UNIVERSITÀ DEGLI STUDI DELLA TUSCIA DI VITERBO

DIPARTIMENTO DI SCIENZE ECOLOGICHE E BIOLOGICHE (DEB)

CORSO DI DOTTORATO DI RICERCA

IN GENETICA E BIOLOGIA CELLULARE XXIV Ciclo.

Proteomics and Metabolomics:

New "omics" platform technologies for biological applications

BIO/11

Coordinatore: Prof. Giorgio Prantera

Firma:……..

Tutor: Dott.ssa Anna Maria Timperio

Firma:…….....

Dottorando: Federica Gevi

Firma:……..……

“You can never write the word end.
As soon as you stop it is time to start again.”

(Pablo Picasso)

-Abstract-

The word '*omics*', is derived from the Greek suffix, '*ome*' (meaning all, every, whole, or complete) and refers to the scientific areas in molecular biology, which aim for a global view on biological systems in order to understand life and its organization as holistic existence. In the mid 1990's, rapid evolution and growing knowledge concerning genomes (the complete genetic sequence of an organism) led to a quick development of the full range of '*omics*' approaches (genomics, transcriptomics, proteomics, metabolomics). The '*omics*' approaches are often applied to investigate samples collected from different populations or upon different physiological states (disease *vs* healthy, treated *vs* untreated, temperature stressed *vs* unstressed, etc.), aiming to find molecules that would differentiate between these classes of samples. This PhD thesis focuses on the application of new "omics" sciences to characterized samples of biomedical interest such as pharmaceutical products. The first section is focused on the characterization of important pharmaceutical products used by hemophiliac patients, trough proteomics approaches. The first objective of this study was to compare the heterogeneity and the high purity of three rFVIII commercially available rFVIII concentrates (Advates , Helixate NexGens and Refacto) before and after thrombin digestion. The second object was to extend the same study to plasma derived factor VIII (Beriate and Emoclot). The second part of my PhD project was designed mainly to the development and optimization of a new HPLC-MS method for metabolomics. Metabolomics is defined as the study of metabolome or the complex system of metabolites that, as intermediates of biochemical reactions, play an important role in the connection between the different pathways involved in a living cell. The metabolites can be seen as the end product of gene expression or activity of the protein (enzyme), thus defining the biochemical phenotype of a biological system. In the last part of my research I used this analytical approach to investigate metabolites in blastocoele fluid that is a routine waste product prior to pre-implantation blastocyst vitrification. Generally embryo assessment is currently performed through the analysis of its morphology. The aim of this study was directed to identify a correlation between quali-quantitative profiles of small molecules of metabolic interest and the outcome of embryo transfer through a simple HPLC-MS assay. Finally I used metabolomics approach for investigation of metabolomics changes in red blood cell during storage. The aim of this work was to confirm and expand existing literature about the rapid fall of glycolytic rate and accumulation of glycolysis end products during first 42 days of storage. Proteome and metabolomic approaches, are widely used for studies complex biological systems and can be used for applications in biological areas.

Abbreviations

CE	Capillary Electrophoresis
ESI	Electrospray ionization
FA	Formic acid
FIVET	Fertilization in Vitri and Embryo Tranfert
FTICR	Fourier Transform Ion Cyclotron Resonance
HPLC	High Performance Liquid Chromatography
LC	Liquid Chromatography
GC	Gas Chromatography
MALDI	Matrix Assisted Laser Desorption/Ionization
MRM	Multiple Reaction Monitoring
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
NMR	Nuclear Magnetic Resonance
SRM	Single Reaction Monitoring
TOF	Time of Flight
UV	Ultraviolet

Index

Chapter 1: Proteomics

Chapter 2: Recombinant clotting factor VIII concentrates: Heterogeneity and high-purity evaluation

Chapter 3: Comparison among plasma-derived clotting Factor VIII by using monodimensional gel electrophoresis and mass spectrometry

Chapter 4: Metabolomics: A new Rapid Resolution Reverse-Phase HPLC strategy to investigate various metabolic species in different biological models

Chapter 5:A mass spectrometry-based targeted metabolomics strategy of human blastocoele fluid: a promising tool in fertility research.

Chapter 6:Alterations of Red Blood Cell metabolome during cold liquid storage of erythrocyte concentrates in CPD-SAGM

Chapter 1

1.1 Proteomics

The word “Proteome” was introduced for the first time by Marc Wilkins in 1994 during the scientific meeting “2D electrophoresis: from protein maps to genomes” held at Siena, Italy. The proteome has been defined as the entire complement of proteins expressed by a cell, organism, or tissue type. At the same time the word “Proteomics” appeared, in analogy with “Genomics”, to describe the scientific field that studies the proteome using a wide range of separation, analytical and bioinformatics tools to characterize and measure the result of gene expression at one time, under specific conditions, in a cell, tissue or organism (Traini 1998) infact proteomics is defined as : “The analysis of the entire protein complement expressed by a genome” (Wasinger 1995). In general, proteomic approaches can be used:

(a) for proteome profiling,

(b) for comparative expression analysis of two or more protein samples,

(c) for the localization and identification of posttranslational modifications,

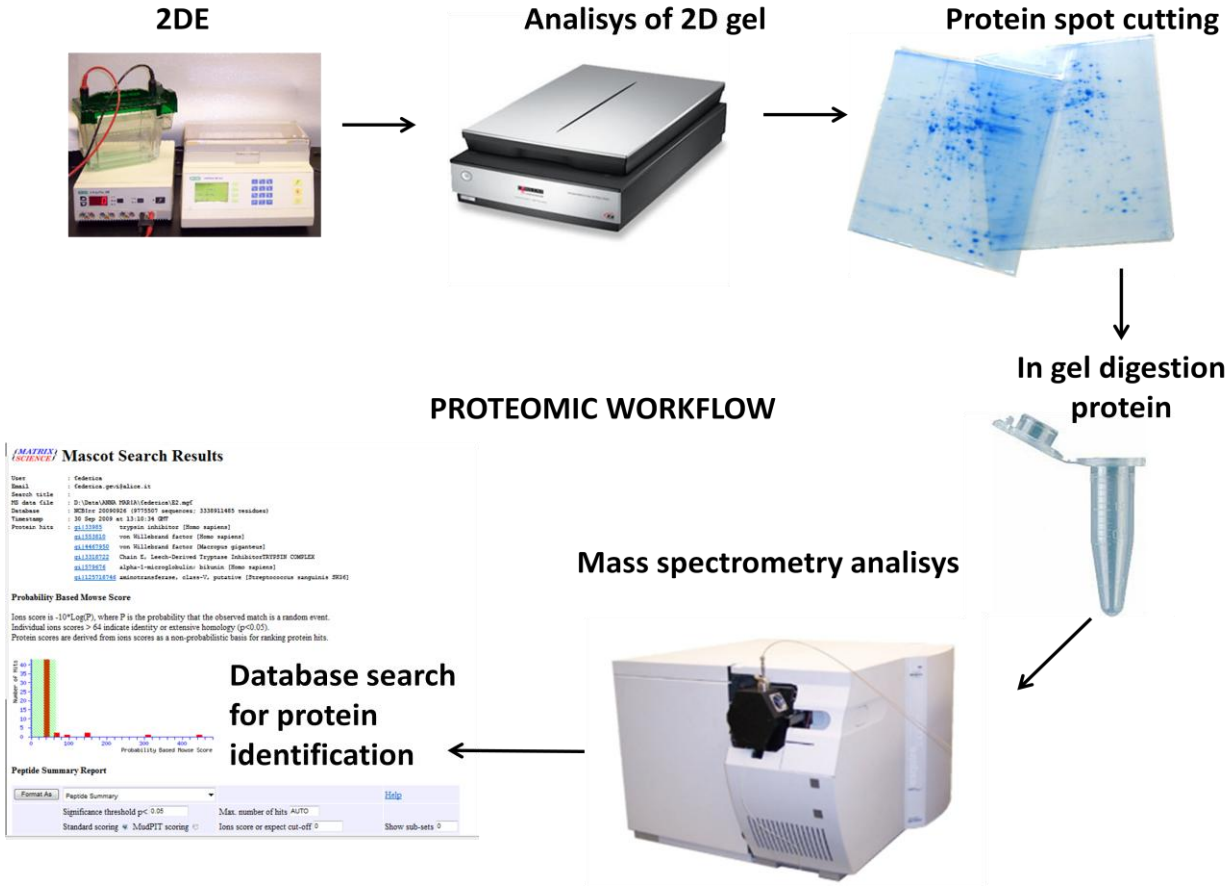
(d) for the study of protein–protein interactions.

Considering the complexity of the proteome, improved methods of analysis are of critical importance to realize the goals of proteomics. As already mentioned, the methodologies employed in proteomics depending on the objectives of the experiment. The general analysis scheme used to study the proteome is referred to as the proteomics workflow. In my study I use this kind of workflow (Fig 1).

