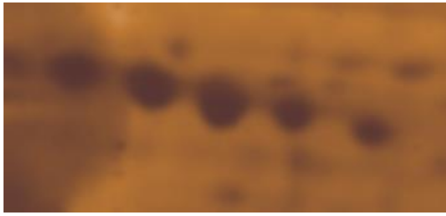


Figure 2-2: A1AT heterogeneity for isoelectric point and molecular weight in 2D PAGE analysis.

The structure is approximately composed by 30% of helical and 40% of β-sheet (Figure 2-3). A methionine in the position 358 of the RCL is a crucial residue for the inhibitory function. Indeed, Met358 and the adjacent Ser359 form a pocket that precisely fits the active site of the serine proteases [5].

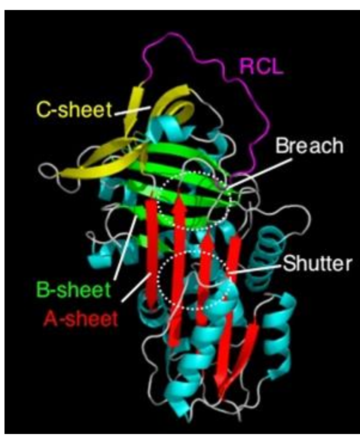


Figure 2-3: The X-ray crystal structure of native human antitrypsin. Five-stranded A-sheet is in red, six-stranded B-sheet in green, and four-stranded C-sheet in yellow. α-helices are shown in cyan. The RCL is at the top of the molecule in magenta.

It is relevant to consider that, in culture, monocytes are able to secrete three different forms of A1AT: the native one (52 kDa) and other two forms complexed with serine proteases (66 kDa and 75 kDa). During complex dissociation, the molecular weight of A1AT is reduced to less than 50 kDa due to the release of the small C-terminal peptide Ser359-Lys39 [5].

Function

A1AT protects tissues from enzymes of inflammatory cells, especially from the neutrophil elastase, proteinase 3, and other serine proteases released from activated human neutrophils during an inflammatory response [1,3,5,7,8]. As all the positive acute phase proteins, circulating A1AT levels increase rapidly (3 to 4 fold) in response to inflammation or infection [3]. Nevertheless, in response to inflammatory cytokines (IL-6, IL-1 and TNFα) and endotoxins, blood monocytes and alveolar macrophages with a local synthesis can contribute to increase tissues A1AT levels as much as 11-fold [1].

"F" variant

The "F" allelic variant of A1AT results from a single amino acid substitution of arginine by cysteine in the position 223 (a single nucleotide substitution of cytosine to thymidine which results in the amino acid exchange) [9]. Structurally, the insertion of the second cysteine (the normal A1AT has only one cysteine at the position 232) may influence the three-dimensional conformation of the protein or allows the interaction with other serum proteins with free-SH groups. It is not likely the interaction between the Cys223 and the Cys232 to produce a disulphide bridge: the distance between these two residues is 23 Å, when the formation of a disulphide bond takes place for residues distance between 4 and 9 Å. The clinical relevance of F variant is the lower affinity for neutrophil elastase that might represent a risk factor to development lung and liver diseases, in particular when coupled with Z allele (PiFZ heterozygote phenotype). The allelic frequency of the F variant is relatively uncommon, approximately 0.002 in

Caucasian population. Moreover, the serum level and the function associated with F allele are at least the 80% of those of the normal M allele [10].

Development of a LC-MS based method for the investigation of an inflammatory condition in the early stages of newborns life

When a LC-MS based method development starts, a few fundamental concepts and instrumental properties need to be taken into consideration. First of all, the detector sensitivity, that is the smallest amount of analyte that can be detected at a certain confidence level, which depends on the chemical physical properties of the molecule of interest. Then, the resolving power, that represents the ability to distinguish between ions differing in the ratio m/z by a small increment. Resolution is expressed as m/Δm, where m is the mass of an ion and Δm is the width of the peak at half height (Figure 2-4).

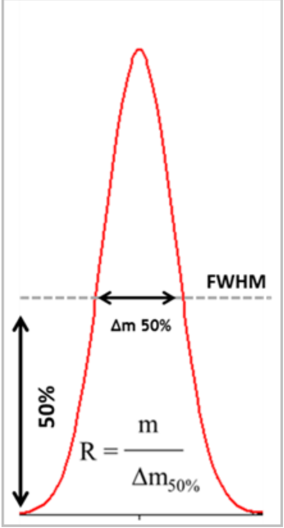


Figure 2-4: scheme of peak resolution.

Different mass spectrometers allow answering different questions. In the present work, UPLC coupled to Synapt G2-S has been exploited for its resolving power (40000 FWHM) to obtain the accurate mass measure of the A1AT peptide of interest and the distinction in between wild type and F variant. Furthermore, the ion mobility function is a powerful tool to investigate the differences of molecules (in our case, the peptides) according to molecular shape. This technique has the capability of separating ions with the same m/z but with different collision cross sections and\or charge states by utilizing

the mobility of an ion in a background gas under the influence of an electric fields. For example, a relevant applicability of this technique is to monitor the copopulated conformational states during the protein folding, that is relevant to elucidating the molecular mechanisms involved in protein aggregation of some diseases [11]. In the present work this approach was evaluated in order to strengthen the identification of the peptides of interest belonging to the wild type and to the F variant. A more sensitive mass spectrometer is required for the quantification of the analytes [12]. For this purpose, HPLC coupled to 4000 Qtrap and UPLC coupled with 5500 Qtrap were exploited in multiple reaction monitoring (MRM) for the quantification of the peptides of interest. Analytical MS-based methods present some undeniable advantages over immunochemical methods: high specificity, throughput multiplex analysis capable to monitor several analytes at the same time, fast method development and cheaper analyses. In the previous PhD thesis within this collaboration project "Bioanalytical solutions to characterize human proteome of clinical interest: healthy newborns and their early infections", L.Tigli has shown the importance of exploiting the most suitable mass spectrometer during the different steps of a LC-MS method development. In particular, targeted analyses were focused on the development of an analytical method based on liquid chromatography coupled to mass spectrometry to qualitatively and quantitatively assay C-reactive protein, procalcitonin and calcitonin within salivary samples. Tigli also investigates the properties of newborn whole saliva samples as biological matrix for quantitative MS-based assays. The experiments performed by Tigli on salivary samples showed the presence of a matrix effect in the determination of the proteins of interest. "Matrix effect" has represented a considerable issue in the development of a reproducible LC-MS based method. In the present work, the detection of the A1AT has been performed after digestion with LysC overcoming a matrix effect.

Materials and methods

Chemicals

Ammonium bicarbonate (ABC)

Acetone

Formic acid (FA)

Iodoacetamide (IAA)

Dithiothreitol (DTT)

Endoproteinase Lys-C from Lysobacter enzymogenes

Sodium dodecyl sulfate (SDS)

PEG400

Acetonitrile

Trizma® base

Water, LC-MS Ultra CHROMASOLV, tested for UHPLC-MS

All these reagents are from SIGMA-ALDRICH

Protein and peptide standard

α1-Antitrypsin from human plasma, Sigma-Aldrich

“F” peptide (VPMMKCLGMFNIQHCKKLSSWV) synthetized by Selleck Chemicals

Bovine serum albumin, Sigma-Aldrich

Porcine haemoglobin, Sigma-Aldrich

Column

Phenomenex 150x2.10 mm, 3.6u

Software

MassLynx V4.1

BiopharmaLynx 1.3.2

Analyst® AB Sciex 1.5

Analyst® AB Sciex 1.5.1

LC-MS instruments

Acquity UPLC Waters®-Synapt G2-S Waters®

Agilent®1100 HPLC-4000 Qtrap, AB Sciex®

Acquity UPLC Waters®-5500 Qtrap, AB Sciex®

LysC activity

Melittin from honey bee venom, Sigma-Aldrich

N-p-Tosyl-Gly-Pro-Lys 4 nitroanilide, Sigma-Aldrich

Cary Series UV-Vis Spectrophotometer, Agilent Technologies

Sample preparation: digestion of A1AT standard

A1AT standard was dissolved in 50 mM ABC and 0.5% SDS. Before the digestion, protein was denatured with 5 mM DTT at 95°C for 5 minutes and then protein was alkylated with 10 mM IAA for 20 minutes at room temperature. 100 mM ABC pH 8.3 was added to reach the final digestion volume. Endoproteinase Lys-C from Lysobacter enzymogenes was added in ratio 1:50 to the protein, and digestion was performed for 18 hours at 37°C.

This digestion was performed in different variants:

- In absence of bovine serum albumin;
- In presence of bovine serum albumin, added in ratio 1:15 [13] with the A1AT, that is approximately the ratio of the plasma concentration;
- In absence of bovine serum albumin but adding the synthetic “peptide F” synthetized by Selleck Chemicals;
- In presence of bovine serum albumin and adding the synthetic “peptide F” synthetized by Selleck Chemicals;
- In presence of PEG400, same ratio of bovine serum albumin;
- In presence of porcine haemoglobin, same ratio of bovine serum albumin;

Sample preparation: digestion of melittin

Melittin was dissolved in ultrapure water. Before the digestion, protein was denatured with 5 mM DTT at 95 °C for 5 minutes and then protein was alkylated with 10 mM IAA for 20 minutes at room temperature. 100 mM ABC pH 8.3 was added to reach the final digestion's volume. Endoproteinase Lys-C from Lysobacter enzymogenes was added in ratio 1:50 to the protein, and digestion was performed for 18 hours at 37 °C.

This digestion was performed in two different variants:

- In absence of bovine serum albumin;
- In presence of bovine serum albumin, same ratio of the A1AT's digestion;

Spectrophotometric assay

The spectrophotometric assays were performed using Cary Series UV-Vis Spectrophotometer. 405 nm, l=0.2 cm. The volumes of the assays are reported in Table 2-1. Buffer: 25 mM Tris-HCl pH 7.7.

total volume (µl)	substrate (µg)	LysC (µg)	BSA (µg)
300	93,75	1,3	59
300	93,75	1,3	0
300	93,75	0,15	59
300	93,75	0,15	0

Table 2-1: reagents volume for the spectrophotometric assays.

Liquid chromatography

Phase A: 0.1% formic acid

Phase B: 0.1% formic acid in acetonitrile

Time (min)	*% PHASE A*	*% PHASE B*	*FLOW (ml/min)*
initial	99	1	0,2
2	99	1	0,2
60	30	70	0,2
65	30	70	0,2

70	20	80	0,2
75	20	80	0,2
80	10	90	0,2
85	10	90	0,2
87	99	1	0,2
100	99	1	0,2

Table 2-2: LC method.

MS parameters

MS PARAMETERS QTRAP	
Declustering Potential	100
Entrance Potential	10
Collision Energy	35
Collision Cell Exit Potential	13
Curtain Gas	30
Collision Gas	Medium
IonSpray Voltage	5500
Temperature	500
Scan Time	500
Ion Source Gas 1	40
Ion Source Gas 2	60

Table 2-3: MS parameters for 4000Qtrap and 5500 Qtrap for MRM analysis of the digested A1AT.

MS PARAMETERS QTRAP	
Entrance Potential	10
Entrance Potential	10
Collision Cell Exit Potential	13
Curtain Gas	30
Collision Gas	Medium
IonSpray Voltage	5500
Temperature	500
Ion Source Gas 1	40
Ion Source Gas 2	60

Table 2-4: MS parameter of 5500 Qtrap for MRM analysis of digested melittin.

Q1	Q3	SEQUENCE	DP	CE
431,2	431,2	Arg(24)-Gln(26)	70	10
303,2	303,2	Arg(22)-Lys(23)	70	10
657,43	487,32	Gly(1)-Lys(7)	90	35
657,43	260,19	Gly(1)-Lys(7)	90	35
756,46	927,56	Val(8)-Lys(21)	100	25
756,46	984,58	Val(8)-Lys(21)	100	25

Table 2-5: DP and CE values of 5500 Qtrap for the MRM analysis of digested melittin.

Results and discussion

Initially, we aimed to select of the most suitable protease for A1AT digestion. For this purpose, three criteria were considered: first, the number of generated peptides, second, the number of generated peptides with an elevate m/z (that is relevant for the production of double and triple charge in ESI source) and third, especially, the production of a peptide containing the amino acid substitution of the "F" variant. This peptide consists in a mutation at residue 223 (R223C). Following these three criteria, we analysed the results of two simulated digestions. The first one using trypsin and the second one using Endoproteinase Lys-C from Lysobacter enzymogenes. In particular, trypsin and LysC are very similar considering to the first two criteria, but as highlighted in Table 2-6, only LysC is able to produce the peptide relevant to characterize the "F" variant (peptide 223-233).

Cleavage sites	Sequence
70-125	ADTHDEILEGLNFNLTEIPEAQIHEGFQELLRTLNQPDSQLQLTTGNGLFLSEGLK
26-69	ITPNLAEFAFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTK
344-365	GTEAAGAMFLEAIPMSIPPE VK
175-191	ELDRDTVFALVNYIFFK
137-154	LYHSEAFTVNFGDTEEAK
275-290	FLENEDRRSASLHLPK
202-217	DTEEEDFHVDQVTTVK
311-328	VFSNGADLSGVTEEAPLK
260-274	LQHLENELTHDIITK
11-25	TDTSHHDQDHPTFNK
244-259	YLGNATAIFFLPDEGK
369-380	PFVFLMIEQNTK
223-233	RLGMFNIQHCK
291-300	LSITGTYDLK
194-201	WERPFEVK
235-243	LSSWVLLMK
1-10	EDPQGDAAQK
301-310	SVLGQLGITK
156-163	QINDYVEK
336-343	AVLTIDEK
388-394	VVNPTQK
381-387	SPLFMGK
130-135	FLEDVK
169-174	IVDLVK
218-222	VPMMK

Table 2-6: peptides obtained in the A1AT simulated digestion using LysC.

Once established the most suitable protease, two different amount of A1AT were digested with LysC, in order to investigate the sequence coverage: 3.5 µg and 10 µg. While the 10 µg amount was established as reference digestion, the reason why 3.5 µg were digested is reported in the first chapter. In particular, an approximate quantification of A1AT in healthy newborn saliva was performed with western blot analysis, exploiting a calibration line obtained with different amount of A1AT standard. A1AT is an acute phase protein that increases in concentration during an inflammatory condition. For this reason, we can assume that in infants the detected amount in healthy newborn saliva could represent the minimal detectable quantity in this biological fluid. Moreover, it is relevant considering that the "F" variant is independent on the inflammatory condition so a diagnostic assay to detect this specific allelic variant through the analysis of saliva has to be efficient also in a normal state.

The peptides obtained for each digestions were analysed with UPLC coupled with Synapt G2-S, and results were submitted to BiopharmaLynx software to estimate the value of sequence coverage for both the digestion. This analysis was performed considering one variable post-translational modification (methionine oxidation) and one fixed modification (cysteine alkylation, performed during the treatment with IAA). Substantial differences in sequence coverage were observed. In particular, 60.9% of the sequence was covered for the 10 µg digestion whereas 14.7% of sequence coverage was covered for the lower digested amount (3.5 µg). Nevertheless, in both cases the peptide RLGMFNIQHCVK containing the position 223 was detected. The m/z corresponding to this peptide was measured in high resolution mode as the triple charged 468.55 $[M+3H+]^{3+}$ (Figure 2-5).

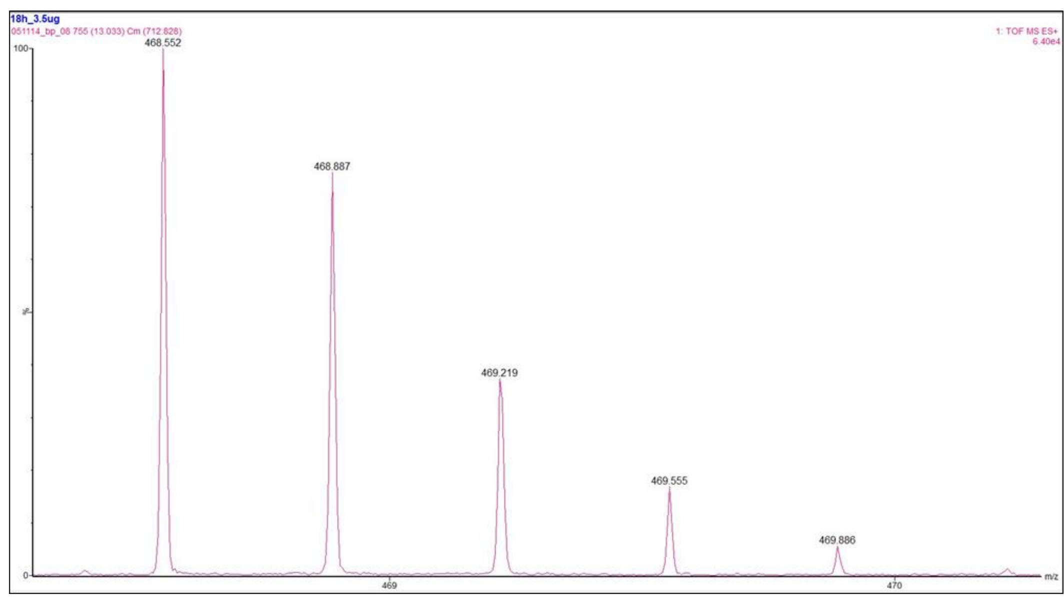


Figure 2-5: pattern distribution of the m/z 468.55 in high resolution.