1.2 Different approaches in proteomics

There are two main approaches to protein identification through mass spectrometry: the bottom-up and top-down approach. In bottom-up mass spectrometry, the proteins in the mixture undergo proteolytic digestion before the introduction into the mass spectrometer. Proteolytic digestion is carried out by proteolytic enzymes, such as trypsin, and consists in the cleavage of the whole protein at specific sites known in advance (specificity of the enzyme). Then the sample is ionized, sometimes after a separation stage, and finally analyzed by the spectrometer. This is what characterizes bottom-up proteomics: that the analytes are cleaved peptides. There are two commonly used methodologies: peptide mass fingerprinting and

tandem mass spectrometry (MS/MS or MS2). In peptide mass fingerprinting, peptide masses obtained from the MS scan are compared to calculated peptide masses.



.

Figure 1 : Workflow in proteomics: a gel-based approach using protein level separation followed by visualization and comparison of test and control samples, with in-gel digestion and LC-MS/MS analysis of the extracted peptides.

These masses are obtained by an in silico cleavage of proteins or gene sequences in a database using the appropriate specificity. In tandem mass spectrometry, a peptide ion is isolated in the mass analyzer and subjected to dissociation to produce product ion fragments. These fragments are then analyzed by a second mass spectrometric measurement. With the product ion spectra, different methods can be applied: a de novo analysis of the whole protein, a de novo reconstruction of small peptide sequences that serve as tags for a database search, or a database search based on a cross-correlation analysis. In top-down proteomics, intact protein molecular ions generated by electrospray ionization are introduced into the mass analyzer and are subjected to gas-phase fragmentation for MS analysis. Therefore its difference from the

bottom-up approach resides in that charged fragments of the whole protein are analyzed instead of proteolytic digests, and they are usually much larger (104 Daltons is a typical parent mass for a top-down spectrum). As in the bottom-up approach, different methods can be applied to attempt to identify proteins from spectra.

1.3 Separative techniques

The complex nature of the proteome demands the use of different analytical technologies to obtain the global picture of the cell. In particular the analysis of proteins, whether on a small or large scale, requires methods for the separation of protein mixtures into their individual components. The two groups of techniques that have come to dominate proteomics are two-dimensional gel electrophoresis (2DE) and multidimensional liquid chromatography (LC).

1.3.1 Two-dimensional gel electrophoresis

Electrophoresis is defined as the movement of charged molecules under an electrical field towards the opposite charged electrode. Due to their varying charges and masses, different molecules move with different velocities and became separated into single fractions (Westermeier 2006). Two-dimensional electrophoresis (2DE) was developed two decades before the term proteomics was coined (Klose 1975; O'Farrell 1975). The 2DE entails the separation of complex protein mixtures by molecular charge in the first dimension and by mass in the second dimension. 2DE analysis provides several types of information about the hundreds of proteins investigated simultaneously, including molecular weight, pI and quantity, as well as possible posttranslational modifications. The underlying principle is that electrophoresis is carried out in a pH gradient, allowing each protein to migrate to its isoelectric point, i.e., the point at which its pI value is equivalent to the surrounding pH and its net charge is zero. Proteins with different pI values therefore focus at different positions in the pH gradient. The pH gradient in an IEF gel can be established to use synthetic carrier ampholytes, which are collections of small amphoteric molecules with pI values corresponding to a given pH range. The second dimension separation in 2DE is generally carried out by standard sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), and separates the proteins according to molecular mass irrespective of charge. The basis of the technique is the exposure of denatured proteins to the detergent SDS, which binds

stoichiometrically to the polypeptide backbone and carries a large negative charge. The presence of tens or hundreds of SDS molecules on each polypeptide dwarfs any intrinsic charge carried by the proteins themselves, and stoichiometric binding means that larger proteins bind more SDS than smaller proteins.

1.3.2 Liquid chromatography separation

LC has several advantages over 2DE as a separation method, including its versatility, sensitivity, and the ease with which it can be automated and integrated with downstream analysis by MS. Unlike gel electrophoresis, LC is suitable for the separation of both proteins and peptides, and can therefore be applied either upstream of 2DE to prefractionate the sample, downstream of 2DE to separate the peptide mixtures from single excised spots, or instead of 2DE as the separation technology of choice. LC methods can exploit different separation principles, such as size, charge, hydrophobicity, and affinity for particular ligands. As is the case for electrophoresis, the highest resolution separations are achieved when two or more separation principles are applied one after the other in orthogonal dimensions. In LC methods used in proteomics, the stationary phase is a porous matrix, usually in the form of packed beads that are supported on some form of column. The mobile phase, a solvent containing dissolved proteins or peptides, flows through the column under gravity or is forced through under high pressure Indeed, the individual fractions produced by such methods are usually ‘anonymous’. Each spot on a 2D gel and each fraction emerging from an HPLC column look very much like any other. In the case of 2DE, even differences in spot size and distribution provide only vague clues about protein identity. The next stage in proteomic analysis is therefore to characterize the fractions and thus determine which proteins are actually present. Currently proteomics may rely on many chromatographic and electrophoresis tools to fractionate the analytes. However, if different approaches are in relation to these techniques of separation, all the strategies have a common essential final step: the mass spectrometry (MS) analysis of peptides or proteins.

1.4 Protein digestion strategies

Protein digestion procedures are used to produce a specific pool of peptides from single proteins, or from complex mixtures of proteins, which is then analyzed by MS. Enzymatic digestion of proteins with proteases produces a reproducible pool of peptides suitable for MS analysis. This feature makes this the preferred and most popular method to obtain the peptide fragments for protein identification. Generally, only one enzyme is used for protein digestion, but it has also been reported the use of multiple proteases to improve protein identification and characterization (Swaney 2010). Trypsin is the most used protease in MS-based proteomics studies, due to several factors: (i) high specificity, since it hydrolyses the peptide bond at the C-terminus of lysine and arginine residues, except when a proline follows in the sequence (Rodriguez 2008) (ii) the peptides produced have basic residues at the C-terminus (arginine and lysine) which make them easily ionisable in the mass spectrometer ; and (iii) the mass range of the obtained peptides, between 800 and 4000 Da, provides excellent MS and MS/MS data. In general, protein enzymatic digestion can be performed by three different approaches: (i) in-gel digestion; (ii) in-solution digestion; and digestion with immobilized enzymes. Each one of these approaches, as well as other new methodologies for protein digestion, will be addressed in the next sections.

1.4.1 Identification through mass spectrometry

The recent advances in proteomics were mainly driven by the increasing ability of MS to detect and characterize low levels of proteins. Mass spectrometers consist of three basic components: an ion source, a mass analyzer, and an ion detector (Mann 2001)

1.4.2 Ion source

Electrospray Ionization (ESI) and Matrix-Assisted Laser Desorption/Ionization (MALDI) are the two most commonly employed ion sources for proteins and peptides. ESI ionizes the analytes out of an aqueous solution and is therefore the ideal method for LC-MS setups. The liquid containing the analytes is pushed through a small capillary to which a potential difference is applied (Fenn 1989). This causes a strong electric eld which in turn causes an accumulation of charged molecules at the surface of the liquid. The solvent molecules, which

are usually more volatile than the analyte, evaporate and force the analyte molecules into closer vicinity. The molecules repel each other and break up the droplets. The process repeats until all solvent molecules are removed and the analytes form lone ions. An ESI source can be operated in positive or negative ion mode. In positive ion mode, the compounds receive a proton as charged adduct, in negative mode they receive an electron.

1.4.3 Mass Analyzer

The mass analyzer measures the mass-to-charge ratio (m/z) of the ionized analyte. For our purposes, its key parameters are sensitivity, mass resolution and mass accuracy (Aebersold 2003). The sensitivity characterizes the ability of the mass analyzer to detect weak signals. Mass resolution and mass accuracy describe how well the analyzer is able to resolve peaks with similar mass and how accurately it measures this mass, respectively.
In proteomics research, four basic kinds of mass analyzers are currently being used:

- time-of-flight (TOF),
- ion trap,
- quadrupole
- fourier transform ion cyclotron resonance (FTICR) analysers.

All four differ considerably in sensitivity, resolution, mass accuracy and the possibility to fragment peptide ions.