Although the accurate mass and the MS/MS analysis confirmed the identity of the m/z 468.5, different retention times were observed according to the different protein quantity. In particular, an elution time of 13.5 minutes was observed for the 3.5 µg digestion while an elution time of 16.2 minutes was obtained for the 10 µg digestion (Figure 2-6).

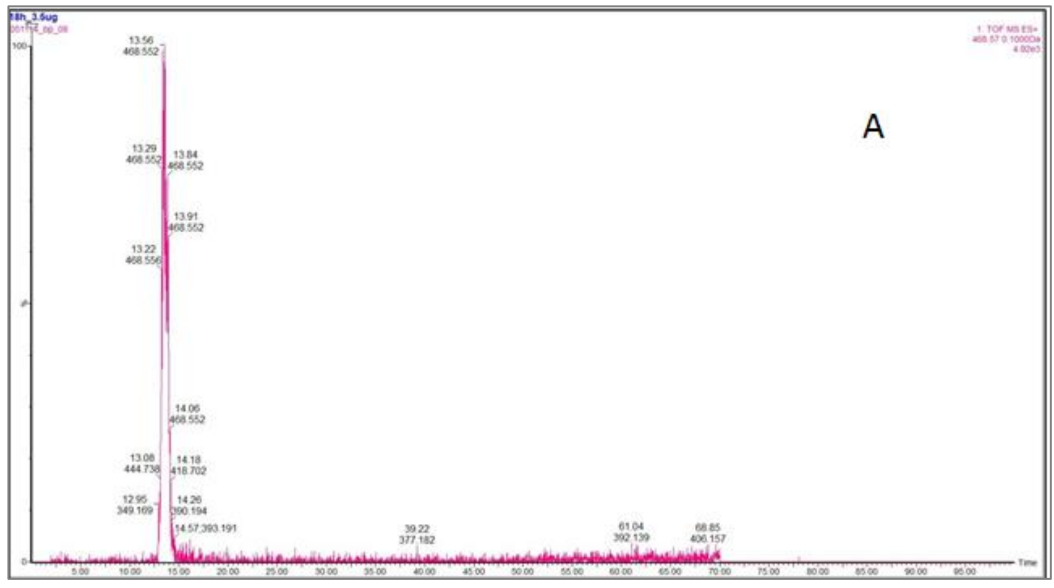


Figure 2-6: chromatogram that shows the different retention time for different amount of digested protein. Chromatogram A is the digestion product of 3.5 µg of A1AT; chromatogram B is the digestion product of 10 µg of A1AT.

In order to investigate this matrix effect (ME), shown as a shift of the elution time, a mixture of the 3.5 µg-digestion and the 10 µg-digestion was performed and analysed with Synapt G2-S. Interestingly, in this case a middle elution time was obtained for the m/z 468.5, specifically 15.2 minutes (Figure 2-7). This phenomenon has been further investigated in the following paragraphs.

Figure 2-7: chromatogram that show the middle elution time observed mixing the 3.5 µg digested solution and the 10 µg digested solution.

Alpha-1-antitrypsin: MS/MS spectrum

MS/MS fragmentation spectrum of the m/z 468.5 was obtained using Synapt G2-S (Figure 2-8). Three intense transitions were selected (468.5/719.3; 468.5/605.3; 468.5/327.2). The experiments in the following paragraph, based on the quantification of the m/z 468.5, was performed using the most intense transition 468.5/327.2. Nevertheless, the co-elution of the three transitions was steadily monitored to improve the confidence of the experiments.

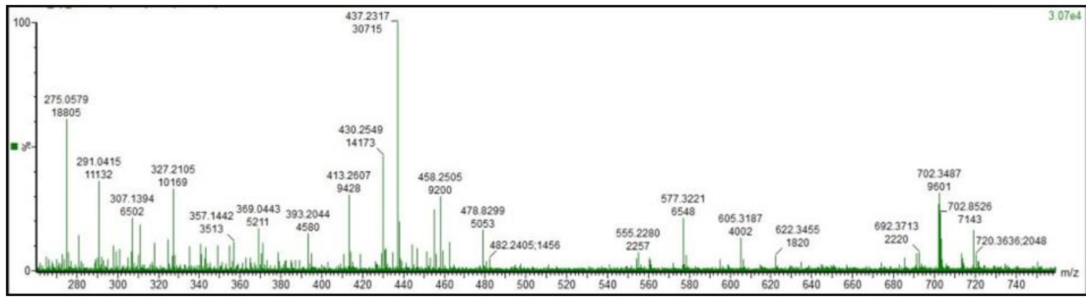


Figure 2-8: MS/MS spectrum of the peptide of interest.

Overcoming the shift of retention time

Matrix effect (ME) can severely affects the quantification methods based on ESI source. Usually ME consists in an unexpected suppression or enhancement of analytes depending on the coeluting matrix. Defining a strategy to overcome this issue is fundamental to preserve the reproducibility, linearity and the accuracy of a LC-MS based method. In fact, ME affects typical mass spectrometric features such as selectivity and specificity [14]. During the development of analytical techniques is relevant to consider ME in particular when reproducible analyses are required [15]. Our purpose is to develop a LC-MS based method to investigate the detection of A1AT in saliva, but also to discriminate the “F” allelic variant. Therefore, the reproducibility, the selectivity and the specificity are fundamental requirements to transpose a LC-MS based method to diagnostic purposes. Several strategies [14,15] to overcome matrix effect are described in literature (for example adjusting the gradient, increasing mobile phase pH, organic extraction step during sample preparation), but in our case being the elution time dependent on the proteins’ concentration, we focused on the amount of digested protein. Macromolecular crowding is an important effect in biochemistry and cell biology that alters the properties of molecules in a solution when high concentrations of macromolecules such as proteins are present. Crowding affects the structure and the state of association of macromolecules in solution and can be defined as the total volume occupancy by macromolecular solutes upon the behaviour of each macromolecular species in that solution [16]. Considering a solution composed of macromolecules, the excluded volume due to the presence of one macromolecule affects the chemical potential and consequently the chemical reactivity of each macromolecules. According to this principle, we have tried to create a crowding effect in the digested samples. In particular, two different amounts of A1AT standard were digested, 3.5 µg and 7 µg. Bovine serum albumin (BSA) was added before the digestion, reducing the volume of “accessible” solvent and consequently increasing the effective concentration of macromolecules. We investigated these samples on HPLC coupled with 4000 Qtrap monitoring both the retention time and the peak area (transition 468.5/327.2 in MRM) in

order to perform a semi-quantitative analysis. The amount of the added BSA was estimated according to the plasma ratio A1AT/serum albumin reported in literature (approximately 1:15), with the intent to recreate a physiological condition and avoiding to alter the equilibrium between these two proteins. The amount of the protease LysC was calculated according to the total quantity of proteins. The results are surprising: in fact, the presence of serum albumin appears to "stabilizes" the elution of the peptide of interest at a specific retention time of 18.6 minutes (Figure 2-9: B and D) both for the 3.5 µg digestion and the 7 µg digestion. Furthermore, also the quality of the chromatogram improves in the presence of serum albumin. In fact, in spectra B and D where serum albumin was added, the peaks appear sharp compared to spectra A and C.

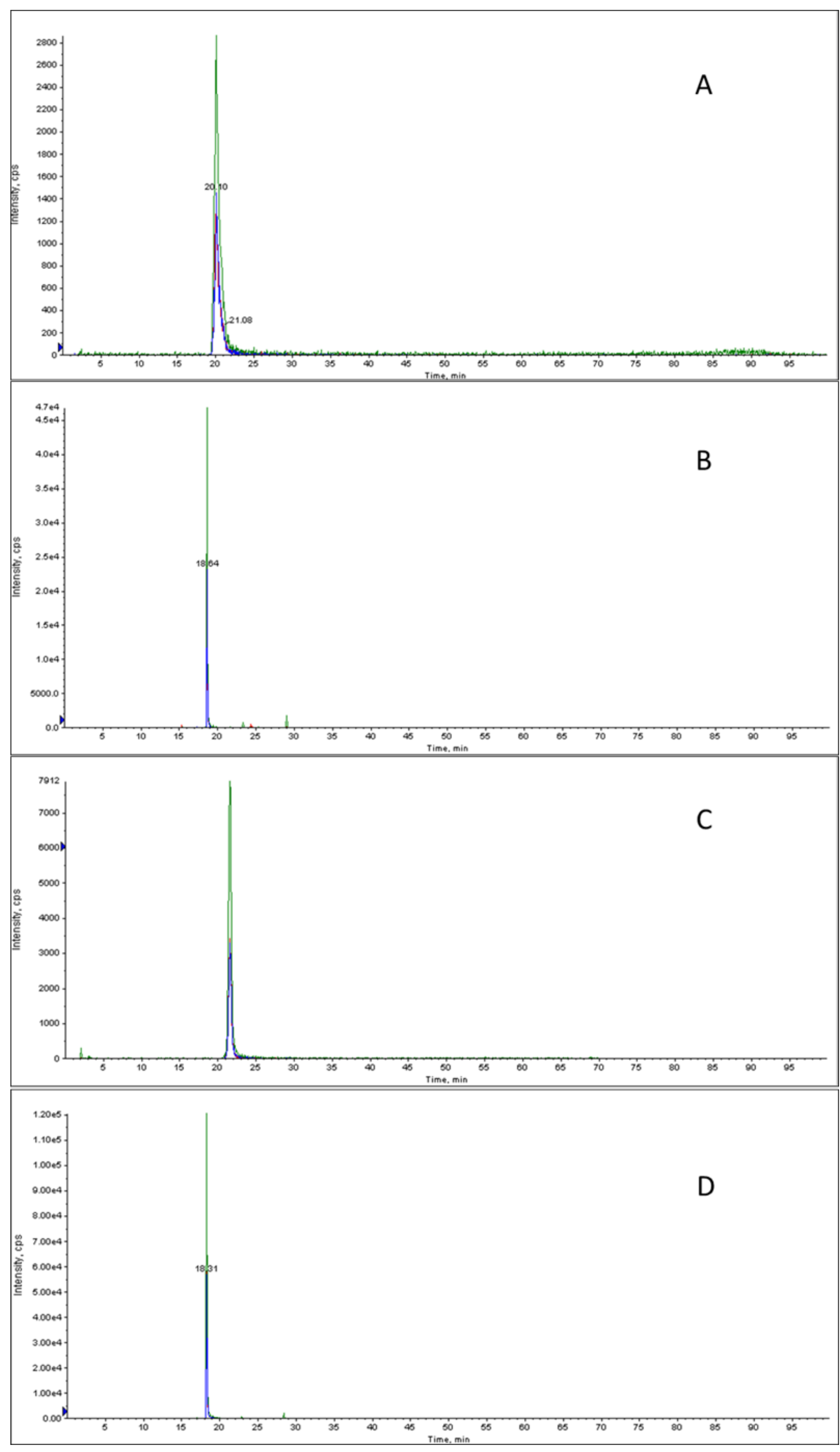


Figure 2-9: A) 3.5 µg digestion; B) 3.5 µg digestion + serum albumin; C) 7 µg digestion; D 7 µg digestion + serum albumin.

The explanation of this result is not yet fully clear but we might suppose that the entropic effect caused by macromolecular crowding affects the activity of the protease LysC, increasing the amount of produced peptide and consequently, in some way, the elution time. Serum albumin has a role of “enhancer” in the digestion as shown in Figure 2-10. In particular, it is surprising how the peak area of the sample “3.5 µg + BSA” is higher than the peak area of the sample “7 µg”. Furthermore, it is relevant considering that serum albumin is one of the most abundant protein in newborn saliva. Therefore, an A1AT digestion performed in presence of BSA is closer to the biological sample condition than the digestion of the only A1AT commercial standard.

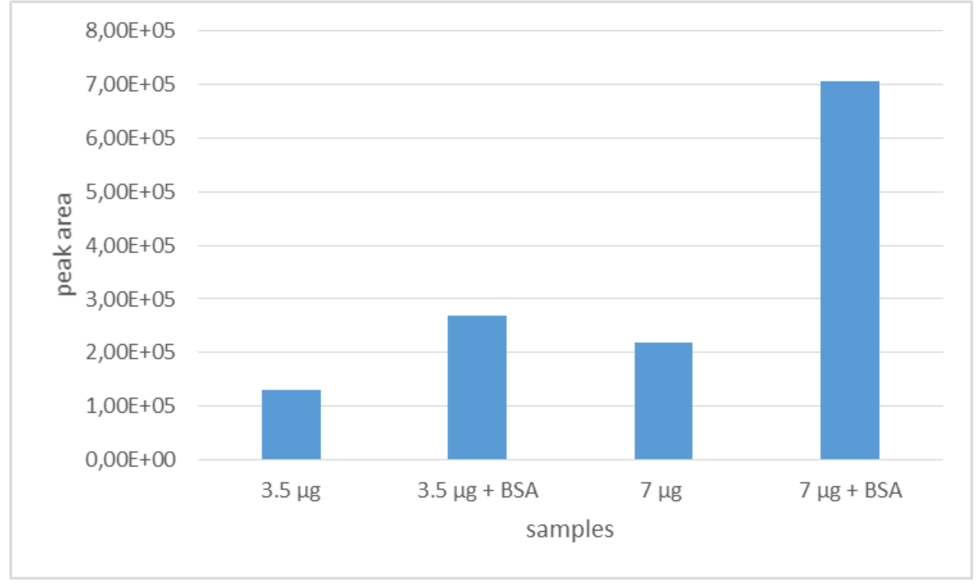


Figure 2-10: peak area of the transition 468.5/327.2 in presence/absence of serum albumin.

LysC protease: a deeper investigation of the activity

In the present work, the protease LysC was employed to obtain the peptide of interest. This protease cleaves proteins on the C-terminal side of lysine amino acids. Lys-C activity is optimal in pH range of 7.0-9.0. Considering the matrix effect previously discussed, we asked whether LysC activity might be affected by serum albumin, and affecting the shift in retention time. For this purpose, both spectrophotometric (λ=405 nm) and MS based assays were exploited. N-p-Tosyl-Gly-Pro-Lys 4-nitroanilide is a chromogenic substrate for coagulation protease, but the presence of a lysine in the

chemical structure makes it also a substrate for LysC. In particular, spectrophotometric assays were performed using the N-p-Tosyl-Gly-Pro-Lys 4-nitroanilide in place of A1AT and the assays were performed for two different amount of LysC in presence/absence of bovine serum albumin.

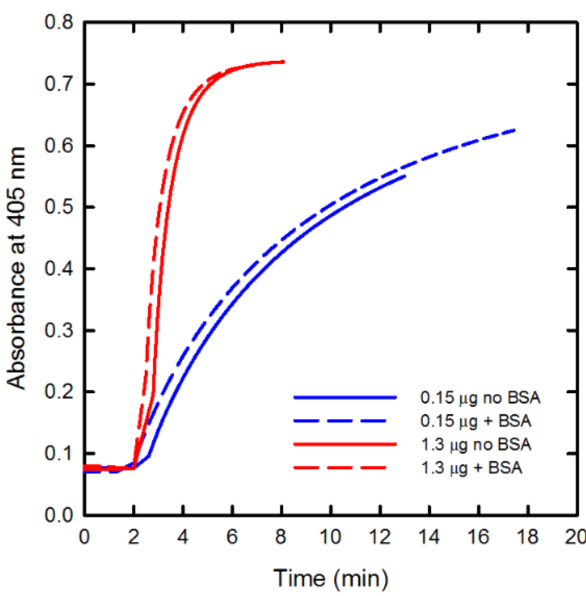


Figure 2-11: LysC activity in presence and absence of serum albumin.

Since for a given amount of Lys C no significant effects were detected in presence or absence of BSA (Figure 2-11) with the commercial substrate, the experiment was not pushed forward to the determination of catalytic parameters. A different approach was instead pursued, and an investigation of this phenomenon performing a LC-MS based activity assay (discussed in the next paragraph), which also requires a lower consumption of LysC than spectrophotometric assay was carried out. In particular, we consider that the shift of retention time might be specifically associated with the substrate A1AT.

Melittin: a different substrate for LysC

Melittin is a small protein containing 26 amino acid residues and is the principal toxic component of bee venom [17]. The presence of three lysine makes melittin a suitable substrate to investigate a digestion preformed with LysC protease. In the

presence and absence of BSA, we investigated the digestion of melittin recreating the A1AT digestion conditions. In particular, the digestion of melittin with LysC leads to the production of four peptides: Arg(24)-Gln(26) m/z 431.2, Arg(22)-Lys(23) m/z 303.2, Gly(1)-Lys(7) m/z 657.4 and Val(8)-Lys(21) m/z 756.4. A solution of digested melittin was analysed in full scan with Synapt G2-S and MS/MS spectra were acquired for each peptide. The same LC method employed for the A1AT was used for these experiments. The transitions reported in Table 2-7 were selected for MRM analysis.

TRANSITIONS
431.2 / 431.2
303.2 / 303.2
657.4 / 487.3
657.4 / 260.1
756.4 / 927.5
756.4 / 984.5

Table 2-7: selected transitions for MRM analysis.

MRM analysis (Figure 2-12) was performed using 5500 Qtrap. In particular, we analysed a digestion of 3.5 µg of melittin (with and without albumin) and a digestion of 10 µg of melittin (with and without albumin). We analysed the resulting spectra according to the retention times and peak areas of the transitions.

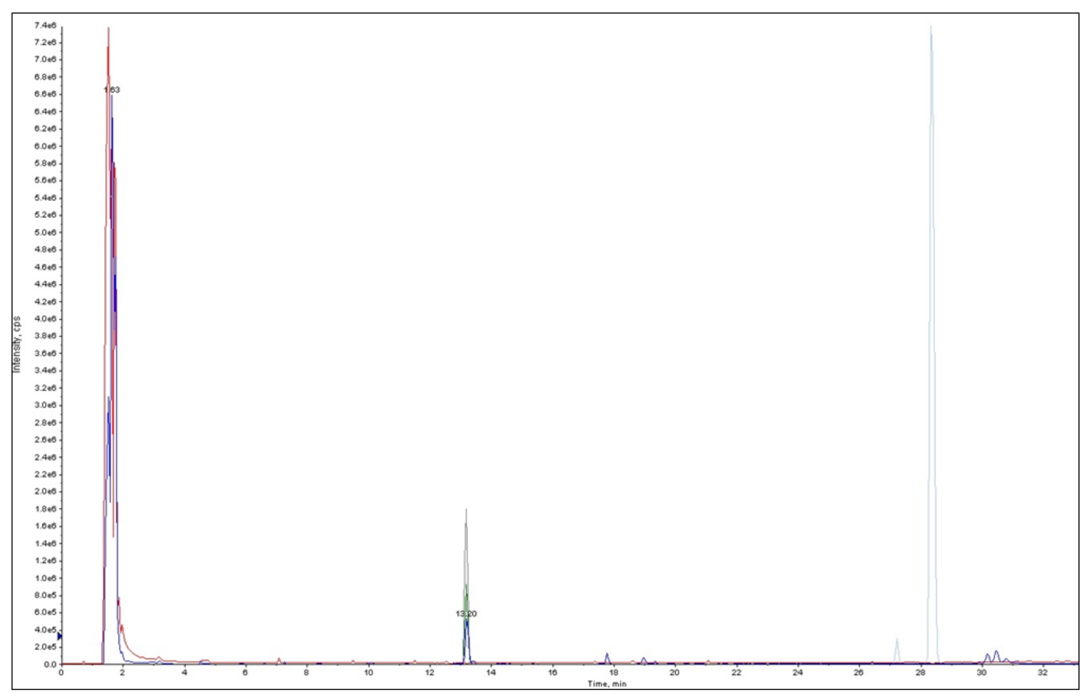


Figure 2-12: MRM spectra of melittin

transition	melittin 3.5 μg + serum albumin	melittin 3.5 μg	melittin 10 μg + serum albumin	melittin 10 μg
431.2 / 431.2	1,6	1,6	1,6	1,6
303.2 / 303.2	1,5	1,6	1,5	1,8
657.4 / 487.3	13,0	13,0	13,0	13,0
657.4 / 260.1	13,0	13,0	13,0	13,0
756.4 / 927.5	28,2	28,2	28,1	28,2
756.4 / 984.5	28,2	28,2	28,1	28,2

Table 2-8: results of retention times (in minutes) for each melittin transition.

Results are reported in Table 2-8. In particular, changing the amount of melittin or adding bovine serum albumin, no significant differences in retention time are observed for each transition of melittin. Obviously, significant differences occur between melittin and A1AT. Melittin is a peptide of 26 amino acids, A1AT indeed presents a protein scaffold that affects the activity of proteases. For this reason, these results might be a

proof that the observed matrix effect for the A1AT is not related to the effect of albumin on the protease but is probably related to a specific effect on the protein structure. These experiments are also relevant to exclude a subnormal activity of LysC that could affects the results of the liquid chromatography.

Figure 2-13: measured area for each transition of melittin. Data obtained with 5500 Qtrap.

Nevertheless, observing Figure 2-13 two transitions (431.2/431.2 and 756.4/927.5) are considerably more intense in the samples containing serum albumin that, also in this case, appears to increase the yield of digestion. Observing the data of LysC activity and the data of melittin digestion, an ultimate experiment was relevant to investigate the matrix effect.

The role of serum albumin as "crowding agent"

Considering the hypothesis that alteration of A1AT conformation is the main effect involved in the observed matrix effect, it remains to investigate the role of serum albumin. In particular, with these experiments we aimed to evaluate the involvement of serum albumin in A1AT digestion using LysC. In other words, which is the effect of a different crowding agent on the peptide retention time? We select a crowding agent that, differently from serum albumin, is not a competitive substrate for LysC. For this purpose two different amount of A1AT were digested. The same amount was also digested in presence of PEG400. 5500 Qtrap was used in MRM mode monitoring the retention time, but also performing a semi-quantitative analysis of the amount of the produced peptide (transition 468.5/327.2).

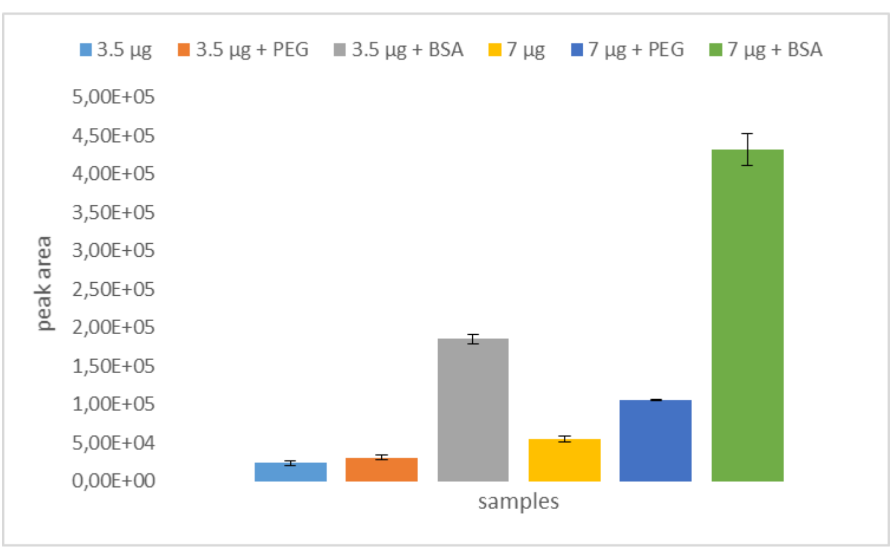


Figure 2-14: semi-quantification performed with the transition 468.5/327.2; digestion in presence of BSA or PEG400.

Observing the Figure 2-14, the higher values of peak area are shown in the digestion in presence of serum albumin (grey and green). Considering the digestions with PEG400, an increasing in peak area is observed (orange vs light blue and dark blu vs yellow) but is considerable lower than the digestion with serum albumin. The samples were also analysed according to the retention time. Serum albumin confirms to stabilize the elution, at different concentration of A1AT, of the peptide of interest. On the other hand, PEG400 has not the same effect (data not shown). These results confirm that the crowding agent probably has to be a substrate for LysC, and the interaction albumin-LysC-A1AT is relevant for increasing the rate of digestion and consequently the elution. Nevertheless, we cannot exclude that this phenomenon is the result of a specific action of the albumin. This protein is known to bind several substrate, and this might affects the retention time of the peptide. To investigate this property, another experiment was performed substituting the albumin with porcine haemoglobin in the digestion of A1AT (also in this case, two different amount of A1AT were digested, 3.5 µg and 7 µg). Haemoglobin was added at the same concentration of bovine serum albumin in the previous experiments. This analysis was performed both with 4000 Qtrap and with 5500 Qtrap, to monitor not only the retention time, but also the amount of A1AT produced peptide (transition 468.5/327.2 in MRM). Surprisingly, considering the retention time, the

effect of haemoglobin is very similar to the effect of serum albumin (Figure 2-15). In fact, observing the spectra B and D, the elution time for the peptide of interest is the same for the two different amounts of digested A1AT, but is different from the retention time obtained in the digestion with BSA. Furthermore, as already observed with serum albumin, significant improvement in peak sharpening were obtained in the presence of haemoglobin.

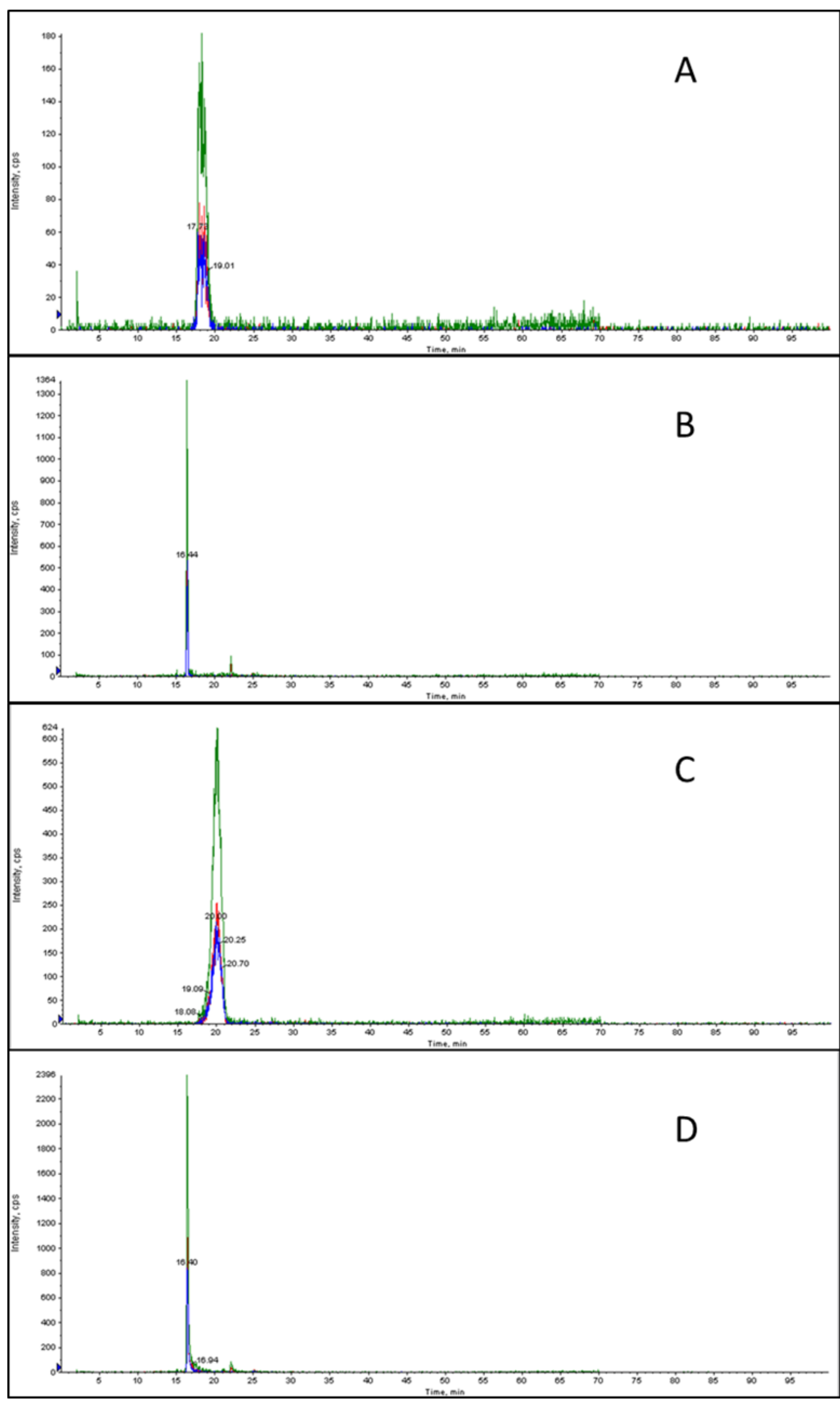


Figure 2-15: A) 3.5 µg digestion; B) 3.5 µg digestion + haemoglobin; C) 7 µg digestion; D) 7 µg digestion + haemoglobin.

Considering the semi-quantitative analysis (Figure 2-16), a trend of "overproduction" has been observed also in the samples with haemoglobin. In particular, the presence of haemoglobin that is a crowding factor but also a substrate for LysC, appears to enhance the yield of digestion. A comparison with serum albumin will be discuss in the next paragraph. The semi-quantitative analysis was performed with 5500 Qtrap to allow the comparison with the data of the digestions in presence of PEG400.

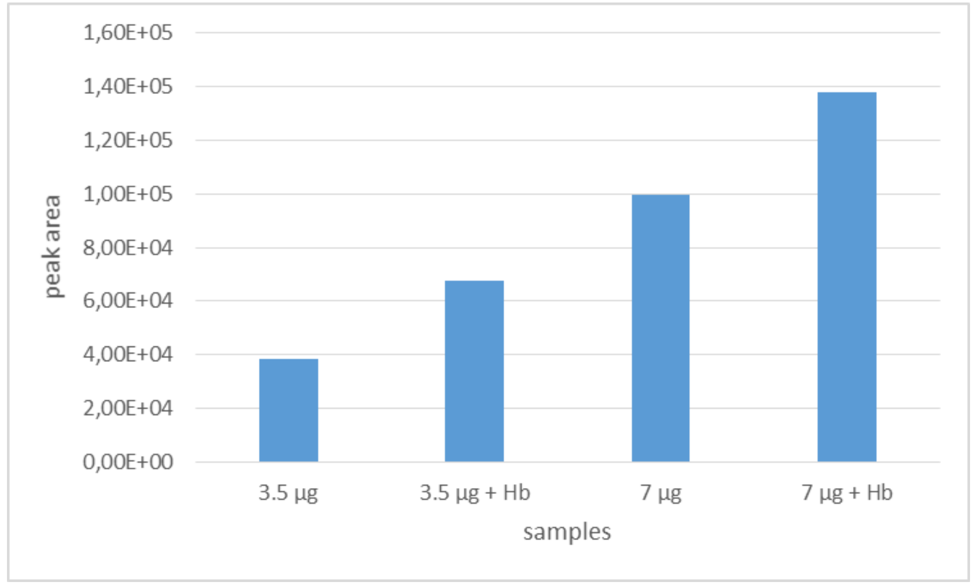


Figure 2-16: semi-quantification performed with the transition 468.5/327.2; digestion in presence of haemoglobin.

A1AT elution: crowding and chromatography effect

Observing all the experiments, we can try to summarize the behaviour of the A1AT after digestion and LC-MS analysis. In particular, at first we aimed to investigate the elution time of the peptide of interest. In fact, the m/z 468.5 is relevant both to quantify A1AT in inflammatory condition and to evaluate the presence of the F phenotype in a patient. Unfortunately, the elution time of this m/z (but this trend has been observed for each A1AT peptides, data not shown) is dramatically affected by the amount of digested protein. This feature represents a considerable issue in the development of a robust and reproducible MS-based method to analyse A1AT. The presence of bovine serum albumin in the digested mixture appears to stabilize the A1AT elution time and to increase the

amount of produced peptide. For this reason, spectrophotometric assays were performed to evaluate the activity of the different amount of LysC protease in the presence and absence of albumin. The different amount of LysC obviously shown different activity, but the presence of serum albumin does not significantly affect the activity of the protease. In a second time, we focused on A1AT structure. In particular, we substituted A1AT with melittin that represents a substrate for LysC protease. Melittin digestions were performed in the presence/absence of serum albumin. The elution time of melittin peptides is not affected by the amount of digested protein, confirming the “specificity” of A1AT in this particular behaviour. Furthermore, we tried to evaluate the specificity also of serum albumin as “crowding agent”. For this purpose A1AT digestion was performed in the presence of PEG400 (not-LysC substrate) and porcine haemoglobin (LysC substrate), added in the same amount of serum albumin. While PEG400 does not reproduce the same effect of serum albumin, haemoglobin stabilizes the peptide elution time exactly as serum albumin did it. However, as reported in Figure 2-17, a marked difference in peak area is observed comparing the digestion with albumin and haemoglobin. Therefore, in conclusion, the presence of other proteins dramatically affect not only the elution time of the m/z 468.5, but also the yield of digestion. The consequence of the crowding effect, due to the presence of other proteins, is a decrease in the void volume. The result is probably an increase in peptides interaction not only with the LysC, but also with the particles of the stationary phase.

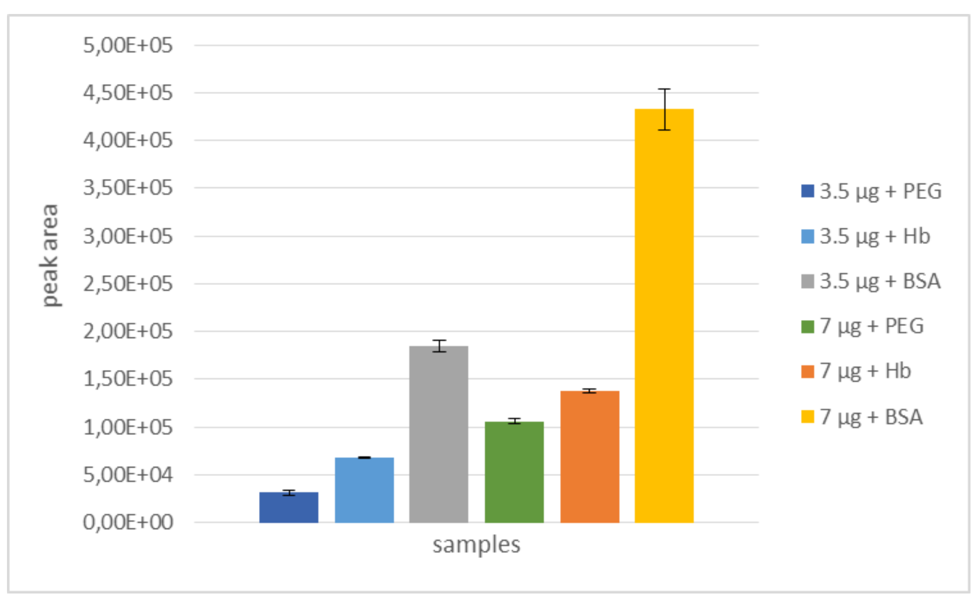


Figure 2-17: peak area comparison between A1AT digestion in presence of serum albumin, PEG400 and haemoglobin. Analysis performed with 5500 Qtrap.