1.4.4 Mass detector

Finally, the mass detector counts the number of ions at each m/z. Possible detectors are photographic plates, faraday cylinders or array detectors. Most detectors need some time to recover after an ion hit. The combination of ion source, mass analyser and detector is usually determined by the application. ESI is most frequently coupled to ion traps (three-dimensional and linear ion traps) and hybrid tandem mass spectrometers like quadrupole time-of-flight (Q/TOF) instruments. In the case of ion traps, the ions are first captured in the centre of the device for a certain time interval and are then scanned from the trap to the detector (Yates 2004).With this type of mass analyser it is not only possible to determine the mass of a given peptide, but also its sequence. Ions with specific m/z ratios can be selected in the ‘trap’ for fragmentation, induced by collision of the ion with an inert gas or a surface in a process called

collision induced fragmentation (CID). This energy causes the peptide ion to fragment at different points, commonly at the peptide bond. The recorded product ions represent the tandem mass spectrum (MS/MS, in particular the application of ESI coupled on-line with high-performance separation techniques such as HPLC, has had a dramatic effect on the sensitivity and the speed with which the primary structure of proteins and peptides can be determined (Aebersold 2001) which contains information on the amino acid sequence. MS-based proteomics has a rowing role in biomedical research where limited sample material is available, as femtomole sensitivity is routinely achieved.

1.5 Proteomic's application

Since the definition of the concept 15 years ago, proteomics has become a very important discipline among the scientific community with multiple applications. Unlike the classical protein biochemistry science, which studies individual proteins emphasizing on structural, function and complete sequence analysis, proteomics investigates complex biological systems to understand the relation between different proteins and their distinct functions within large networks. By studying these complex systems at the proteome level, scientists can obtain better knowledge about biological functions, understand the changes in cellular regulation mechanisms caused by disease states, identify disease biomarkers and develop new drugs or therapeutic approaches. In recent years, proteomic technologies have led to enormous advances in basic research and medicine. On the clinical side, proteomics is viewed as a promising new approach that will speed the discovery and validation of protein biomarkers that correlate with disease and allow for assessment of therapeutic regimens. Proteomics research is applicable to just about every area of biochemical investigation including health, agriculture, fisheries and forestry. In my thesis I used proteomic analysis to characterized pharmaceutical products and to identified possible contaminations that could cause problems in Haemophiliac patient. In the first study, we have focused our investigation on rFVIII. Recombinant FVIII were first introduced in the early 1990s, as biosynthesized blood coagulation factor prepared using recombinant DNA technology (Wood 1994; Eaton 1987).

1.6 References :

Aebersold R., Goodlett D.R. (2001) Mass spectrometry in proteomics. Chem. Rev. 101: 269-295.

Aebersold R., Mann M. (2003) Mass spectrometry-based proteomics. Nature 422:198-207.

Eaton D.L., Hass P.E., Riddle L., Mather J., Wiebe M., Gregory T., Vehar G. A. (1987) Characterization of factor VIII. Biol. Chem. 262: 3285–3290.

Fenn J. B., Mann M., Meng C. K., Wong S. F., Whitehouse C. M. (1998) Electrospray ionization for mass spectrometry of large biomolecules. Science 246: 64–71.

Yates J. R. Mass spectral analysis in proteomics (2004) Annu. Rev. Biophys. Biomol. Struct. 33: 297–316.

Klose J. (1975) Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. Human Genetics 26:231–243.

Mann M., Hendrickson R. C., Pandey A., (2001) Analysis of proteins and proteomes by mass spectrometry. Annual Review of Biochemistry 70: 437–473.

O'Farrell P. H. (1975) High resolution two dimensional electrophoresis of proteins. Journal of Biological Chemistry 250:4007–4021.

Rodriguez J., Gupta N., Smith R.D., Pevzner P.A. (2008) Does trypsin cut before proline? J. Proteome Res 7: 300-305.

Swaney D.L., Wenger C.D., Coon J.J. (2010) Value of using multiple proteases for large-scale mass spectrometry-based proteomics. J. Proteome Res. 9:1323-1329.

Traini M., Gooley A.A., Ou K., Wilkins M.R., Tonella L., Sanchez J.C., Hochstrasser D.F., Williams K.L, (1998) Towards an automated approach for protein identification in proteome projects. Electrophoresis, 19:1941-1949.

Wasinger V.C., Cordwell S.J., Cerpa-Poljak A., Yan J.X., Gooley A.A., Wilkins M.R., Duncan M.W., Harris R., Williams K.L., Humphery-Smith I . (1995) Progress with gene-product mapping of the mollicutes: Mycoplasma genitalium. Electrophoresis 16:1090-1094.

Westermeier R. (2006) Preparation of plant samples for 2-d electrophoresis. Proteomics 2006 2:56-60.

Wilkins M.R., Williams K.L., Appel R.D., Hochstrasser D.F. (1997) Proteome Research: New Frontiers in Functional Genomics (Principles and Practice). Berlin: Springer 2:35-64.

Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J., Keyt B., Seeburg P.H., Smith D. H. (1994) Expression of active human factor VIII from recombinant DNA clones. Nature312:330-337.

Chapter 2

Recombinant clotting factor VIII concentrates: Heterogeneity and high-purity evaluation.

2.1.1 Introduction: Haemophilia A

Hemophilia is a bleeding disorder caused by a deficiency in one of two blood clotting factors: factor VIII (Haemophilia A) or factor IX (Haemophilia B). Hemophilia affects males much more frequently (1 in 10.000) than females (1 in 100.000.000). This occurs because a critical blood clotting gene is carried on the X chromosome. Since males only carry one X chromosome, if that is defective, hemophilia will immediately show up. Females, on the other hand, carry two X chromosomes. If only one is defective, the other normal X chromosome can compensate. The woman will have normal blood clotting; she will simply be a carrier of the recessive defective gene. This fact will be discovered if some of her children are hemophiliacs. Naturally, women hemophiliacs are rare because it takes two defective X chromosomes in order for the condition to be seen. Hemophilia has played an important role in Europe's history, for it suddenly cropped up in the children of Great Britain's Queen Victoria. It became known as the "Royal disease" because it spread to the royal families of Europe through Victoria's descendants. Queen Victoria had always been worried about the quality of the blood of the British royal family.

2.1.2 History of Treatment in haemophiliac patient

In 1960 the average life expectancy for severe hemophilia A was less than 20 years, and the quality of life was generally devastating from joint bleeding complications or intracranial hemorrhage (Ikkalat 1982; Triemstra 1995). In the late 1950s and much of the 1960s, fresh frozen plasma (FFP) was the mainstay of treatment for hemophilia A and hemophilia B. Then in the early 1980s, the epidemic of HIV devastated the hemophilia community, and by 1985 nearly 75% of the severe hemophilia A patients in the US had acquired HIV. The ironic tragedy was that the source of the HIV infection was the plasma-derived factor VIII products that had promised to be life-saving. As the community dealt with the horrors of HIV, a second epidemic gradually became apparent. Over 95% of severe hemophilia A patients had also acquired hepatitis C infection through plasma-derived preparations of factor VIII (Goedert 2007). So in 1980s the need for safer haemophilia treatments became crucial to the

haemophilia community. The successful cloning of the factor VIII gene in 1984 was a major breakthrough, allowing production of recombinant human factor VIII (rfactor VIII). Recombinant FVIII, first introduced in the early 1990s, significantly reduced the risk for pathogen transmission. The FVIII gene is one of the largest proteins to be expressed for therapeutic purposes (Jiang 2002). The gene has been expressed in full length or with the B-domain deleted (Refacto®). Once the gene is selected for transfection it is spliced into a vector during the first part of recombinant product processing. The vector is then inserted into Chinese hamster ovary (CHO) or baby hamster kidney (BHK) cells. (Eriksson 2007; Pipe 1998). An inoculum, meeting strict criteria of viability, sterility, and growth rate, is then fermented in a bioreactor using a batch-refeed process. In culture, the cells then secrete rFVIII into the medium, which is harvested. The rFVIII-rich media undergoes a stepwise purification process (usually unique to the manufacturer), but generally includes multiple chromatography columns. The purification processes are designed to isolate rFVIII and remove or inactivate any remaining contaminants. The bulk processing is designed to ensure consistency in the quantitative and qualitative composition of all preparations. There are a number of different recombinant products in use currently, produced by various pharmaceutical companies. This products referred to as 'first, second or third generation' depending on the way they are manufactured are present.

- The 'first generation' of recombinant products use animal products in the culture medium and contain human albumin (a human blood product) added as a stabilizer. 'Recombinate' was the first generation recombinant factor VIII product manufactured by Baxter.
- 'Second generation' products use animal-derived materials in the culture medium but do not have added albumin and instead use sucrose or other non-human derived material as a stabiliser. Second generation Factor VIII products are 'Kogenate FS' made by Bayer (also labelled as 'Helixate Nexgen' and distributed by ZLB), 'Refacto®' manufactured by Wyeth.
- 'Third generation' products have no human or animal protein used in their preparation. The new third generation factor VIII products is. 'Advate®', from Baxter.