Alpha-1-antitrypsin: proof of concept of quantitative analysis

A1AT quantification must be performed using the most sensitive and selective MS assay. Multiple reaction monitoring (MRM) meets these requirements. As shown in the paragraph "Alpha-1-antitrypsin: MS/MS spectrum" three transitions (468.5/719.3; 468.5/605.3; 468.5/327.2) were selected to monitor the peptide of interest in MRM mode using 4000 Qtrap and the calibration line was obtained following the most intense transition 468.5/327.2. In particular, 7 µg of A1AT standard were digested in presence of BSA and the following concentrations were obtained diluting the 0.07 µg/µl solution: 0.035 µg/µl, 0.01 µg/µl, 0.005 µg/µl, 0.001 µg/µl and 0.0005 µg/µl. The semi-quantitative analysis of the transition 468.5/327.2 was exploited to obtain the calibration line (Figure 2-18).

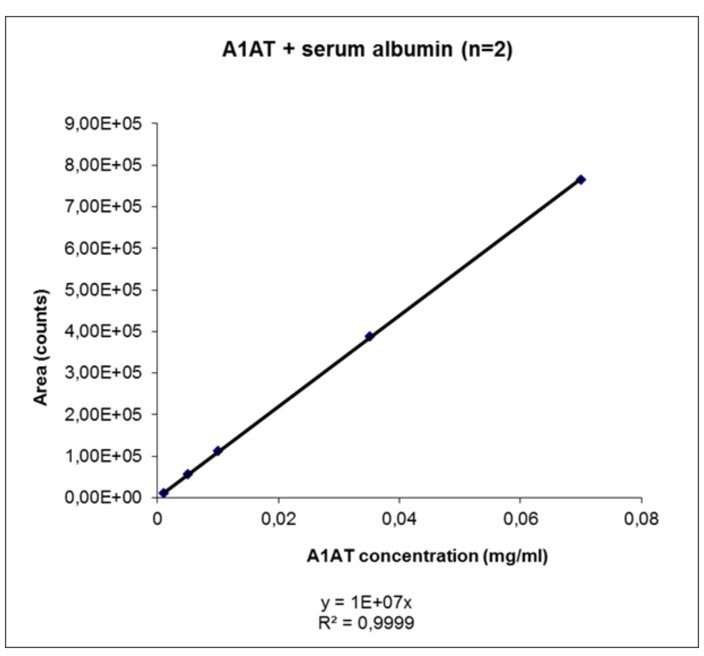


Figure 2-18: calibration curve obtained in presence of serum albumin.

The same analysis was also performed in the absence of bovine serum albumin obtaining a second calibration curve (Figure 2-19).

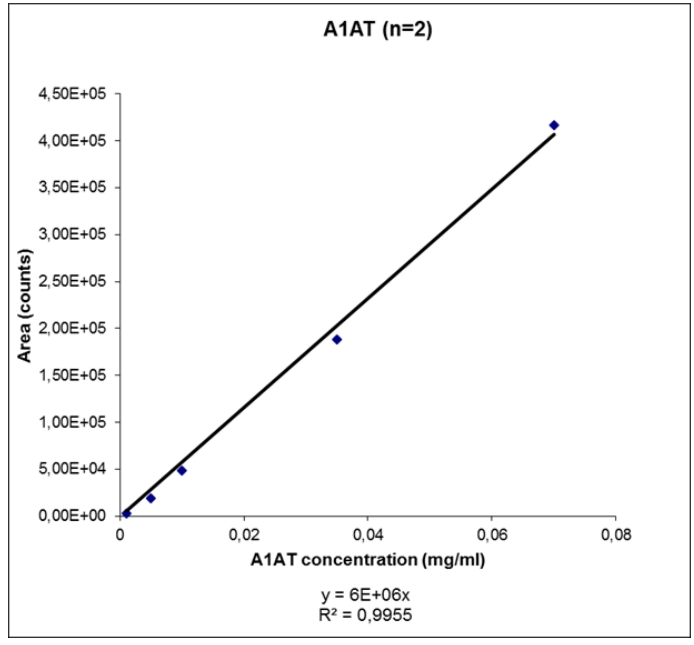


Figure 2-19: calibration curve obtained without serum albumin.

Surprisingly, increasing the sample complexity with serum albumin, an improvement in the accuracy of the analysis was observed. In fact, an R^2 =0.9999 was observed with in the sample where BSA is present, while in the absence of albumin A1AT digestion leads to a lower value of 0.9955. Furthermore, in presence of serum albumin the retention time of the peptide is steady, unlike in the absence of serum albumin were different retention times were observed according to sample dilution. Exploiting the calibration line, LOQ (limit of quantification, the concentration for which signal to noise ratio is 10) was estimated at 0.5 μg/ml in the presence of serum albumin, whereas without albumin the value increase at 1 μg/ml. Once again, serum albumin appears to increase the amount of produced peptide and in a semi-quantitative analysis this feature dramatically impacts on the limit of quantification. Furthermore, it is relevant considering that the condition in the presence of bovine serum albumin is more representative of a hypothetical biological sample condition. In fact, serum albumin is one of the most abundant protein in newborn saliva. Therefore, considering the concentration range, the A1AT calibration line in the presence of albumin can be exploited to quantify this acute phase protein within saliva samples both in infected and healthy newborns saliva.

Alpha-1-antitrypsin: investigation of the “F” allelic variant

The so-called “F” variant of A1AT, consisting in a mutation at residue 223 (R223C) is considered a potentially emphysema risk factor because it exhibits a lower affinity for the neutrophil elastase. In particular, this allelic variant has not generally considered clinically relevant because the circulating levels are very similar to those of the normal variant. Nevertheless, the “F” variant may predispose to the development of lung and liver diseases in PiFZ heterozygote subjects. In the present work, we investigated the “F” variant in order to develop a reproducible and sensitive assay to discriminate this important allelic variant from the normal one (phenotype PiMM). The high resolution MS, the ion mobility and MRM analysis were exploited to discriminate the peptide RLGMFNIQHCVK (normal variant) from the “peptide F” CLGMFNIQHCVK and consequently to build a LC-MS based method that allows the detection of this specific variant within biological samples.

In particular, a peptide including R223C substitution (VPMMKCLGMFNIQHCKKLSSWV) was synthetized by Selleck Chemicals. The peptide includes the two cleavage sites of LysC also present in the whole protein. This feature is crucial to investigate the “F” variant both in the phase of digestion and in the MS analysis. In particular, A1AT 0.07 µg/µl was digested in the presence of “peptide F” 0.07 µg/µl and serum albumin. The peptides were analysed with UPLC coupled with Synapt G2-S. Results showed the m/z 468.5 that originates from the A1AT digestion and the m/z 469.902 $[M+3H+]^{3+}$ deriving from the digestion of the synthetic peptide, corresponding to the peptide F (Figure 2-20).

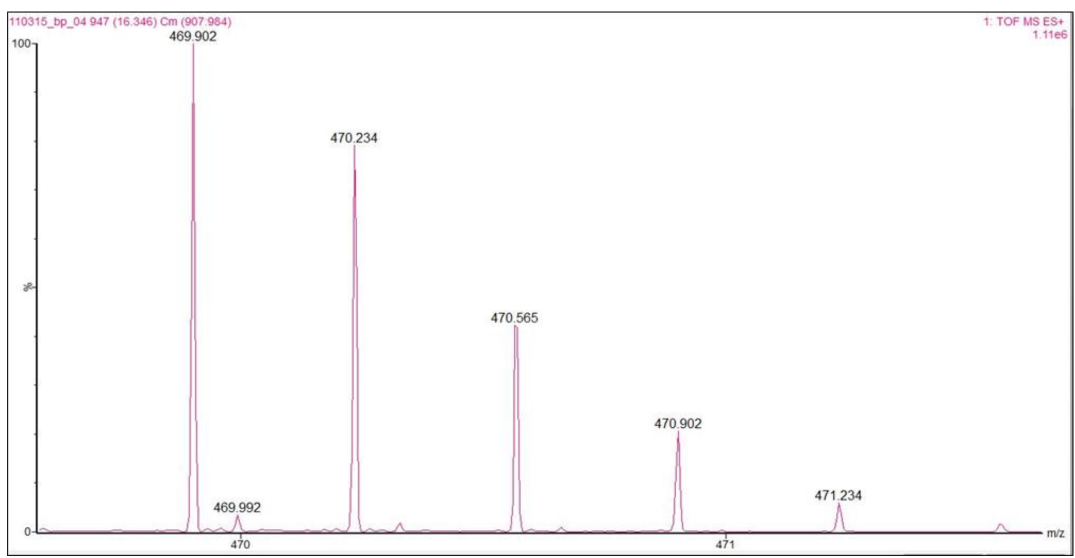


Figure 2-20: isotopic distribution of m/z 469.9.

In fact, the two lysine in the synthetic peptide VPMMKCLGMFNIQHCKKLSSWV represent the cleavage sites of the protease LysC, producing the “F” peptide CLGMFNIQHCVK. The first parameter that we considered for the characterization of the peptide F is the elution time. In particular, we evaluated how R223C substitution can affect the physicals and chemicals features of the peptides, the interaction with the stationary phase of the column and consequently the retention time. Observing the chromatogram, no significant differences in elution times were appreciated: in particular,

an elution time of 16.2 minutes was observed for the m/z 468.5 compared to an elution time of 16.4 minutes for the peptide F (Figure 2-21).

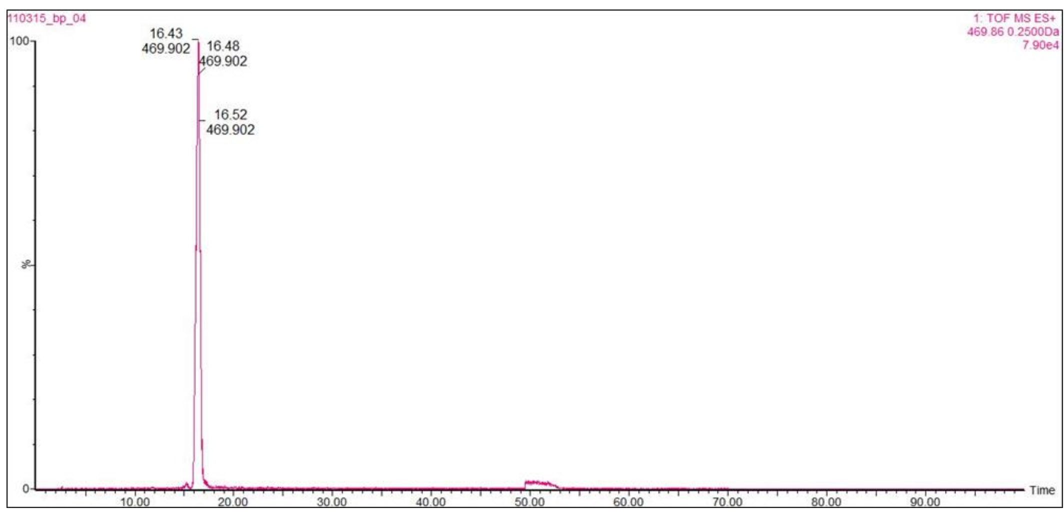


Figure 2-21: chromatogram of the m/z 469.9.

Observing that these two peptides are characterized by the same retention time we considered the possibility to discriminate the two variants according to the shape of the peptides. For this purpose, Synapt G2-S was exploited for the ion mobility. In particular, the operating principle of an ion mobility mass spectrometer (IMS) is based on the mobility of a gas-phase ion under both the positive influence of an electrical field gradient and a negative influence of a cross-flow of buffer gas. In these conditions, the mobility of an ion essentially depends on its mass, charge and collision cross section (shape): ion charges being equal, small and compact ions with small collision cross sections drift more quickly than large and extended ions with large collision cross sections.

A mixture of A1AT, serum albumin and “F” peptide was digested and analyzed. The parameter observed in an IM analysis is the “drift time”; two molecules with identical shape are characterized by equal value of drift time. This parameter is extrapolated from the mobilogram (that is reported in Figure 2-22). Observing the Figure 2-22, two peaks are shown corresponding to the drift time of the analytes of interest. Both the m/z 469.9 (“F” peptide) and the m/z 468.5 showed a drift time of 5.21 minutes. These results

confirm that the substitution R223C does not significantly affects the shape of the peptide of interest and this is substantially in agreement with the data of retention time obtained in liquid chromatography.

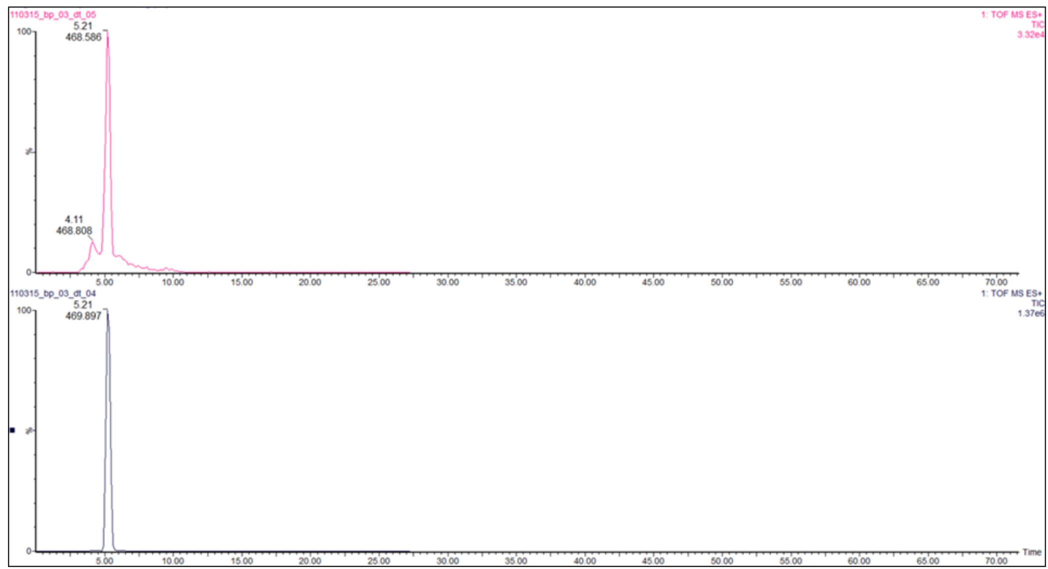
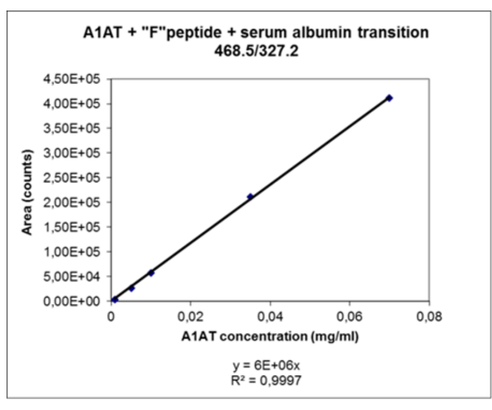


Figure 2-22: on the top the mobilogram of the m/z 468.5; on the bottom the mobilogram of the m/z 469.9.

Since retention time and drift time cannot be used to discriminate between the two variants of A1AT, we exploited the fragmentation of the peptides. In particular, after digestion the MS/MS analysis of both "the normal" and the "F" peptide was performed using Synapt G2-S obtaining different fragmentation spectra. The most intense transition 468.6/327.2 was selected to monitor the peptide of interest of A1AT in MRM analysis, and the transition 469.8/246.1 was selected to monitor the A1AT F variant.

These analyses were performed with a mixture of digestion composed of A1AT, synthetic "F" peptide and bovine serum albumin at different concentrations in order to build a calibration line for both the normal variant and the "F" variant. Once again, we exploited the high sensitivity of 4000 Qtrap. The results are shown in Figure 2-23.



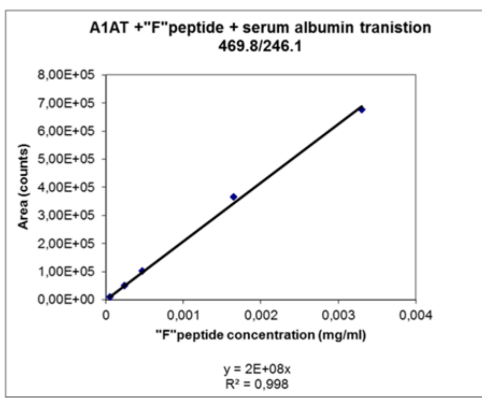


Figure 2-23: calibration lines of the peptides of interest.

The first crucial observation is that, through MS/MS analysis, the "F" variant of A1AT is univocally discriminated from the normal variant. In fact, the fragmentation spectra are different and specific for each peptides. The calibration lines reported in Figure 2-23 allow the quantification of both the normal A1AT (relevant in inflammatory condition) and the "F" variant. The calibration line of the F variant is fundamental to investigate how the "F" peptide responds in terms of ionization compared to the normal one. Transposing the method to biological samples, the use of a sensitive mass spectrometer such as 4000 Qtrap in MRM mode is essential for the sensitivity and the reproducibility of the analysis. Nevertheless, it is relevant to consider that the use of internal standard is required for the quantification of the A1AT within biological samples. Moreover, with these experiments, we also estimated the minimal quantifiable concentration of the "F" peptide; in particular, the LOQ was determined at 50 ng/ml for the "F" peptide. The possibility to reach these low concentration values is crucial because the "F" allelic variant is independent on the inflammatory condition. Therefore, this sensitivity allows detecting this allelic variant also in a healthy newborn saliva, where the A1AT is not necessary differently expressed. In the biological sample investigation, it is crucial to consider the extreme variability of the elution time according to the amount of digested protein. The presence of bovine serum albumin allows building the calibration line that is more similar to a biological sample than the presence of the only A1AT standard. Nevertheless, the detection and the quantification of this protein must be validated also by other techniques. For example, the coupling of the immune-blot and LC-

MS approaches might be essential to improve the confidence in the identification of this protein within biological samples avoiding false positives.

Conclusions

The development of diagnostic tools allowing the early detection of an ongoing infective state represents an urgency in neonatal clinical care units. The acute phase protein A1AT was detected to be differently expressed in infected newborn saliva. In the present work, we characterized this protein with the aim to build a reproducible and sensitive LC-MS based method for the detection and the quantification of A1AT in biological fluid and, in particular, in newborns saliva. The selected quantifier peptide RLGMFNIQHCVK (468.5) allows both quantification of A1AT and the detection of the “F” variant. In particular, the “F” phenotype CLGMFNIQHCVK (m/z 469.9) is discriminated from the normal variant according to the MS/MS fragmentation spectrum. In fact, both the drift time and the retention time do not permit a distinct separation of the two variants. Furthermore, the elution time of the m/z 468.5 is dramatically affected by the amount of digested protein. For this reason, the A1AT digestion was performed in presence of three different crowding agents and serum albumin appear both to stabilize the elution time and to optimize the yield of digestion. Taken into account that the A1AT is not a specific marker of inflammation, the coupling of the detection of this protein with the metabolomics profile of newborn saliva sample (in particular, phospholipids and eicosanoids) might represent a “start point” to define a more complete framework of an inflammatory condition. Therefore, the present work has set the bases for the development of novel approaches to detect biomarkers of inflammation in newborn saliva.

Bibliography

1) Janciauskiene, S., Wrenger, S., & Welte, T. (2013). Immunoregulatory Properties of Acute Phase Proteins—Specific Focus on α1-Antitrypsin.
2) Gruys, E., Toussaint, M. J. M., Niewold, T. A., & Koopmans, S. J. (2005). Review: Acute phase reaction and acute phase proteins. Journal of Zhejiang University. Science. B, 6(11), 1045-1056
3) Khan, F. A., & Khan, M. F. (2010). Inflammation and acute phase response. International Journal of Applied Biology and Pharmaceutical Technology, 1(2), 312-321.
4) Huntington, J. A. (2011). Serpin structure, function and dysfunction. Journal of Thrombosis and Haemostasis, 9(s1), 26-34.
5) Kalsheker, N. (1989). Alpha1-antitrypsin: structure, function and molecular biology of the gene. Bioscience reports, 9, 129-138.
6) Kolarich, D., Weber, A., Turecek, P. L., Schwarz, H. P., & Altmann, F. (2006). Comprehensive glyco-proteomic analysis of human α1-antitrypsin and its charge isoforms. Proteomics, 6(11), 3369-3380.
7) Brown, W. T. (1982). Alpha 1-antitrypsin: apparent molecular weight heterogeneity shown by two-dimensional electrophoresis. American journal of human genetics, 34(2), 195-208
8) Bergin, D. A., Hurley, K., McElvaney, N. G., & Reeves, E. P. (2012). Alpha-1 antitrypsin: a potent anti-inflammatory and potential novel therapeutic agent. Archivum immunologiae et therapiae experimentalis, 60(2), 81-97.
9) Sinden, N. J., Koura, F., & Stockley, R. A. (2014). The significance of the F variant of alpha-1-antitrypsin and unique case report of a PiFF homozygote. BMC pulmonary medicine, 14(1), 132.
10) Okayama, H., Brantly, M., Holmes, M., & Crystal, R. G. (1991). Characterization of the molecular basis of the alpha 1-antitrypsin F allele. American journal of human genetics, 48(6), 1154-1158.

11) Smith DP, Giles K, Bateman RH, Radford SE, Ashcroft AE. Monitoring Copopulated Conformational States During Protein Folding Events Using Electrospray Ionization-Ion Mobility Spectrometry-Mass Spectrometry. Journal of the American Society for Mass Spectrometry. 2007;18(12-9):2180-2190.
12) Doerr, Allison. "Mass spectrometry-based targeted proteomics." *Nature methods* 10.1 (2013): 23-23.
13) Tuazon, Sherilyn Alvaran; Scarpaci, Anthony P (2012). Staros, Eric B, ed. "Serum protein electrophoresis". Medscape. Retrieved 2 October 2013.
14) Cappiello, A., Famiglini, G., Palma, P., Pierini, E., Termopoli, V., & Trufelli, H. (2008). Overcoming matrix effects in liquid chromatography− Mass Spectrometry. Analytical chemistry, 80(23), 9343-9348.
15) Smeraglia, J., Baldrey, S. F., & Watson, D. (2002). Matrix effects and selectivity issues in LC-MS-MS. Chromatographia, 55(1), S95-S99.
16) Minton, A. P. (2005). Influence of macromolecular crowding upon the stability and state of association of proteins: predictions and observations. Journal of pharmaceutical sciences, 94(8), 1668-1675.
17) Terwilliger, T. C., Weissman, L., & Eisenberg, D. (1982). The structure of melittin in the form I crystals and its implication for melittin's lytic and surface activities. Biophysical journal, 37(1), 353.

3- Lipids profile investigation of healthy and infected newborn saliva: phospholipids and eicosanoids as marker of inflammation

Abstract

In the present work, we analysed newborn saliva samples in order to investigate whether phospholipids and mediators of inflammation (eicosanoids) are subject to variation in inflammatory condition. In particular we provide an overview of the general composition of these molecules in saliva samples and some ideas to optimize the sample preparation, which is a crucial step to analyse these molecules. The study has been conducted with samples collected during the first 48 hours of life of the newborns. We have subdivided newborn saliva samples in three groups: the first one of healthy newborns, the second of potentially infected newborns and the third of infected newborns (high blood level of C-reactive protein and positive swab for pathogenic bacterium). A targeted lipidomic approach using UPLC coupled with 5500 Qtrap was employed to monitor the major phospholipids classes, while mediators of inflammation (also called PUFA-derived lipid mediators) were also investigated with 5500 Qtrap but in multiple reaction monitoring (MRM). Three classes of phospholipids were detected for each group of analysis and a weak trend has been observed in lysophosphatidylcholines composition according to the inflammatory condition. On the other hand, phospholipids were detected in high concentration but the profiles appear to be very similar for each group. Mediators of inflammation are probably present in trace amount, and below the limit of detection of the analytical platform adopted, especially for what concern the sample preparation procedure. Finally we discussed the results according to the salivary proteome previously identified and we framed some hypothesis for future investigation of newborn saliva samples.

Introduction

In the previous chapters, we described the healthy newborn salivary proteome and, exploiting the image analysis, we detected a different expression of the alpha-1-antitrypsin in infected newborn saliva. The A1AT, belonging to the acute phase proteins class, makes the present work very interesting for the development of a non-invasive saliva-based diagnostic assay in infants. Nevertheless, we cannot forget that this protein does not a represent a specific marker of inflammation. In fact, a quantitative analysis of this protein in biological fluids might confirm an “in place” inflammatory condition. However, A1AT is not uniquely associated with inflammation [1]. In this chapter we considered the possibility to explore the metabolomics pattern in newborn saliva in order to provide a full yet preliminary characterization of this biological fluid, both in healthy and infected conditions. Salivary metabolomics pattern, and in particular the lipid profile, might represent a support of the A1AT quantification to confirms an inflammatory condition in newborns.

Lipids of salivary secretions can be subdivided in three classes: neutral lipids, glycolipids and phospholipids. The neutral lipids are represented by free fatty acids, cholesterol, cholesteryl esters, and mono-, di- and triacylglycerols. In the present work we focused on phospholipids detection that are subdivided in six classes:

I. Phosphatidylcholines (PC);
II. Phosphatidylethanolamines (PE);
III. Sphyngomyelins (SM);
IV. Phosphatidylglycerols (PG);
V. Phosphatidylinositols (PI);
VI. Phosphatidylserines (PS);

The salivary phospholipids exhibit a high content of phosphatidylcholine, phosphatidylethanolamine and sphingomyelin. Generally, the lipid content of serum is 50-60 times greater than that of saliva, and both salivary neutral lipids and salivary phospholipids usually originate from serum transudates [2].

The fatty acid constituents of salivary phospholipids range in size from C_{12} to C_{26} and consist of saturated and unsaturated species. Phospholipids are an energy store, a signal for cells communication [3,4] and a structural components of cell membranes. This biological component also plays a key role in inflammatory processes. In particular, when cells are stimulated, the arachidonic acid (AA, composed of 20 carbon atoms with 4 unsaturation) is released from membrane phospholipids by phospholipase A2 (Figure 3-1) that hydrolyzes the acyl ester bond. The AA is then processed by either cyclooxygenases (COX), lipoxygenases (LOX) or cytochrome P450 monooxygenases (CYP). These enzymes lead to the production of the so called “PUFA-derived lipid mediators” (for example prostaglandins, leukotrines, lipoxins) that are responsible of the inflammatory process [5]. For this reason, during the inflammation a progressive decrease in phospholipids concentration is usually observed due to the activity of the protein phospholipase A2. Simultaneously, an increase of “lysoforms” (phospholipids without a fatty acid) and free-fatty acid can occurs. In the present work we reported a preliminary investigation of newborn saliva according to phospholipids, lysoforms and PUFA-derived lipid mediators.

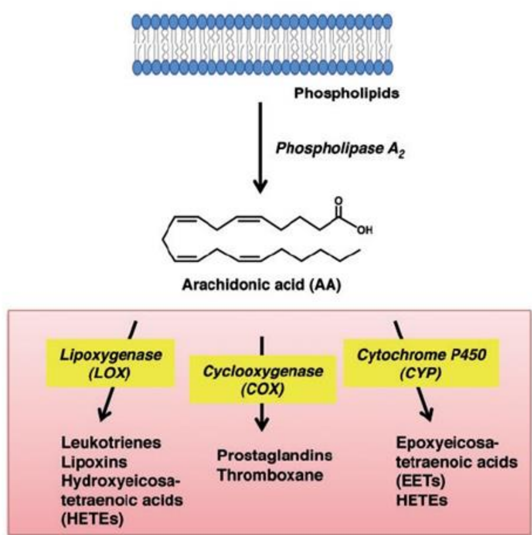


Figure 3-1: eicosanoid production from AA. From “Mediator lipidomics in acute inflammation and resolution”, Makoto Arita.

Salivary lipid pattern is well known and characterized [2], but the conditions in the first 48 hours of life has never been reported, as well as the inflammatory mediators composition. We reported a case study in which saliva were collected according to ethical committee procedures, during the first 48 hours of life of the newborns. The collected samples are characteristic of different status of the inflammation (healthy newborns, potentially infected newborns and infected newborns). Phospholipids and inflammatory mediators were extracted with specific procedures and analysed. For this purpose an high sensitive mass spectrometry (triple quadrupole instrument) platform was employed for the identification of these molecules. In particular, we exploited MS/MS targeted experiments of precursor ions and neutral loss (these experiments will be explained in the paragraph "materials and method") to discriminate the main phospholipids classes and to perform a semi-quantitative analysis [6]. The analysis of PUFA-derived lipid mediators was performed in multiple reaction monitoring (MRM)). In this case we also employed and reported the use of commercial standards to evaluate the limit of detection of these molecules and provide a stronger rationale for the results obtained and shown below. The MRM transitions of PUFA-derived lipid mediators have been monitored according to literature [7,8].

Materials and methods

Chemicals

Acetonitrile

Formic acid (FA)

Sodium dodecyl sulfate (SDS)

Trizma® base

Isopropanol

Methanol

Ammonium formate

Acetic acid

Etilacetate

Glycerol

$PGF_{2\alpha}$ standard

PGE_2 standard

Phosphatidylinositol commercial standard

Phosphatidylserine commercial standard

Phosphatidylglycerol commercial standard

Water, LC-MS Ultra CHROMASOLV, tested for UHPLC-MS

All these reagents are from SIGMA-ALDRICH®

Saline NaCl 0.9%

Software

Analyst AB Sciex® 1.5.1

www.lipidmaps.org online software

SPE

Waters® Oasis HLB Cartridge

LC-MS instruments

Acquity UPLC Waters®-5500 Qtrap AB Sciex®

Phospholipids: targeted mass spectrometry experiment setup

Targeted lipid class profiling is best executed on triple quadrupole/linear ion trap (LIT) instrument that is characterized by high sensibility in precursor ion scanning. For this purpose, UPLC coupled with 5500 Qtrap Absciex® was exploited in positive and negative mode.

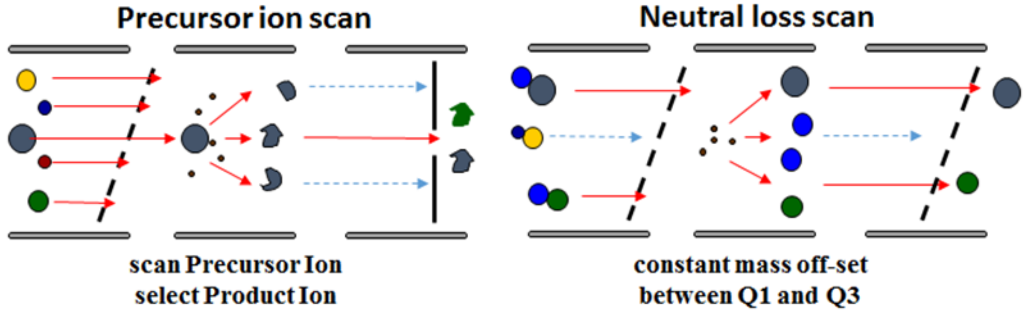


Figure 3-2: description of precursor ion and neutral loss experiments.

Precursor ion scan: The ions formed in the ion source are introduced into Q1 that scans a wide mass range, then ions are fragmented into product ions in Q2. Finally, Q3 only transmits the mass of the selected fragment towards the detector. The spectrum obtained shows all parent ions that produce a selected charged fragment while the chromatogram shows the masses of precursors ions (Figure 3-2).

Neutral loss scan: The ions formed in the ion source are introduced into Q1, then ions are fragmented into product ions in Q2. Finally, Q3 nearly scans the same mass range as Q1 and the mass range is lowered by the mass of a neutral fragment. The spectrum obtained illustrates all parent ions that have lost a selected neutral fragment while chromatogram shows the masses of precursor ions (Figure 3-2).

PHOSPHOLIPID	MOLECULE	FRAGMENT
Phosphatidylcholines		Neutral loss
Lisophosphatidylcholines		Neutral loss
Sphingomyelins		Neutral loss
Phosphatidylethanolamines		Precursor ion
Phosphatidylglycerols		Precursor ion
Phosphatidylinositols		Precursor ion
Phosphatidylserines		Neutral loss
Phosphatidylcholines/lysophosphatidylcholines		Precursor ion

Table 3-1: molecular fragments exploited for the identification of the phospholipids classes.

Precursor ion scan mode and neutral loss scan mode were exploited in positive and negative mode to discriminate the phospholipids classes, according to the product fragment (Table 3-1 and Table 3-2).

Molecule	Fragment	Mode
Phosphatidylcholines	m/z 184	Precursor ion positive
Lisophosphatidylcholines	m/z 184	Precursor ion positive
Phosphatidylethanolamines	m/z 164	Precursor ion positive
Sphyngomyelins	m/z 59	Neutral loss positive
Phosphatidylglycerols	m/z 153	Precursor ion negative
Phosphatidylinositols	m/z 241	Precursor ion negative
Phosphatidylserines	m/z 87	Neutral loss negative

Table 3-2: m/z of the fragments for each class.

MS/MS settings: precursor ion and neutral loss mode	
Paramethers	Values
Mass range	400-920 Da
Scan rate	200 Da/s
Declustering potential	+/-70
Entrance potential	+/-10
Collision energy	+/-40
Collision cell exit potential	+/-13

Table 3-3: MS/MS settings on 5500 Qtrap for phospholipids analysis.