Helixate NexGens® are produced introducing human FVIII genes into baby hamster kidney (BHK21) cell line while Advate® and Refacto® derived from Chinese hamster ovary (CHO) cells (Gomperts 1992; Jiang 2002; Eriksson 2001; Pipe1998).

The FVIII human gene has been expressed in full length (FLrFVIII), such as Helixate NexGens ® and Advates®, or with the B-domain deleted (BDDrFVIII), like ReFacto®. To this

regard, it is worthwhile to remark that initially, the B-domain was believed to have an insignificant function; however, in vitro studies have shown that B-domain interacts with chaperone proteins that have a processing role with ultra-cellular structures, although further investigations are mandatory (Pipe 1998). This might have an impact on the production production of rFVIII, although this should not influence its function in vivo. Indeed, while its biological role is not yet fully understood, B-domain removal has been found to dramatically improve the yield of rFVIII (Dorner 1987). Genetically engineered cell lines secrete rFVIII into a defined cell culture medium. Different methods are used to express, isolate, harvest and purify the rFVIII preparation. In all cases, a multistep purification process, including several chromatographic steps (like monoclonal antibody immunoaffinity chromatography, anion exchange chromatography, gel filtration), is needed for the isolation and the purification of rFVIII from the raw product of fermentation. The purification processes are designed to isolate rFVIII and remove or inactivate any remaining contaminants. The bulk processing is designed to ensure consistency in the quantitative and qualitative composition of all preparations. Single-dose vials are then prepared aseptically, and the liquid is lyophilized to produce a powder, which can be reconstituted with diluents prior to administration. Quality controls of the end product aim at establishing product stability and the absence of contaminants, mainly through SEC-HPLC and 1-D SDSPAGE, but not 2-DE (Kim 2007). So for all this ragion CNS decided to finance a project which premise the characterization of products general in use. In this study, we exploited the 2-DE approach to assess and compare the heterogeneity and the purity grade of commercially rFVIII Helixate NexGens®, Refacto® and Advate® preparation used in the treatment of hemophilia A before and after the thrombin cleavage. In literature are just present this kind of study in fact, Clifton (2009) performed a proteomic characterization of plasma derived FVIII in order to identify eventual impurities. The authors concluded that proteomic investigations are recommended in addition to routine biochemical and functional analyses, as to improve monitoring of production processes. Our data confirm that a proteomic approach can improve the quality assurance process also for different rFVIII concentrates.

2.1.3 Factor VIII's structure

Human clotting factor VIII (FVIII) is a non-enzymatic plasma glycoprotein that plays an essential role in the intrinsic pathway of the blood coagulation cascade (Vehar,1984). Native protein is synthesized as a single polypeptide chain of 2332 amino acids, which is preceded by a 19-residue hydrophobic signal peptide of importance for the secretion. The protein contains six main domains: three A type domains, a central B domain and two C-type domains. In detail, domains can be distinguished as follows: three small acidic peptide regions (30–40 designed by ''a'') join (i) the A1 and A2 domains (residues 337–372); (ii) the A2 and B domains (residues 711–740); and (iii) the last one connects the B domain with the amino terminal of the A3 domain (residues 1649–1689). Therefore, the domain organization order of FVIII is as follows: (NH2)-A1-a1-A2- a2-B-a3-A3-C1-C2 (COOH). (Fig 2) The B domain, which represents 40% of the mass of FVIII, has no recognized function, as it does not appear to contribute to procoagulant activity (Andersson 1986).In plasma, FVIII circulates in an inactive pro-cofactor form in complex with von Willebrand factor, a 2050 amino acid glycoprotein. Interaction with von Willebrand factor stabilizes the FVIII molecule and facilitates the cleavage and activation of FVIII (Fulcher 1983). Under physiological conditions, FVIII is processed to a series of divalent metal ion-linked heterodimers by cleavage at the B-A3 junction: two fragments are generated (i) a variable heavy chain (90–210 kDa) consisting of the A1-A2 and (ii) a heterogeneous fragment of partially proteolyzed B domains, which is associated with a light chain (80 kDa) consisting of the A3-C1-C2 domains (Fay 1986; Fay 2004). Conversion of profactor FVIII to active FVIIIa by thrombin (or factor Xa) is associated with specific proteolytic cleavages in both the heavy and the light chains (Lenting 1998). When factor VIIl is activated by thrombin, a 41-amino-acid peptide is liberated from the N-terminus of the light chain by cleavage at Arg l689 and removal of the B domain is completed by cleavage al Arg740 followed by cleavage of the heavy chain between the A1 and A2 domains at Arg372 (Eaton 1986; Fay 1986). This results in the formation of the active factor VIIIa molecule comprising in Al/light chaindimer associated with the A2 domain (Fay 1991).

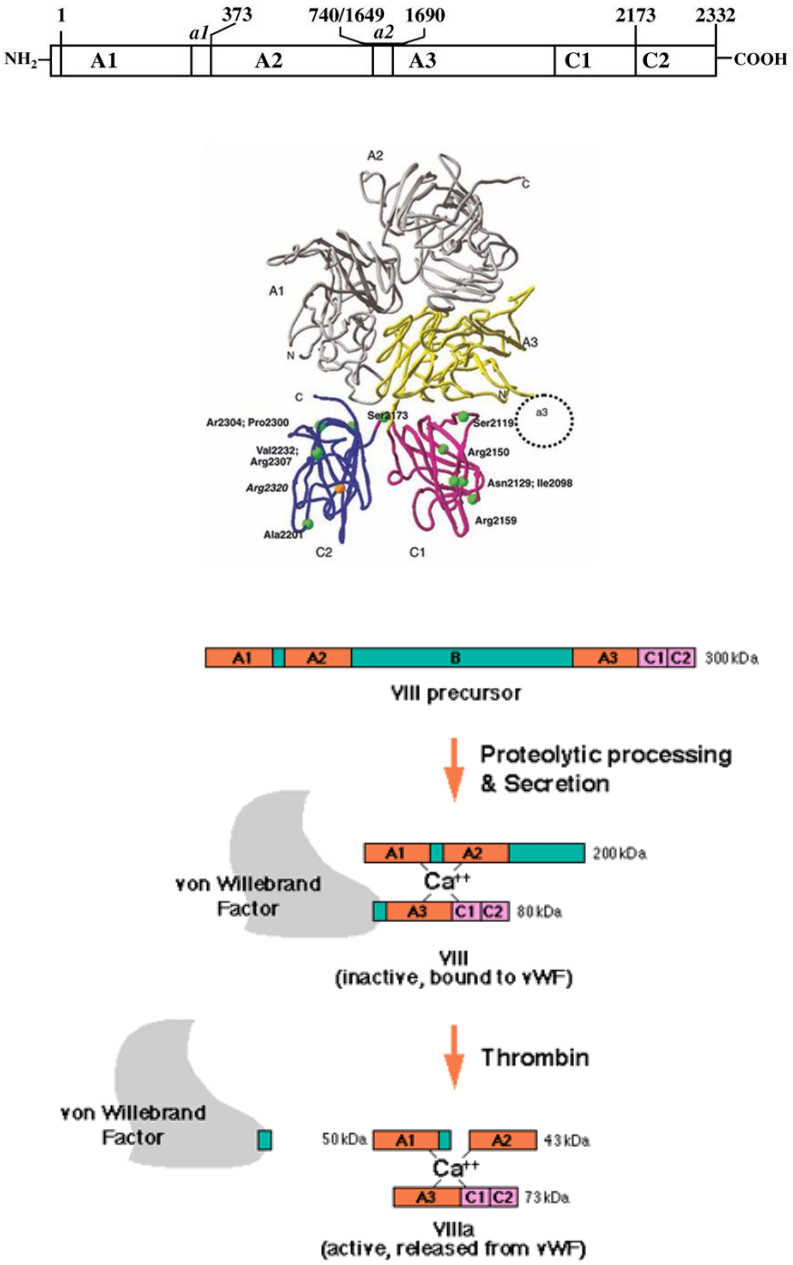


Figure 2: A) The protein structure of factor VIII. There are 2332 residues in factor VIII, composed of 3 distinct domain types in the arrangement (NH2) A1—A2—B—A3—C1—C2 (COOH). B) Activation of Factor VIII by Thrombin.

2.2 Materials and methods

2.2.1 rFVIII

We have used two batches of three approved rFVIII preparations obtained from different manufacturers: Helixate NexGen® 1000 UI (CSL Behring, Philadelphia, USA), Refacto® 250 UI (Wyeth Europa, Maidenhead, UK) and Advate® 250 UI (Baxter, Vienna, Austria). We reconstituted the lyophilized rFVIII preparations according to the manufacturers' instructions and then we stored at -80°C in 1-mL aliquots. We quantified total protein trough a Bradford assay (Bio-Rad Laboratories, Hercules, CA, USA) using bovine serum albumin (Sigma Aldrich, St. Louis, MO, USA) as standard.