Inflammatory mediators: targeted mass spectrometry experiment setup

MRM analysis was performed to investigate PUFA-derived lipid mediators in saliva samples. In particular, the following transitions were monitored according to the results in literature (Table 3-4):

	Mediator	Q1	Q3	Precursor	Class	Pathway
1	PGE2	351.2	189.5	AA	Prostanoid	COX
2	PGF2α	353.2	193.2	AA	Prostanoid	COX
3	TXB2	369.2	169.1	AA	Thromboxane	COX
4	LTB4	335.2	195.1	AA	Leukotriene	LOX
5	LTD4	495.2	177.1	AA	Leukotriene	LOX
6	12-HETE	319.2	179.2	AA	Alcohol	LOX
7	15-HETE	319.2	219.2	AA	Alcohol	LOX
8	5-HETE	319.2	115.1	AA	Alcohol	LOX
9	9-HETE	319.2	123.0	AA	Alcohol	LOX
10	8-HETE	319.2	155.1	AA	Alcohol	LOX
11	LTB5	333.2	195.1	EPA	Leukotriene	LOX
12	RVE1	349.2	195.0	EPA	Resolvin	LOX
13	TBX3	367.2	169.1	EPA	Thromboxane	COX
14	15-HEPE	317.2	219.2	EPA	Alcohol	LOX
15	12-HEPE	317.2	179.1	EPA	Alcohol	LOX
16	5-HEPE	317.2	115.1	EPA	Alcohol	LOX
17	PGD3	349.2	269.2	EPA	Prostanoid	COX
18	RVD1	375.2	141.0	DHA	Resolvin	LOX
19	9-HODE	295.2	171.1	LINOLEIC ACID	Alcohol	LOX
20	TBX1	371.2	171.1	LINOLEIC ACID	Thromboxane	COX
21	PGF1α	355.2	293.2	LINOLEIC ACID	Prostanoid	COX
22	LXA4	351.0	217.0	AA	Lipoxin	LOX

Table 3-4: transitions for the identifications of inflammatory mediators.

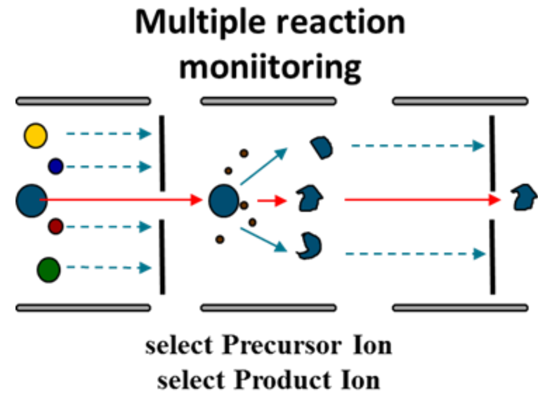


Figure 3-3: description of MRM experiment.

MRM: The ions formed in the ion source are introduced into Q1 where a specific parent mass is selected. Then, ions fragmentation take place in Q2 while the function of Q3 is to select the generated specific fragment mass. Mass Chromatogram shows the intensity of product ions (Figure 3-3).

MRM settings	
Paramethers	Values
Mass range	400-920 Da
Scan rate	200 Da/s
Declustering potential	-70
Entrance potential	-10
Collision energy	-35
Collision cell exit potential	-13

Table 3-5: MRM parameters for inflammatory mediators analysis in 5500 Qtrap.

In the Table 3-6 are reported the concentrations of inflammatory mediators commercial standards (PGE_2 and $PGF_{2\alpha}$) employed to build the calibration line and to determine the limit of detection (LOD) and the limit of quantification (LOQ). The LOD is determined when the signal to noise ratio has value 3. The LOQ is determined when the signal to noise ratio has value 10.

The commercial standards were solubilized in physiological saline solution.

concentration (mg/ml)
0.01
0.001
0.0005
0.0001
0.00001
0.000005
0.0000025

Table 3-6: prostaglandin concentration.

Liquid chromatography experiment setup

PHOSPHOLIPIDS: Liquid chromatography was performed using waters XSELECT CSH130 C18 column 4.6 x 150 mm, 2.5 µm. Phase A was composed of acetonitrile/water (60:40) with 10 mM ammonium formate and 0.1% formic acid, while phase B was isopropanol/acetonitrile (90:10) with 10 mM ammonium formate and 0.1 % formic acid. Flow rate was setting to 0.6 ml/min and column was warmed up to 55°C. In Table 3-7 is reported the gradient:

Time	%A	%B
initial	60	40
1	57	43
1.1	50	50
6	46	54
6.2	30	70
15	10	90
15.5	10	90
16	60	40
30	60	40

Table 3-7: LC method for phospholipids

INFLAMMATORY MEDIATORS: Liquid chromatography was performed using waters acquity UPLC HSS T3 column, 1.8 µm. Phase A was composed of 8.3 mM acetic acid pH 5.7 while phase B was a mixture acetonitrile 65%-methanol 35%. Flow rate was setting to 0.2 ml/min and column was warmed up to 65 °C. In Table 3-8 is reported the gradient:

Time	%A	%B
initial	55	45
12	25	75
14	2	98
25	2	98
26	55	45
35	55	45

Table 3-8: LC method for inflammatory mediators.

Sample preparation

Swabs were introduced in Tris 40mM-SDS 1% solution and three sonication cycles were applied for 3 minutes to improve the extraction of saliva. The extracted mixture was separated in two equal aliquots: one employed for the analysis of phospholipids and the second one for the investigation of the inflammatory mediators. The following protocol was applied to extract and to concentrate the phospholipids:

- ✓ 500 µl of sample solution were treated with 2 ml of chloroform/methanol 2:1;
- ✓ The solution was left at room temperature for 5 minutes;
- ✓ The solution was vortexed for 30 seconds and centrifuged at 4000 rpm for 5 minutes;

- ✓ The organic phase containing the phospholipids was separated and dried with nitrogen;
- ✓ The residual was resuspended in 500 µl of isopropanol/acetonitrile/water 2:1:1.

The following protocol was applied to obtain the purification of the inflammatory mediators using SPE columns:

- ✓ Conditioned with 1 ml of etilacetate;
- ✓ Centrifuged at 500 rpm for 3 minutes;
- ✓ Conditioned with 1 ml of methanol;
- ✓ Centrifuged at 500 rpm for 3 minutes;
- ✓ Conditioned with 1 ml of methanol;
- ✓ Centrifuged at 500 rpm for 3 minutes;
- ✓ Conditioned with 1 ml of 5% methanol + 0.1% acetic acid;
- ✓ Centrifuged at 500 rpm for 3 minutes;
- ✓ Sample loading (500 µl);
- ✓ Centrifuged at 500 rpm for 3 minutes;
- ✓ Second cycle of sample loading;
- ✓ Centrifuged at 500 rpm for 3 minutes;
- ✓ Washed with 1 ml of 5% methanol + 0.1% acetic acid;
- ✓ Centrifuged at 500 rpm for 3 minutes;
- ✓ First elution: 250 µl of methanol;
- ✓ Centrifuged at 500 rpm for 3 minutes;
- ✓ Second elution: 1 ml of ethyl acetate (in a tube with 6 µl of 30% glycerol in methanol);
- ✓ Centrifuged at 500 rpm for 3 minutes;
- ✓ Mixed first and second eluates;
- ✓ Dried with nitrogen;
- ✓ Resuspended in 100 µl of methanol.

These experiments were performed to obtain a comparison between healthy newborns saliva (day 1 and day 2) and infected newborn saliva. In particular, we compared a pool (N=2 newborn) for each group of analysis (Table 3-9).

- Group 1: healthy newborn saliva;
- Group 2: potentially infected newborn saliva;
- Group 3: infected newborn saliva.

Analytical reproducibility was evaluated performing a double extraction for each pool of analysis.

NEWBORN	GROUP	SEX	AGE (WEEKS)	WEIGHT (g)	DELIVERY	SWABS
5922	1	M	39	3170	SPONTANEOUS	EN. FAECALIS
9967	1	M	39	3400	SPONTANEOUS	CANDIDA
2034	2	M	39	3700	SPONTANEOUS	NEG
5320	2	F	37	3270	SPONTANEOUS	NEG
10623	3	M	41	3850	CAESAREAN	POSITIVE SBE
10639	3	M	39	3750	SPONTANEOUS	POSITIVE SBE

Table 3-9: data samples.

Results and discussion

Phospholipids

The peak areas of phosphatidylcholines (PC), lysophosphatidylcholines (LPC) and sphyngomyelins (SM) detected in group 1, group 2 and group 3 are reported in Table 3-10, where also the CV% is reported to evaluate the analytical reproducibility.

GROUP 1			GROUP 2			GROUP 3		
molecule	Average area (n=2)	CV%	molecule	Average area (n=2)	CV%	molecule	Average area (n=2)	CV%
LPC (C18:1)	5,75E+07	0,02%	LPC (C18:1)	6,68E+07	24,1%	LPC (C18:1)	1,02E+08	7,2%
LPC (C16:0)	1,87E+07	1,7%	LPC (C16:0)	3,30E+07	7,2%	LPC (C16:0)	3,32E+07	10,9%
LPC (16:1)	1,38E+07	3,4%	LPC (16:1)	9,32E+06	28,4%	LPC (16:1)	1,75E+07	10,2%
SM (dC18:1,NC16:0)	3,53E+08	3,5%	SM (dC18:1,NC16:0)	3,28E+08	19,7%	SM (dC18:1,NC16:0)	4,00E+08	3,8%
PC (C16:0,C14:0)	2,39E+08	0,6%	PC (C16:0,C14:0)	2,15E+08	17,3%	PC (C16:0,C14:0)	2,80E+08	11,6%
m/z 718,3 (attachment A)	2,57E+07	6%	m/z 718,3 (attachment A)	3,29E+07	18,1%	m/z 718,3 (attachment A)	1,66E+07	0,1%
PC (C16:0,C16:1)	9,88E+07	3,7%	PC (C16:0,C16:1)	2,38E+08	21,9%	PC (C16:0,C16:1)	1,07E+08	19,8%
PC (C16:0,C16:0)	1,01E+08	5,7%	PC (C16:0,C16:0)	5,78E+08	14,9%	PC (C16:0,C16:0)	9,19E+07	5,9%
m/z 744,8 (attachment A)	9,88E+07	1,6%	m/z 744,8 (attachment A)	1,06E+08	16,9%	m/z 744,8 (attachment A)	8,55E+07	9,7%
PC (C15:0, C18:1)	7,71E+07	1,3%	PC (C15:0, C18:1)	9,17E+07	21,9%	PC (C15:0, C18:1)	4,77E+07	5,5%
PC (C16:0, C18:2 / C16:1, C18:1)	1,35E+08	1,3%	PC (C16:0, C18:2 / C16:1, C18:1)	1,76E+08	24,3%	PC (C16:0, C18:2 / C16:1, C18:1)	1,43E+08	6,7%
PC (C16:0,C18:1)	3,61E+08	3,2%	PC (C16:0,C18:1)	6,07E+08	25,1%	PC (C16:0,C18:1)	4,59E+08	8,2%
PC (C17:0,18:1/C17:1,18:0)	3,15E+07	1,4%	PC (C17:0,18:1/C17:1,18:0)	3,28E+07	20,8%	PC (C17:0,18:1/C17:1,18:0)	2,09E+07	3,14%
PC (C18:1,C18:1	1,90E+08	0,8%	PC (C18:1,C18:1	2,71E+08	16,7%	PC (C18:1,C18:1	2,36E+08	10%
PC (C18:0, C18:1)	2,07E+08	3,1%	PC (C18:0, C18:1)	2,66E+08	14%	PC (C18:0, C18:1)	2,82E+08	0
m/z 675,7 (attachment A)	9,61E+06	10,1%	m/z 675,7 (attachment A)	1,34E+07	19%	m/z 675,7 (attachment A)	1,44E+07	7,4%
m/z 677,8 (attachment A)	7,36E+06	6,2%	m/z 677,8 (attachment A)	8,18E+06	23%	m/z 677,8 (attachment A)	8,81E+06	9,1%

Table 3-10: peak area and CV% for PC and SM.

The peak area and CV% of phosphatidylethanolamines (PE) and lysophosphatidylethanolamines (LPE) detected in group 1, group 2 and group 3 are reported in Table 3-11.

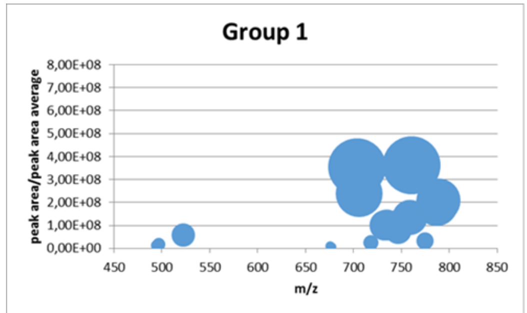
GROUP 1			GROUP 2			GROUP 3		
molecule	Average area (n=2)	CV%	molecule	Average area (n=2)	CV%	molecule	Average area (n=2)	CV%
LPE (C13:0)	4,08E+06	5,4%	LPE (C13:0)	3,33E+06	12,1%	LPE (C13:0)	1,33E+06	2,7%
m/z 696,7 (attachment B)	9,21E+05	5%	m/z 696,7 (attachment B)	9,10E+05	15,1%	m/z 696,7 (attachment B)	7,49E+05	31,1%
PE (C20:4, C18:0vinil)	5,01E+06	11,4%	PE (C20:4, C18:0vinil)	5,59E+06	21,7%	PE (C20:4, C18:0vinil)	4,53E+06	3,5%
PE (C18:2, C18:2)	1,15E+06	9,4%	PE (C18:2, C18:2)	1,17E+06	12,7%	PE (C18:2, C18:2)	1,56E+06	9,7%
PE (C18:0, C18:1)	1,58E+06	2,2%	PE (C18:0, C18:1)	1,56E+06	34,6%	PE (C18:0, C18:1)	1,94E+06	4,8%
PE (C18:1, C20:4)	8,43E+05	17,7%	PE (C18:1, C20:4)	1,06E+06	18,9%	PE (C18:1, C20:4)	9,00E+05	8,8%

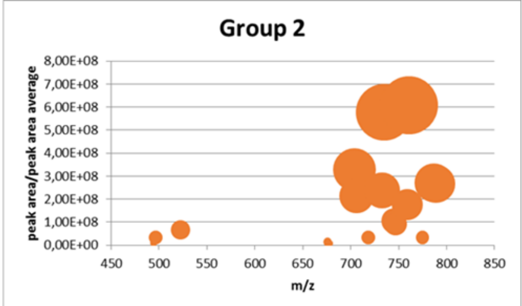
Table 3-11: peak area and CV% for PE.

Only these three classes of phospholipids were detected in the saliva samples. In particular, no signals were observed in negative polarity. This modality is considerably less sensitive than the positive one and with the aim to evaluate the limit of detection we previously analysed a phospholipids standard mixture. Standard solution of phospholipids was also used for the optimization of the MS methods. A LOD concentration of 0.1 mg/ml was estimated in negative mode for phosohatidylinositols, 0.4 mg/ml for the phosphatidylserines and 0.45 mg/ml for phosphatidylglycerols. All the PC, LPC and SM identified are reported in Table 3-10. The m/z 703.8 is the sphyngomyelin (dC18:1, NC16:0) that gives a signal in the experiment of precursor ion 184 [9]. All the analysed phospholipids are ionized as single charge. No signals were detected in the neutral loss 59 experiment. In our opinion an interesting value is represented by the CV%. This value is the percentage ratio between the standard deviation and the average and, in the present

work, it permits to evaluate the analytical reproducibility of the extraction method. Considering the Table 3-10, high values of CV% were observed especially in the sample belonging to group 2. This trend is observed also in Table 3-11 but to a lesser extent. We can postulate that the majority of the variability is ascribed to the extraction method, which is suitable to extract and to concentrate the phospholipids of a biological sample but it could not be optimal for heterogeneous biological suspension. This could affects the reproducibility and the yield. It can be argued that this variability could reflect a significant biological diversity in between the group samples especially regarding the results of Group 2.

Taking into account this analytical result, an attempt was made to highlight qualitative (composition) global differences in the pattern of PC and PE patterns. In Figure 3-4 and in Figure 3-5 are reported the lipid profile of the phospholipid detected classes. In particular, in these charts it is represented for each group of babies the impact of the peak area of a single phospholipid molecule (bubble diameter) on the whole class average peak area. This particular plotting allows a quick overview of the lipid pattern within each group and define a potential fingerprint of selected lipids of each group of babies.





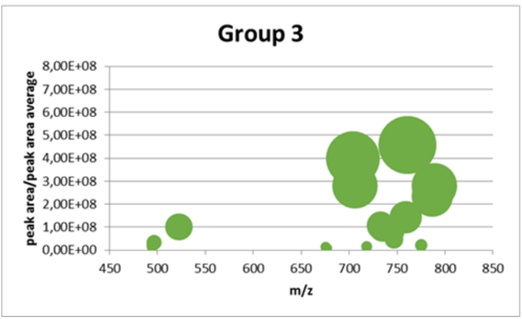
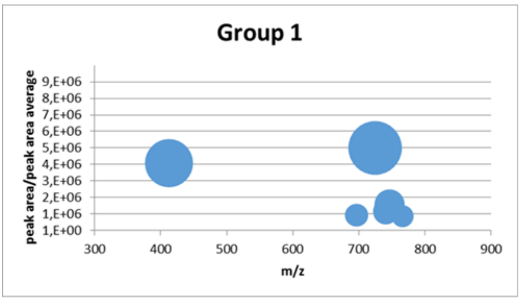
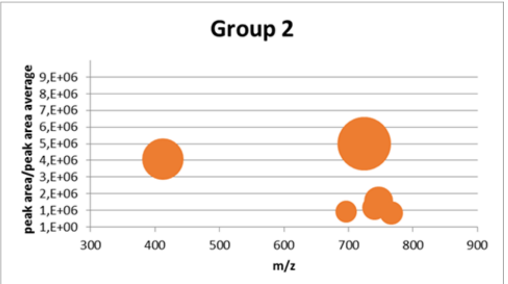


Figure 3-4: PC and SM profiling of newborns saliva in different inflammatory conditions.