2.2.2 Digestion of rFVIII with thrombin

To cleavage rFVIII with thrombin we prepare a solution with rFVIII preparations (2 mg) and 4 mg thrombin (Sigma Aldrich) in 20 mL of a cleavage buffer containing 50mM Tris (pH 8.0), 150mM NaCl, 2.5mM CaCl and 0.1% 2-mercaptoethanol. Reactions were run at room temperature for 20 min.

2.2.3 SDS-PAGE

SDS-PAGE was performed according to the method of Laemmli (1970). The rFVIII preparations (30 mg), before and after thrombin digestion, were loaded onto a 0.75-mm-thick 5–16% w/v acrylamide gradient gel. The 1-D gel was stained with Blue Silver (Shevchenko,1996). The apparent molecular weight of bands was determined using a wide molecular weight range calibration kit for SDS-PAGE (Sigma Aldrich).

2.2.4 2-DE

Protocol A: Four hundred micrograms of reconstituted sample was precipitated in 80% v/v acetone. After incubation at 41°C for 90 min, the precipitate was pelleted by centrifugation at 12 000 g for 20 min at 41°C. After washing with the same solution, the pellet was air-dried and then solubilized in the focusing solution containing 7M urea, 2M thiourea, 4% w/v

CHAPS, 0.5% w/v pH 4–7 carrier ampholyte (Bio-Rad Laboratories), Tris 40mM (pH 8.8), 5mM TBP and 10mM acrylamide. Before focusing, the sample was incubated in this solution for 3 h at room temperature, under agitation. To prevent over-alkylation excess, acrylamide was destroyed by adding an equimolar amount of DTE. Three hundred microliter of the resulting protein solution was then used to rehydrate 17 cm-long IPG 4–7 (Bio-Rad Laboratories) for 8 h. IEF was carried out on a Protean IEF cell (Bio-Rad Laboratories) with a maximum current setting of 50 mA/strip at 20°C. The total product time voltage applied was 80 000 Vh for each strip. For the second dimension, the IPG strips were equilibrated for 30 min in a solution containing 6M urea, 1% w/v SDS, 20% v/v glycerol and 375mM Tris-HCl (pH 8.8), with gentle agitation. The IPG strips were then laid on an 11% T SDS PAGE gel and proteins were visualized by Blue Silver method (Laemmli 1970).Protocol B: After precipitation in 80% v/v acetone, the air-dried pellet was solubilized in 100 mL of 7M urea, 2M thiourea, 4% w/v CHAPS, 0.5% w/v pH 4–7 carrier ampholyte, Tris 40mM (pH 8.8), 5mM TBP and 10mM acrylamide. Dried 17 cm-long IPG 4–7 strips were rehydrated overnight at room temperature with 800 mL of 7Murea, 2M thiourea, 2% w/v CHAPS, 0.5% w/v pH 4–7 carrier ampholyte (pH 6.3). The sample solution (100 mL) was loading into rehydrate IPG strip using the cup-loading method, and IEF was carried out on a Protean IEF cell with a maximum current setting of 50 mA/strip at 201C. After IEF, IPG strips were equilibrated for 30 min in a solution containing 6M urea, 1% w/v SDS, 20% v/v glycerol and 375mM Tris-HCl (pH 8.8), and laid on an 11% T SDS PAGE gel. Proteins were visualized by Blue Silver method. Protocol C: To remove the salt contaminations, the reconstituted sample (400 mg) was precipitated in 80% v/v acetone. The obtained air-dried pellet was than solubilized in 100 mL of 8M urea, 4% w/v CHAPS, 0.5 % w/v pH 4–7 carrier ampholyte, Tris 40mM (pH 8.8), 5mM TBP and 10mM acrylamide. Dried 17 cm-long IPG 4–7 strips were rehydrated with 800 mL of 8M urea, 2M thiourea, 2% w/v CHAPS and 0.5% w/v pH 4–7 carrier ampholyte (pH 6.3). Cup-loading method was used to load the sample solution (100 mL) into rehydrate IPG strip. After the first dimension step, IPG strips were equilibrated for 30 min in a solution containing 6M urea, 3% w/v SDS, 20% .

2.2.5 In gel digestion

Protein spots were carefully excised from blue-silver-stained gels and subjected to in-gel trypsin digestion according to Shevchenko with minor modifications. The gel pieces were swollen in a digestion buffer containing 50mM NH4HCO3 and 12.5 ng/mL of trypsin (modified porcine trypsin, sequencing grade, Promega, Madison, USA) in an ice bath. After 30 min, the supernatant was removed and discarded, 20 mL of 50mM NH4HCO3 was added to the gel pieces and digestion allowed to proceed at 37°C overnight. The supernatant containing tryptic peptides was dried by vacuum centrifugation. Prior to mass spectrometric analysis, the peptide mixtures were redissolved in 10 mL of 5% formic acid (FA).

2.2.6 Protein identification by MS/MS

Peptide mixtures were separated using a nanoflow-HPLC system (Ultimate; Switchos; Famos; LC Packings, Amsterdam, The Netherlands). A sample volume of 10 mL was loaded by the autosampler onto a homemade 2-cm fused silica precolumn (75 mm id; 375 mm od; Reprosil C18- AQ, 3 mm, Dr. Maisch, Ammerbuch-Entringen, Germany) at a flow rate of 2 mL/min. Sequential elution of peptides was accomplished using a flow rate of 200 nL/min and a linear gradient from Solution A (2% ACN; 0.1% FA) to 50% of Solution B (98% ACN; 0.1% FA) in 40 min over the precolumn in-line with a homemade 10 to 15 cm resolving column (75 mm id; 375 mm od; Reprosil C18-AQ, 3 mm, Dr. Maisch GmbH, Ammerbuch-Entringen, Germany). Peptides were eluted directly into a High Capacity ion Trap (model HCTplus, Bruker-Daltonik, Germany). Capillary voltage was 1.5–2 kV and a dry gas flow rate of 10 L/min was used with a temperature of 230°C. The scan range used was from 300 to 1800 m/z. Protein identification was performed by searching the National Center for Biotechnology Information non-redundant database (NCBInr, version 20081128, www.ncbi.nlm.nih.gov) using the Mascot program (in-house version 2.2, Matrix Science London, UK). The following parameters were adopted for database searches: complete carbamido methylation of cysteines and partial oxidation of methionines, peptide mass Tolerance 71.2 Da, fragment mass tolerance 70.9 Da, missed cleavages 2. For positive identification, the score of the result of (10LogP) had to be over the significance threshold level (po05). Although high Mascot scores are obtained with values greater than 60, when proteins were identified with only one peptide, a combination of automated database search and manual interpretation of peptide

fragmentation spectra was used to validate protein assignments. In this manual verification, the mass error, the presence of fragment ion series and the expected prevalence of C-terminus containing ions (Y-type) in the high mass range were all taken into account. Moreover, replicate measurements have confirmed the identity of these protein hits Sequencing of A2 cleavage fragments Thrombin digested rFVIII preparations (30 mg) were run on 5–16% w/v acrylamide gradient gel at 50 mA and transferred for 2 h at 100 mA on a Hybond-P PVDF membrane in 10mM CAPS, 10% ethanol (pH 11.0). After staining with Coomassie blue, visible bands (molecular mass between 36.0 and 45.0 kDa) were cut and subjected to N-terminal sequencing for ten cycles, using an automatic protein microsequencer Prosize 492 cLC (Applied Biosystems). Amounts of protein sequenced 0.5 pM were considered significant and above the background level associated with the experimental conditions. All sequences were identified as thrombin cleavage fragments of rFVIII.

2.3 Results

To evaluate the structural heterogeneity, the presence of contaminants in rFVIII samples and the effect of thrombin cleavage, SDS-PAGE was carried out on both intact (Fig. 3) and thrombin-digested (Fig. 4) preparations from three commercial products: Helixate NexGen®, Refacto® and Advate®. 1-D gel of Helixate NexGen® and Advate® (Fig. 3A) showed multiple bands in the region between 180 and 110 kDa (bands 2, 3, 4 and 11, 12). In the case of Refacto®, only one band was observed at 170 kDa (band 7), corresponding, as expected, to the entire protein expressed without B domain and plus the two main bands at 90 and 80 kDa (bands 8, 9). To identify these bands, gel bands were subjected to an in-gel tryptic digestion and the proteins were identified by MS (Table 1) Figure 3B reports in bold the peptides identified by MS/MS in each band; this determined that all the bands observed in Helixate NexGen® and Advate® between 180 and 110 kDa are truncated forms of the full-length protein, where most cleavage occurs in the B domain. The intensity and the number of these bands were predominant in the second generation of FLrFVIII material from BHK cells (Helixate NexGen®) rather than in third generation FLrFVIII material from CHO cells (Advate®). Clearly, these truncated forms are absent in Refacto®, in which the expressed gene is B-domain deleted. Otherwise, FLrFVIII preparations Helixate® and Advate®, showed a similar banding pattern, with a major species at 200 kDa corresponding to the full length protein (A1-A2-B-A3-C1-C2) containing the domain B, (respectively bands 1 and 10), and

two chains: heavy at 90 kDa (A1-A2, bands 5 and 13) and light at 80 kDa (A3-C1-C2, bands 6 and 14). In Fig. 3A, two bands are recognizable as belonging to A3-C1-C2 domains in all rFVIII (bands 6, 9, 14) that differ in molecular weight of 2.5 kDa. In this case, we hypothesized that the observed bands may be produced upon differential cleavage of the protein or rather result from different glycosylation moieties. The last hypothesis is perhaps more reliable, as it can be supported with observations from literature, such as Shen et al. who identified sites of glycosylation and metal binding in the A3-C1-C2 domains (the ‘‘light chain’’) studying the tertiary structure and domain organization of rFVIII. Although SDS-PAGE patterns of the rFVIII apparently differed, they were rather comparable upon thrombin digestion, which yielded similar products (Fig. 4A). Figure 4B is a schematic representation of the fragmentation pattern generated upon thrombin digestion of precursors. In the case of Helixate NexGen® and Advate®, along with the cleavage products of thrombin, which were recorded in the region between 66 and 36 kDa, the isolated B domain (bands 15 and 22) was recovered almost intact at about 135 kDA (bands 15 and 22). Thus, the heterogeneity observed above, belonging to different forms of three rFVIII, is removed upon thrombin treatment and the same final products are obtained. By MS/MS analysis of bands excised from gel, it can be assessed that after thrombin activation the precursor bands at 200 (Helixate NexGen®, Advate®), 170 (Refacto®), 90 (heavy chain) and 80 kDa (light chain) were processed to 73 (bands 16, 19, 23), 50 (bands 17, 20, 24) and 43 kDa (bands 18, 21, 25) polypeptide fragments. Bands 16, 19 and 23 correspond to A3-C1-C2 domain. Bands 17, 20, 24 (A1) and bands 18, 21, 25 (A2) originated from the cleavage of bands 5, 8, 13 (A1-A2). Interestingly, in all rFVII preparations, bands corresponding to the A2 domain were split into two bands, which differed in apparent molecular mass of about 1.5 kDa. Figure 4C reports in bold the peptides recognized by MS in A2 bands. Although MS/MS recognized that the two unexpected bands both belonged to the A2 domain, by using N-terminal sequencing, we could state that the two forms of A2 protein were present that differed in molecular weight of 1254.4 Da, corresponding to a specific peptide sequence (LLSKNNAIEPR). Thrombin digestion of A3-C1-C2 (bands 16, 19, 23 in Fig.4A) confirmed that two forms of this domain are present, even if Mascot results indicated a list of identical peptides for all of them and we were not able to apply Edman sequence, due to intrinsic technical limitation relative to high molecular weight. However, from SDS-PAGE profiles, we could state that two forms of these domains are present in the final product, which are characterized by different molecular weights.

*-

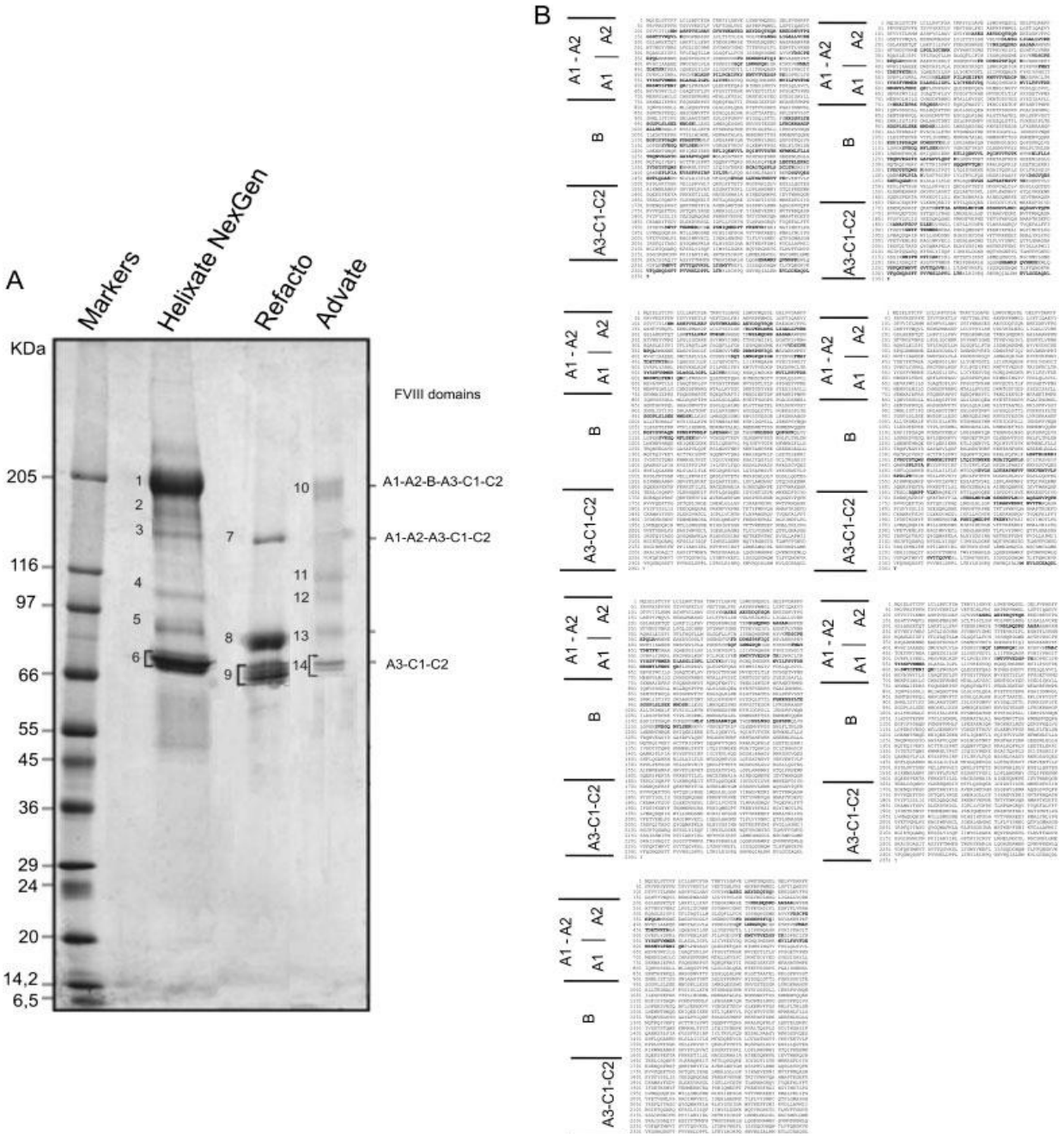


Figure 3:(A) 1-D SDS-PAGE of three different rFVIII preparations. Forty microgram of each rFVIII was separated on 5–16% w/vacrylamide gradient SDS gel and visualized by staining with blue silver. (B) Multiple bands in Helixate NexGen® (2, 3, 4) and Advate ® (11, 12), which represent truncated forms of the full-length protein bands 1 and 10, respectively, obtained by Mascot search. The proteins are marked by band number as reported Table 1