Observing Figure 3-4 we can affirm that no significant differences are highlighted comparing the group 1 (healthy newborn saliva) and group 3 (infected newborn saliva). Weak differences instead are observed in the group 2 saliva, which actually was the group mostly affected by the lower analytical reproducibility based on the results reported before. This difference worths a further careful investigation.





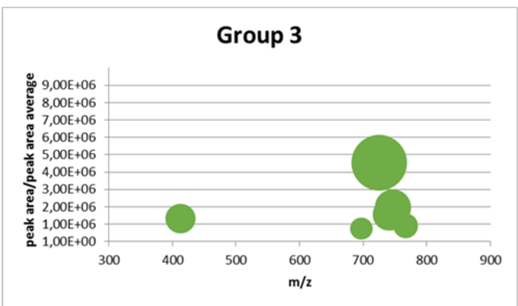


Figure 3-5: PE profiling of newborns saliva in different inflammatory conditions.

In the graphs above, no significant differences were detected in the PE fingerprints in between the three groups (Figure 3-5) , only exception is the analyte at m/z 412,4. This m/z corresponds to the lysoform PE (C13:0/0:0) and this species is more abundant in group 1 and group 2 compared to group 3 with respect to the other PE molecules. These results, indicating that no significant differences can be detected on the basis of the qualitative fingerprint of the PE family in between the groups, more strongly support the hypothesis that a more careful investigation of the PC pattern previously described (Figure 3-4) needs to take place, to figure out whether the PC fingerprint could be a marker of a particular systemic condition.

Lysoforms

The lysoforms and in particular the lysophosphatidylcholine m/z 496.5 (LPC C16:0) are very diagnostic of an inflammatory condition. The phospholipase A2 (PLA2) releases arachidonic acid and other fatty acids from membrane phospholipids. The AA is then metabolized through the pathway of cyclooxygenase and lipoxygenase to produce the eicosanoids. This process also lead to the production of lysophospholipids. The lysoforms are the precursors of lysophosphatidic acid and sphingosine-1-phosphate, which have a

crucial role in inflammation. The hydrolysis of the DPPC (dipalmitoilphosphatidylcholine, PC C16:0, C16:0) is an early event that leads to the accumulation of lysophophatidylcholines (LPC). LPC has a crucial role in the onset of acute lung injury. It exhibits a chemotactic action against mononuclear cells playing a key role in cell recruitment during inflammation.

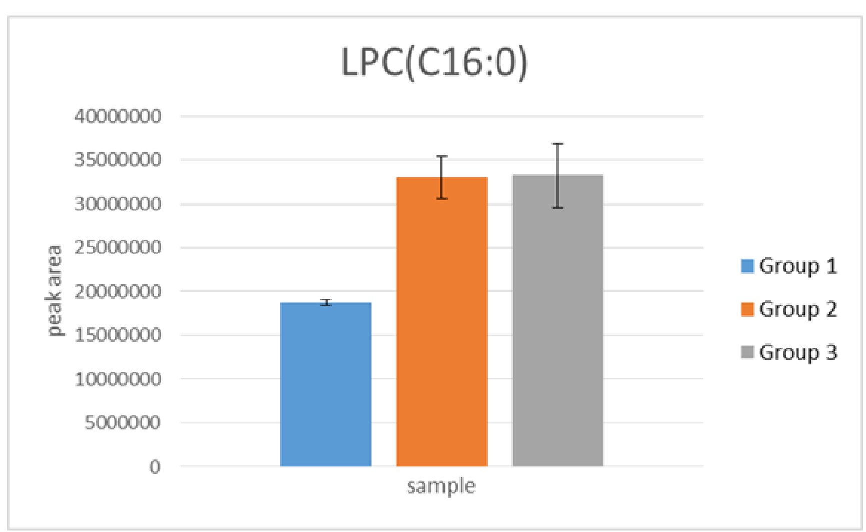


Figure 3-6: LPC (C16:0) detection in group 1, group 2 and group 3. (N=2 injections).

The Figure 3-6 shows a progressive increase of the LPC (C16:0) according to the inflammatory condition. The same trend is also observed for other two detected LPCs (Figure 3-7 and Figure 3-8).

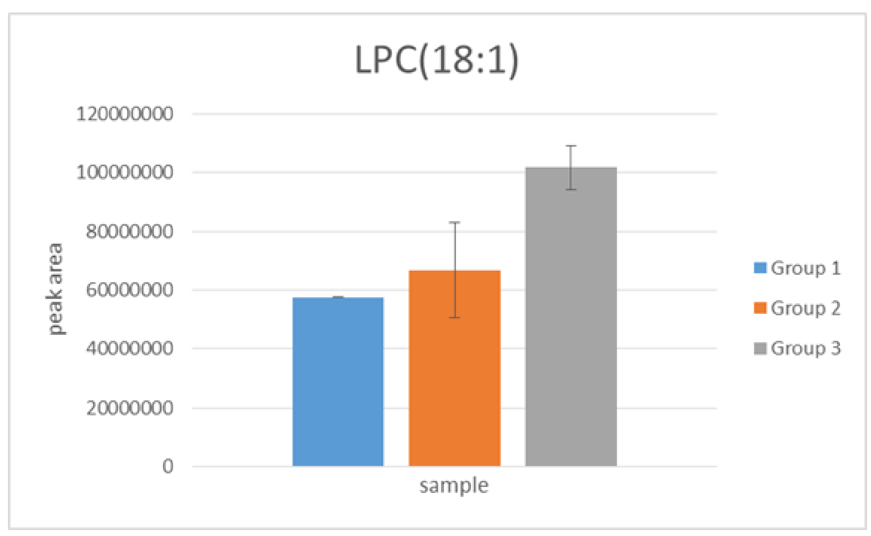


Figure 3-7: LPC (C18:1) detection in group 1, group 2 and group 3 (N=2 injections).

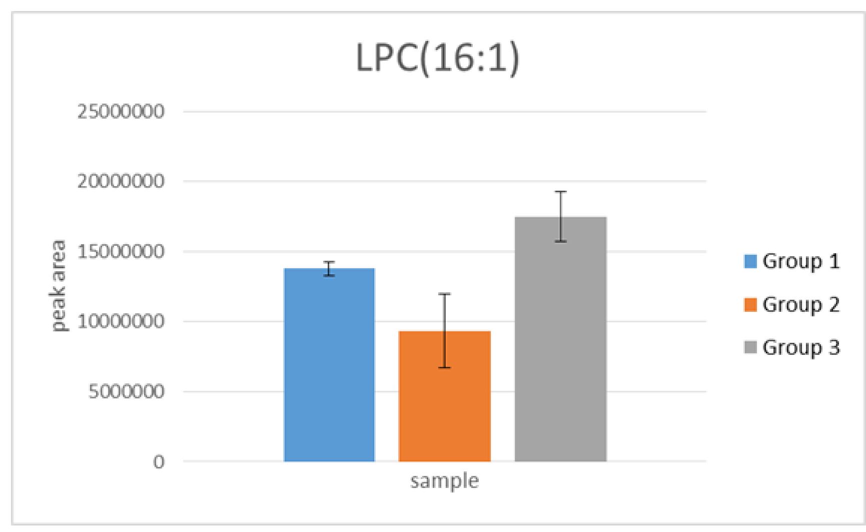


Figure 3-8: LPC (C16:1) detection in group 1, group 2 and group 3 (N=2 injections).

The LPC composition appears to be representative of an inflammatory condition that increase from group 1 to group 3. LPC C16:0, the lysoform that previously was shown to increase in group 3 is the main product of DPPC which is in turn the main target of the phospholipase A2 activity.

Inflammatory mediators

After the characterization according to lipid profiling, we also investigated the samples to look for the PUFA-derived lipid mediators (inflammatory mediators). The data resulting from the lysophosphatidylcholines, that could be a proof of ongoing inflammation, can be supported by the detection of these molecules that play a key role in inflammation. We have selected the mediators (Table 3-4 in "materials and methods") to be followed in MRM with specific transitions according to the literature. For the selection we considered the enzyme of processing (LOX or COX), reaction intermediates, final products, pro-inflammation molecules and molecules that resolve the inflammation. In other words, we covered all the pathways of the arachidonic acid and also different stages of inflammation. Differently from phospholipids, the PUFA-derived lipid mediators require a solid phase extraction. The extraction was performed using the Waters Oasis HLB (hydrophilic-lipophilic-balanced) Cartridge. These cartridges are made from a specific ratio of two monomers, the hydrophilic N-vinylpyrrolidone and the lipophilic divinylbenzene. In comparison with the traditional silica-based C18 SPE, these cartridges guarantee a higher relative hydrophobic retention. A solid phase extraction is a crucial and fundamental step to concentrate these analytes that are usually present in trace amount in biological samples. Furthermore, these steps also allow to clean up the samples with the removal of proteins and hydrophilic solutes.

We analyzed the three groups (two newborns for each group) that were previously analysed according to the phospholipids composition in order to evaluate a potential correlation with the inflammatory mediators. Each preparation was performed in double to evaluate the analytical reproducibility.

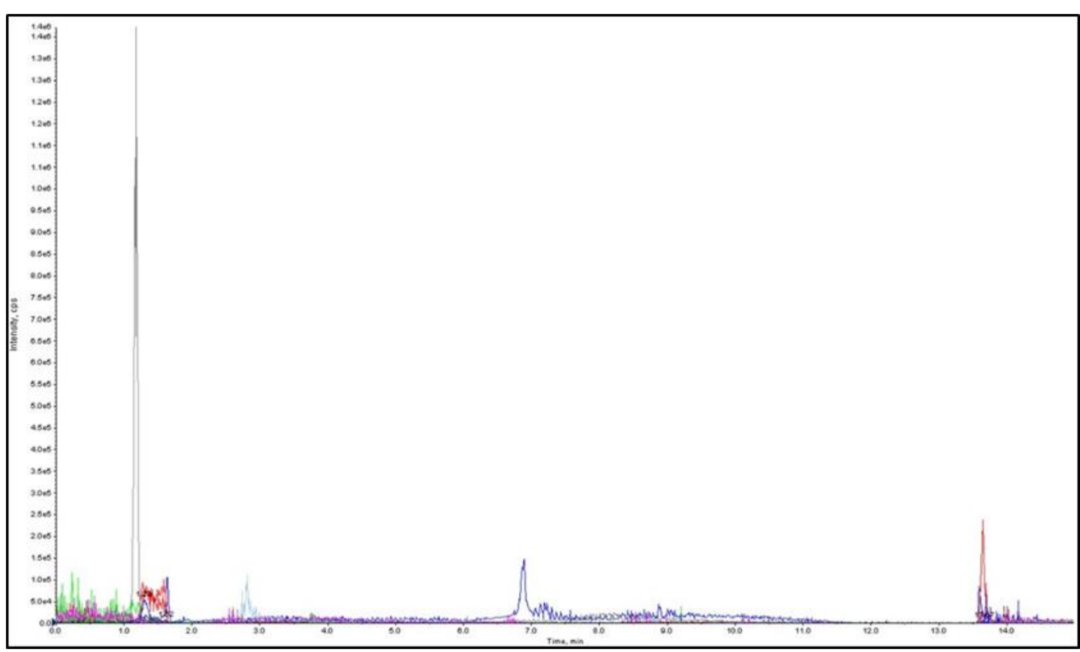


Figure 3-9: MRM spectrum of inflammatory mediators within healthy newborn saliva.

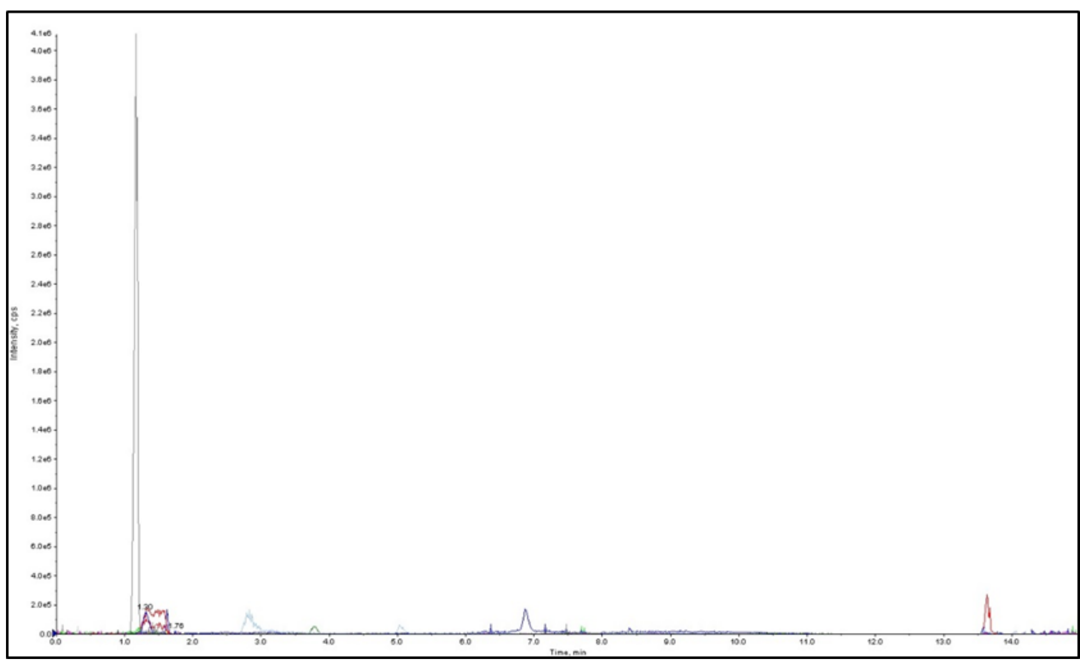


Figure 3-10: MRM spectrum of inflammatory mediators within potentially infected newborn saliva.

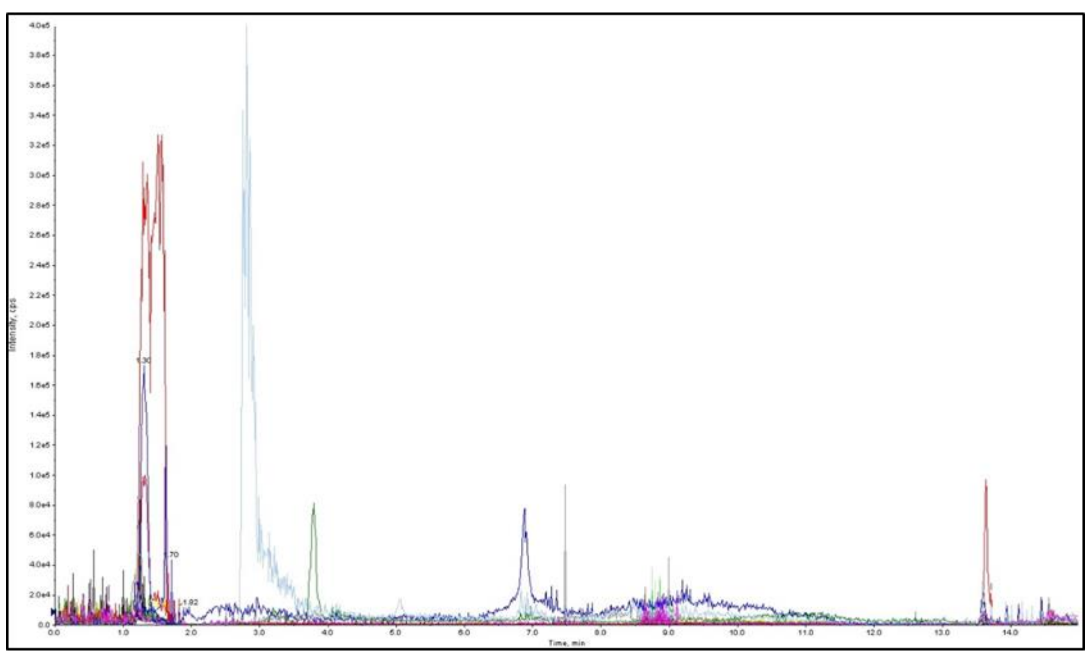


Figure 3-11: MRM spectrum of inflammatory mediators within infected newborn saliva.

In Figure 3-9, Figure 3-10 and Figure 3-11 the MRM spectra obtained for each group are reported. Based on the retention time and on the MRM transitions, it can be stated that PUFA-derived lipid mediators were below the limit of detection of the experimental platform adopted in this study.

Quantitative considerations about eicosanoids

The data collected for the inflammatory mediators can be summarized with the hypothesis that these molecules, in these saliva samples, are under the limit of detection (LOD). Therefore, for the future implementation of the analytical method to investigate PUFA-derived lipid mediators, it is relevant to know both the minimum detectable concentration and the minimum quantifiable concentration. For this purpose PGE_2 commercial standard and $PGF_{2\alpha}$ commercial standards were pooled in saline solution and analysed in MRM with 5500 Qtrap at different concentration to obtain the calibration line (Figure 3-12).