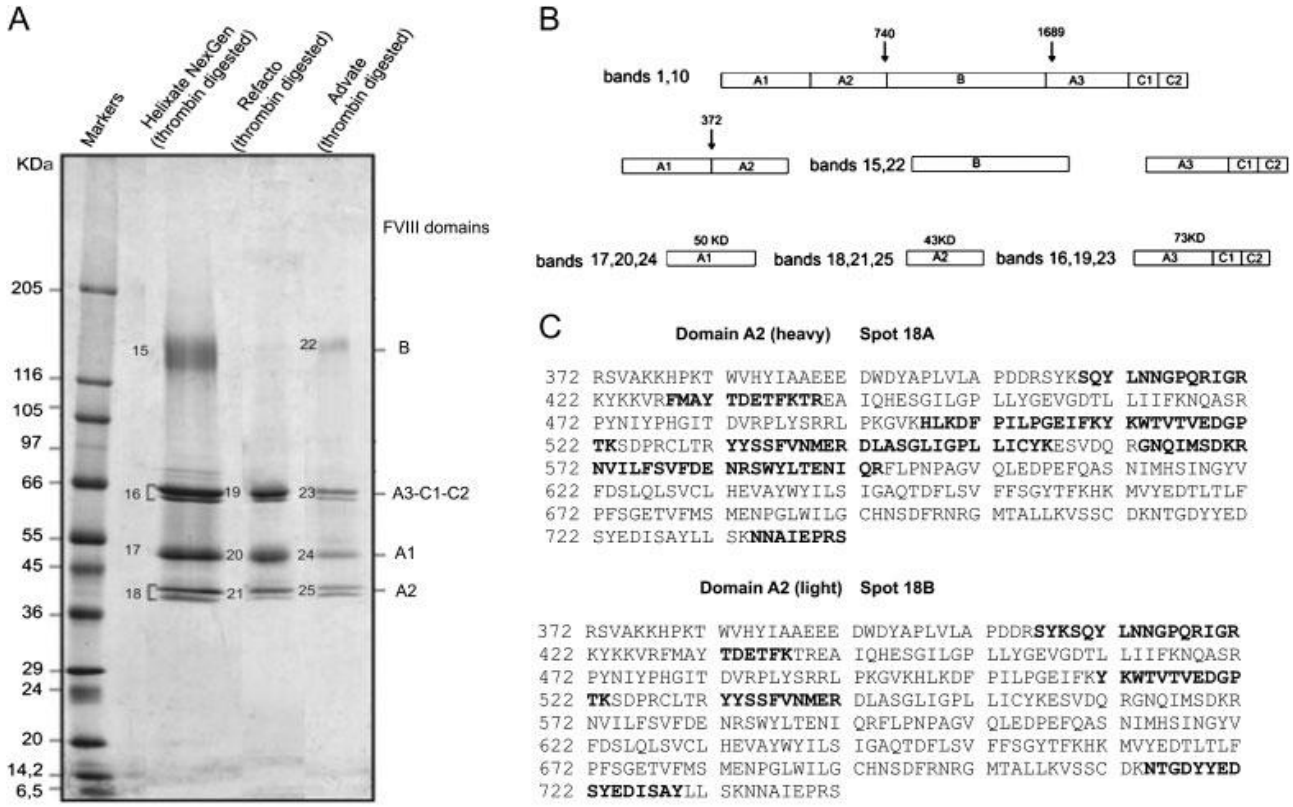


Figure 4:(A) 1-D SDS-PAGE of three different rFVIII preparations after thrombin activation. Forty micrograms of each rFVIII was separated on 5–16% w/v acrylamide gradient SDS gel and visualized by staining with blue silver. (B) A schematic representation of the fragmentation pattern generated upon thrombin digestion of precursors. (C) Peptides recognized by MS using Matrix Science (http://www.matrixscience.com) Mascot MS/MS ion search program in A2 bands. The proteins are marked by band number as reported in Table 3.

NEX GEN®						
Spot number	**Protein name**	**ID**	**Mascot score %**	**p/M r**	**Number mathced peptides**	**NCBI Accession number**
1	Coaugulation Factor VIIItutto	6.51	855	239233	19	gi|119593052
2	coagulation factor VIII [Homo sapiens]A1A2B	6.95	1427	268250	18	gi|182383
3	coagulation factor VIII, procoagulant component (hemophilia A), isoform CRA_b [Homo sapiens]A1A2	6.51	1075	239233	34	gi|119593052
4	coagulation factor VIII, procoagulant component (hemophilia A), isoform CRA_b [Homo sapiens]A1A2B	6.51	843	239233	21	gi|119593052
5	Chain A, Crystal Structure Of Human Factor VIII B	5.92	664	86866	17	gi|171849087
6	Chain A, Crystal Structure Of Human Factor VIII B	5.92	301	86866	7	gi|171849087
7	B domain-deleted coagulation factor VIII [synthetic construct]A1 A2 B	6.36	652	168754	20	gi|157863686
8	Chain B, Crystal Structure Of Human Factor VIII	7.46	1124	79418	39	gi|171849088
9	Chain B, Crystal Structure Of Human Factor VIII	7.46	1382	79418	57	gi|171849088

10	Chain B, Crystal Structure Analysis Of Coagulation Factor VIII(A3 c1 c2)	7.94	411	89376		gi|183448130
11	coagulation factor VIII, procoagulant component (hemophilia A), isoform CRA_b [Homo sapiens]A3	6.51	465	239233	13	gi|119593052
12	proapolipoprotein	5.45	95	28944	3	gi|178775
NEX-GEN® digested with thrombin						
Spot number	**Protein name**	**ID**	**Mascot score %**	**p/M r**	**Number mathced peptides**	**NCBI Accession number**
1	coagulation factor VIII [Homo sapiens dominio B	6.95	598	268250	20	gi|182383
2	coagulation factor VIII [Homo sapiens]		435	268250	14	gi|182383
3	coagulation factor VIII [Homo sapiens]B-A3	7.01	435	268250	14	gi|31499
3	heat shock 70 kDa protein 5 isoform	5.13	360	72493	7	gi|253982049
4	Chain B, Crystal Structure Analysis Of Coagulation Factor VIII A3 C1-C2	7.94	523	89376	20	gi|183448130
5	factor VIII [Homo sapiens]c2c1 A3	7.01	771	89376	27	gi|182803
6	coagulation factor VIII, procoagulant component (hemophilia A), isoform CRA_b [Homo sapiens]A1	6.95	1067	239233	41	gi|119593052
7	Chain A, Crystal Structure Of Human Factor VIIICatena A2	5.92	533	86866	41	gi|171849087
7	haptoglobin [Homo sapiens]	6.27	163	38941	5	gi|1212947
8	coagulation factor VIII [Homo sapiens] A2	7.01	325	268382	11	gi|31499
9	coagulation factor VIII, procoagulant component [Homo sapiens]A2	6.95	114	25056	4	gi|150036899
ADVATE®						
Spot number	**Protein name**	**ID**	**Mascot score %**	**p/M r**	**Number mathced peptides**	**NCBI Accession number**
1	Chain B, Crystal Structure Analysis Of Coagulation Factor VIII tutto	6.51	280	89376	8	gi|183448130
2	coagulation factor VIII [Homo sapiens]B-C1-C2 A3	6.97	724	268337	16	gi|119593053
3	Chain A, Crystal Structure Of Human Factor VIII B	5.92	269	86866	7	gi|171849087
4	B domain-deleted coagulation factor A1 A2	6.36	405	168754	10	gi|157863686
5	Chain B, Crystal Structure Analysis Of Coagulation Factor VIII A3 C1C2	7.46	151	89376	4	gi|183448130
6	coagulation factor VIII, procoagulant component (hemophilia A), isoform CRAc [Homo sapiens]C2	6.97	149	268337	4	gi|119593053
ADVATE® digested with thrombin						
Spot number	**Protein name**	**ID**	**Mascot score %**	**p/M r**	**Number mathced peptides**	**NCBI Accession number**
1	Chain B, Crystal Structure Analysis Of Coagulation Factor VIII C1C2	7.94	325	59720	6	gi|183448130
2	Chain A, Crystal Structure Of Human Factor VIII C1C2A2	5.92	452	86866	13	gi|171849087
3	Structure Of HumanFactor VIII A1 frammento più lungo	5.92	227	86866	6	gi|171849087
4	Chain A, Crystal Structure Of Human Factor VIII A2	5.92	426	86866	13	gi|171849087
5	Chain A, Crystal Structure Of Human Factor VIII A2	5.92	385	86866	11	gi|171849087
6	Full=Prothrombin; AltName: Full=Coagulation factor II	5.97	93	71886	3	gi|135806
REFACTO®						
Spot number	**Protein name**	**ID**	**Mascot score %**	**p/M r**	**Number mathced peptides**	**NCBI Accession number**

1	B domain-deleted coagulation factor VIII [synthetic construct]tutto	6.36	1006	168754	28	gi|157863686
2	B domain-deleted coagulation factor VIII [synthetic construct]A1A2	6.36	1141	168754	52	gi|157863686
4	Chain B, Crystal Structure Analysis Of Coagulation Factor VIII c2	7.94	368	89376	12	gi|183448130
5	Chain B, Crystal Structure Of Human Factor VIII c2	7.46	1372	79418	43	gi|171849088
6	Chain B, Crystal Structure Analysis Of Coagulation Factor VIII c2	7.94	544	89376	17	gi|183448130

REFACTO® digested with thrombin						
Spot number	**Protein name**	**ID**	**Mascot score %**	**p/M r**	**Number mathced peptides**	**NCBI Accession number**
1	Chain A, Crystal Structure Of Human Factor VIII A1		421	86866	13	gi|171849087
2	B domain-deleted coagulation factor VIII [synthetic construct]	6.36	961	168754	25	gi|157863686
3						
4	Chain B, Crystal Structure Analysis Of Coagulation Factor VIII c2 un frammento	7.94	148	89376	3	gi|114667176
5	B domain-deleted coagulation factorVIII [syntheticconstruct]	7.94	362	168754	9	gi|157863686
6	glutathione S-transferase [unidentified cloning vector]	5.74	257	27323	7	gi|595706
7	Chain B, Crystal Structure Of Human Factor VIII c1c2	7.46	764	79418	21	gi|171849088
49	Chain B, Crystal Structure Of Human Factor VIII- A1a2			79418		gi|171849088
50	B domain-deleted coagulation factorVIII [synthetic	6.36	559	168754	17	gi|157863686
51	Chain A, Crystal Structure Of Human Factor VIII A2	5.92	676	86866	19	gi|171849087

Table I: Identification through HPLC-ESI-MS-MS of the band obtained in monodimensional gel of three commercially rFVIII products (Helixate NexGen®, Refacto® and Advate® before and after thrombin digestion .