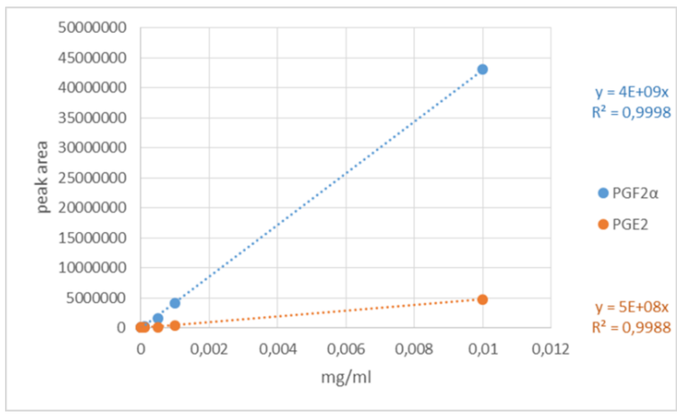


Figure 3-12: calibration line of the eicosanoids commercial standards.

The LOD and the LOQ were estimated according to the signal to noise ratio and are reported in Table 3-12. The MRM spectrum is reported in Figure 3-13.

	$PGF_{2\alpha}$	PGE_2
LOQ	5 ng/ml	10 ng/ml
LOD	2.5 ng/ml	5 ng/ml

Table 3-12: LOQ and LOD of eicosanoids standard.

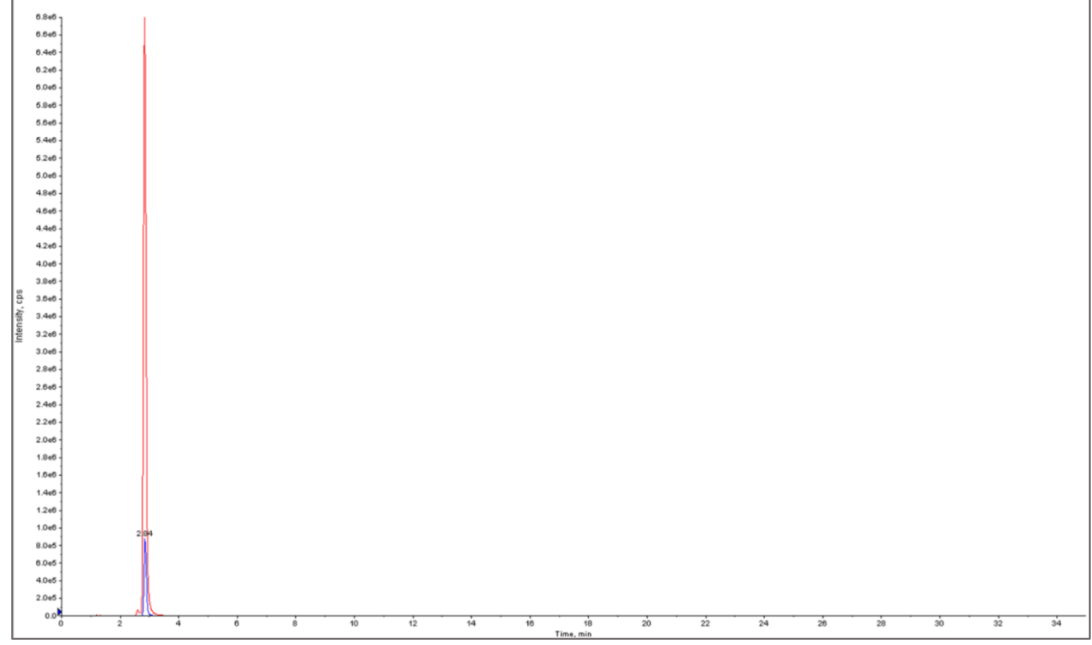


Figure 3-13: MRM spectrum of eicosanoids standard solution (coelution of the two analysed prostaglandins).

In Figure 3-14 the structures of the two prostaglandins employed to build the calibration line are reported.

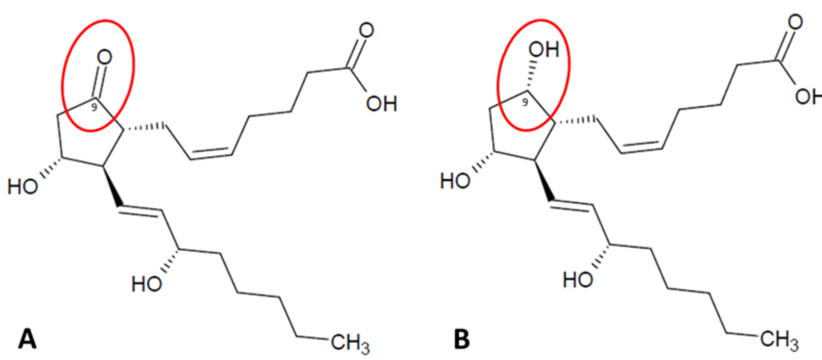


Figure 3-14: A) PGE_2; B) $PGF_{2\alpha}$.

As shown in Table 3-13 and confirmed observing the spectrum in Figure 3-13, PGE_2 and $PGF_{2\alpha}$ ionized in different ways. These two molecules differ only in the oxidation state of carbon 9 (red circle). This chemical difference does not affect the elution time (as shown in the spectrum) but has dramatically impacts on the ionization in ESI source. This is well appreciated in LOD and LOQ values, but also in the peak area. In fact, starting from equal concentration values (0.01 mg/ml) the percentage MRM ratio, that is approximately constant for each concentration, confirms that the peak area of the PGE_2 is more or less the 10% of the $PGF_{2\alpha}$ peak area (Table 3-13).

concentration mg/ml	peak area $PGF_{2\alpha}$	peak area PGE_2	% PGE_2/ $PGF_{2\alpha}$ MRM ratio
0,000005	1,49E+04	1,85E+03	**12%**
0,00001	1,91E+04	2,25E+03	**12%**
0,0001	2,52E+05	2,25E+04	**9%**
0,0005	1,62E+06	1,41E+05	**9%**
0,001	4,09E+06	3,60E+05	**9%**
0,01	4,31E+07	4,73E+06	**11%**

Table 3-13: MRM ratio (N=2 injections).

Observing the chemical structures in Figure 3-14, an explanation of this behaviour can be attributed to the different number of negatively ionisable sites. The PGE_2 is composed of three sites that in ESI are negatively ionisable, in particular one carboxylic acid and two hydroxyl functions, while in $PGF_{2\alpha}$ the ionisable sites are four: one carboxylic acid and three hydroxyl functions. A second hypothesis might be that the two molecules show different characteristics of fragmentation. In other words, in this MRM analysis a single transition has been followed for each compound. Observing the peak area we could hypothesize that $PGF_{2\alpha}$ is easier to fragment than PGE_2, and consequently in semi-quantitative MRM analysis its signal is higher than the signal of the PGE_2. Extending these hypothesis to a biological sample it is quite clear that the detection of the inflammatory mediators is very difficult. In fact these molecules, having similar chemical-physical characteristics, are hardly discriminable in liquid chromatography. Nevertheless these differences are responsible of a specific behaviour during mass spectrometry analysis. Therefore, as shown in Table 3-13 and in Figure 3-13, the peak area might be surprisingly low considering the concentration. To significantly increase the confidence in the identification of inflammatory mediators, internal standard should be mandatory and more representative of the eicosanoids population than only two standards as carried out in the present work.

Healthy… or infected?

In the present work we have investigated the ongoing inflammation in newborn saliva considering three points of view: the decrease of phospholipids amount, the increase of lysoforms and the detection of inflammatory mediators. These three mechanisms, belonging to different phases of the inflammation, are complementary and equally important to obtain an overview of an inflammatory condition. The results of this investigation can be also interpreted considering the protein pattern identified and described in the first chapter. In particular, two hypothesis might help us to explain the similarity between healthy and infected saliva. The first one is based on a principle that we already illustrated about the proteome, in particular the presence of some identified proteins confirms that during the first 48 hours of life newborns are "constitutively" in an

inflammatory condition. Transposing this principle to phospholipids and mediators pattern, we can't exclude the possibility that also these molecules, like proteins, are referred to a constitutional mild inflammatory state also observed for healthy newborns. Therefore, this hypothesis can support why the phospholipids and inflammatory mediators profiles are not so different comparing a healthy and an infected condition. The second hypothesis is related with a protein identification of the first chapter. Why the phospholipids haven't significantly changed their concentration in inflammation? Why only a weak variation of the lysoforms has been observed? Considering these two questions, the results obtained with the inflammatory mediators are not unexpected. Which is the common precursor on the basis of all these processes? Phospholipase A2 activation represents the triggering event for these processes (Figure 3-15).

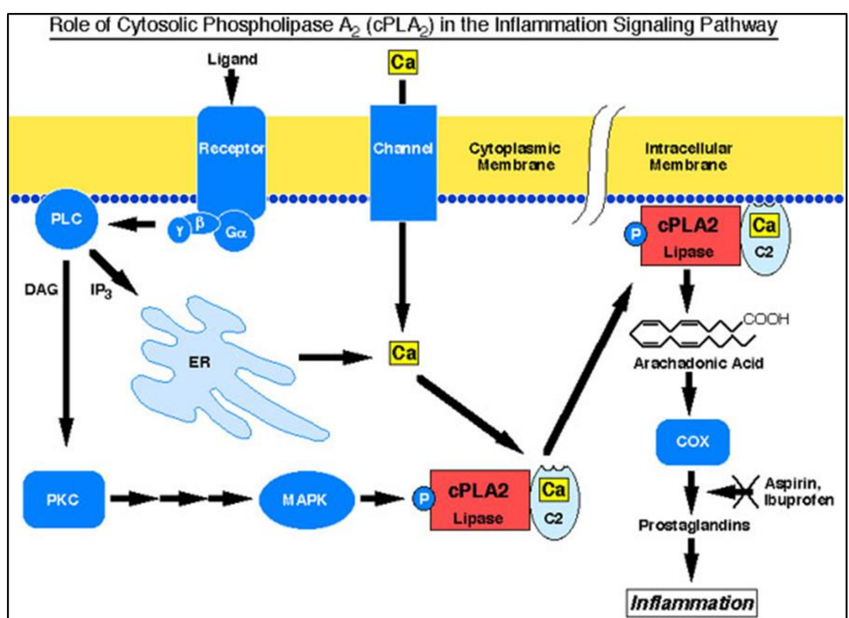


Figure 3-15: PLA2 mechanism.

In the first chapter, we identified in newborn saliva the protein annexin A1. Annexin A1 both suppresses phospholipase A2, thereby blocking eicosanoid production, and inhibits various leukocyte inflammatory events. Since phospholipase A2 is required for the biosynthesis of the potent inflammatory mediators, prostaglandins, and leukotrienes, annexin A1 may have potential anti-inflammatory activity. For example, the

glucocorticoids drugs stimulate production of this protein. In other words, the presence of this protein in saliva might prevent the investigation of an inflammatory condition exploiting the lipid pattern. Annexin A1 activity is probably the main explanation for the weak differences observed comparing healthy and infected newborn saliva.

Conclusions

In the present work we investigated the newborn saliva with the aim to characterize the lipid pattern in healthy and infected conditions. A high sensitive LC-MS platform has been exploited for MS/MS targeted experiments to identify and to obtain a semi-quantitative valuation of the main phospholipid classes. Results were obtained for sphingomyelins, phosphatidylcholines, lysophosphatidylcholines, phosphatidylethanolamines and lysophosphatidylethanolamines. Overall, no huge significant differences were observed in saliva of newborns at different stage of inflammation. A weak trend was observed for the lysophosphatidylcholines in inflammatory conditions, and a different phosphatydilcholine fingerprint was determined for Group 2. The mediators of inflammation are under the limit of detection. The investigation of these molecules in saliva will requires the use of commercial standard to optimize both the sample preparation and the concentration. Finally, we can preliminary say that the present results are in agreement with the results of the first chapter were a "physiological" inflammatory conditions were detected also in healthy newborn saliva. The presence of the protein annexin A1, that is a phospholipase A2 inhibitor, is a valuable reason why no significant variations has been observed in the lipid pattern. The present work provides an overview of lipids composition in newborn saliva samples collected in the first 48 hours of life. Therefore this study lays the groundwork for future works for a deeper investigation to look for eicosanoids in saliva and to describe a complete framework of the inflammatory pattern in newborn saliva.

Overall, a proof of concept for the discovery of diagnostic assays based on newborn saliva samples was provided. The results clearly suggest that the development of such a tool cannot be framed within the knowledge acquired on adult samples, but a careful and detailed validation needs to be carried out in light of the fact that the newborn state within the 48 hours after delivery represents *per se* a peculiar physiological condition which cannot be easily predicted even when "healthy" conditions are envisaged (Group1).

Attachment A: potential identifications from Lipidmaps

Input Mass	Matched Mass	Delta	Abbreviation	Formula	Ion
718,3	718,5382	0,2382	PC(31:1)	$C_{39}H_{77}NO_8P$	[M+H]+
718,3	718,5745	0,2745	LPC(32:1)	$C_{40}H_{81}NO_7P$	[M+H]+
718,3	718,5745	0,2745	PC(O-32:1)	$C_{40}H_{81}NO_7P$	[M+H]+
718,3	718,5745	0,2745	PC(P-32:0)	$C_{40}H_{81}NO_7P$	[M+H]+
Input Mass	**Matched Mass**	**Delta**	**Abbreviation**	**Formula**	**Ion**
744,8	744,5902	0,2098	PC(O-34:2)	$C_{42}H_{83}NO_7P$	[M+H]+
744,8	744,5902	0,2098	PC(P-34:1)	$C_{42}H_{83}NO_7P$	[M+H]+
744,8	744,5538	0,2462	PC(33:2)	$C_{41}H_{79}NO_8P$	[M+H]+
Input Mass	**Matched Mass**	**Delta**	**Abbreviation**	**Formula**	**Ion**
675,7	676,4912	0,7912	PC(28:1)	$C_{36}H_{71}NO_8P$	[M+H]+
675,7	676,5276	0,8276	PC(P-29:0)	$C_{37}H_{75}NO_7P$	[M+H]+
Input Mass	**Matched Mass**	**Delta**	**Abbreviation**	**Formula**	**Ion**
677,7	678,5069	0,8069	PC(28:0)	$C_{36}H_{73}NO_8P$	[M+H]+
677,7	678,5432	0,8432	LPC(29:0)	$C_{37}H_{77}NO_7P$	[M+H]+
677,7	678,5432	0,8432	PC(O-29:0)	$C_{37}H_{77}NO_7P$	[M+H]+

Attachment B: potential identifications from Lipidmaps

Input Mass	Matched Mass	Delta	Abbreviation	Formula	Ion
696,7	696,4963	0,2037	LPE(34:5)	$C_{39}H_{71}NO_7P$	[M+H]+
696,7	696,4963	0,2037	PE(P-34:4)	$C_{39}H_{71}NO_7P$	[M+H]+
696,7	696,4599	0,2401	PC(30:5)	$C_{38}H_{67}NO_8P$	[M+H]+
696,7	696,4599	0,2401	PE(33:5)	$C_{38}H_{67}NO_8P$	[M+H]+

Bibliography

1) Janciauskiene, S. M., Bals, R., Koczulla, R., Vogelmeier, C., Köhnlein, T., & Welte, T. (2011). The discovery of α1-antitrypsin and its role in health and disease. Respiratory medicine, 105(8), 1129-1139.
2) Slomiany, B. L., Murty, V. L., & Slomiany, A. (1985). Salivary lipids in health and disease. Progress in lipid research, 24(4), 311-324.
3) Dowhan, W. (1994). Role of Phospholipids in Cell Function. In Biological Membranes: Structure, Biogenesis and Dynamics (pp. 1-21). Springer Berlin Heidelberg.
4) German, J. B., Gillies, L. A., Smilowitz, J. T., Zivkovic, A. M., & Watkins, S. M. (2007). Lipidomics and lipid profiling in metabolomics. Current opinion in lipidology, 18(1), 66-71.
5) Arita, M. (2012). Mediator lipidomics in acute inflammation and resolution. Journal of biochemistry, 152(4), 313-319.
6) Pelizzi, N., Catinella, S., Barboso, S., & Zanol, M. (2002). Different electrospray tandem mass spectrometric approaches for rapid characterization of phospholipid classes of Curosurf®, a natural pulmonary surfactant. Rapid communications in mass spectrometry, 16(24), 2215-2220.
7) Strassburg, K. (2013). Targeted lipidomics of oxylipins (oxygenated fatty acids). Waters App note.
8) Masoodi, M., Mir, A. A., Petasis, N. A., Serhan, C. N., & Nicolaou, A. (2008). Simultaneous lipidomic analysis of three families of bioactive lipid mediators leukotrienes, resolvins, protectins and related hydroxy-fatty acids by liquid chromatography/electrospray ionisation tandem mass spectrometry. Rapid Communications in Mass Spectrometry, 22(2), 75-83.
9) M.C. Sullards, Analysis of sphingomyelin, glucosylceramide, ceramide, sphingosine, and sphingosine 1-phosphate by tandem mass spectrometry, *Methods Enzymol.*, 2000, 312, 32-45.