2.3.1 Optimization of 2-DE protocol

Protein identification of bands observed in the first electrophoretic dimension did not reveal any contaminants besides the presence of full or truncated forms of rFVIII. Therefore, we performed a 2-DE analysis with the final aim to investigate traces of contaminants present in the three rFVIII preparations. To this regard, although 2-DE has proved to be a valuable method for the analysis of complex protein mixtures, such as blood plasma, the procedures utilized for the preparation of the samples are critical for the acquisition of high-quality results (Kim 2007). Helixate NexGen®, Refacto® and Advate® preparations are not as complex as blood plasma, although they contained recombinant proteins having the same physiochemical properties of the corresponding human plasma proteins. Protocol B differed from protocol A only for sample application method. In B, we used the cap-loading method, which produced the greatest number of detectable spots (Shevchenko1996). However, best

results were obtained with protocol C in which in addition to the cap-loading method, we used only 8M urea as a chaotropic agent and equilibrated the IPG strip at pH 6.8 in the presence of 3% w/v SDS. The caploading option provides simple sample loading into rehydrate IPG strip, improving sample uptake of very high molecular-weight proteins (Gorg 2004). The presence of thiourea chaotropic agents enhanced the total spot number, as shown by Pock et al. 1998 especially in the high-molecular-weight range. Our results, however, were focused on contaminants: in this respect, the reduction of pH value from 8.8 to 6.8 inequilibration buffer, increased the quality of isoelectrofocusing and the quantity of total spot number, especially in the range of 45-36 kDa without using thiourea. The low pH value reduced the net negative charge of the acid glycoprotein promoting the SDS binding and facilitating the transfer to the second dimension gel. On the basis of this improvement in 2-D map quality, protocol C was chosen to perform the rFVIII preparations analysis (Fig. 5).

2.3.2 Analysis of FVIII preparation by 2-D IEF-SDS PAGE

Two batches of three approved rFVIII preparations obtained from different manufacturers were analyzed by 2-D IEF-SDS- PAGE (Fig. 6) using the tested protocol C (see above). To evaluate the batch-to-batch reproducibility, three technical replicates were performed for each batch. Image analyses showed a high batch-to-batch reproducibility of the protein composition for all the three rFVIII preparations. The majority of the detected spots (Helixate NexGen®, 98.3%; Refacto®, 99.2%; Advate®, 99.1%) did not differ between the two batches for each of the rFVIII preparation analyzed (Brigulla 2006). A match set was created matching the protein patterns of six replicate gels for each rFVIII preparation including only those spots that were present in at least five technical replicates. Gel spots form Helixate NexGen®, Refacto® and Advate® master maps were subjected to an in-gel tryptic digestion and the proteins contained in each were identified by MS. Comparison of the 2-DE protein

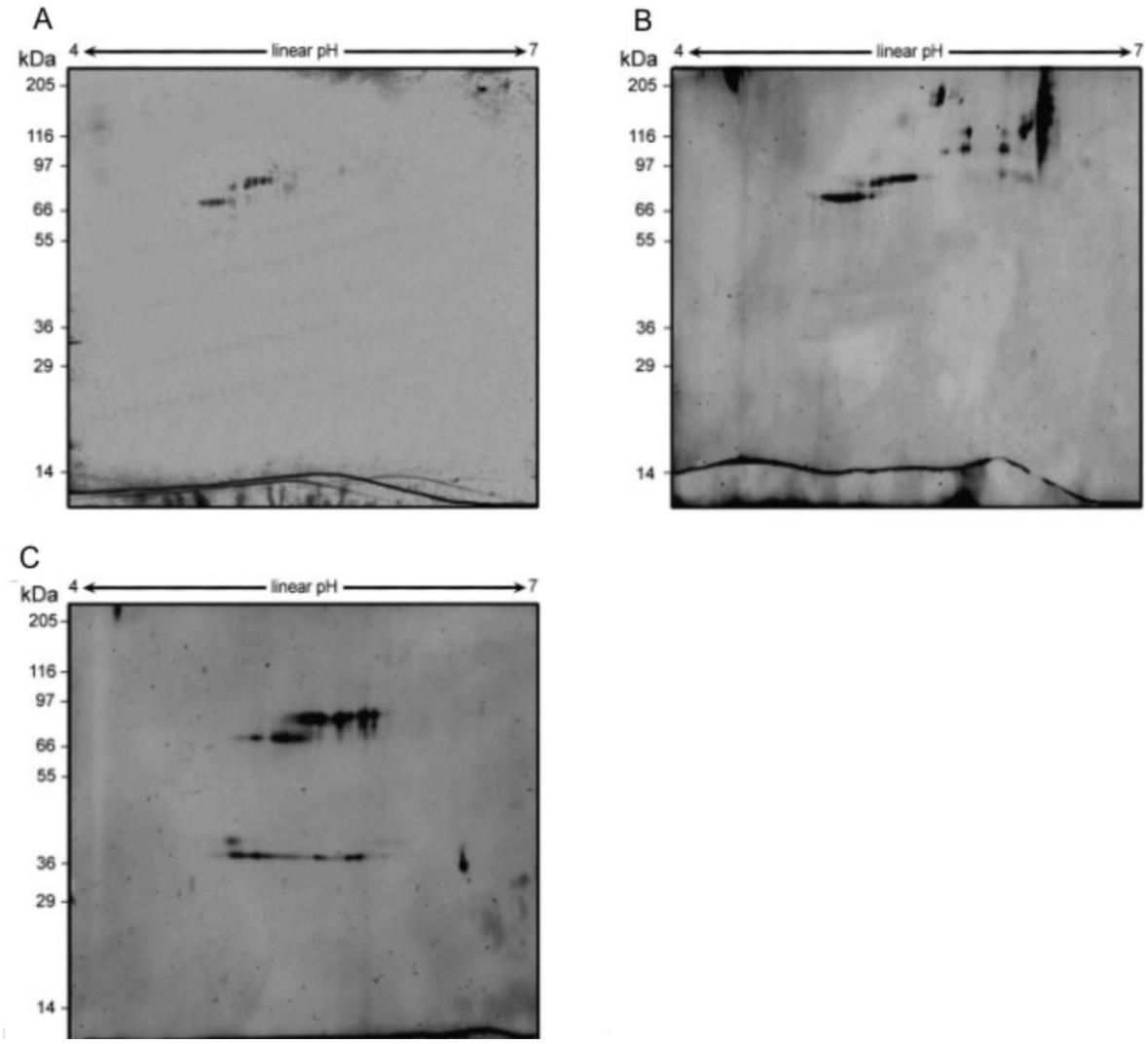


Figure 5: Comparison of three different IEF procedures (panels A, B and C) for the 2-DE analysis of Refacto preparation. The proteins (about 400 mg) were first separated in a linear pH gradient of 4–7, followed by separation in an SDS-PAGE (11%) and Blue Silver staining. For a detailed description of separation protocols.

and the proteins contained in each were identified by MS. Comparison of the 2-DE protein patterns revealed major differences in the protein composition among the three rFVIII preparations obtained from the different manufacturers. In all the three rFVIII preparation 2-DE maps, it was possible to characterize the A1-A2 heavy chain (Helixate NexGens® spots: 2, 3, 4; Refacto® spots: 14, 15, 16, 17, 18, 19; Advate ®spots: 30, 31,32, 33, 34, 35, 36) and the A3-C1-C2 light chain (Helixate NexGen® spots: 5, 6, 7; Refacto® spots: 20, 21, 22, 23; Advate® spots: 37, 38, 39, 40, 41). Curiously, both heavy and light chains showed a different

number of isoforms among the rFVIII preparations. Helixate NexGen® 2-DE map showed to contain three protein spots derived from the manufacturing process: (i) spot 1, which was identified as the heat shock protein 70 kDa; (ii) spots 8, 9, 10 and 11 identified as the plasma glycoprotein haptoglobin; and (iii) spots 12 and 13 identified as the plasma lipoprotein proapolipoprotein. All these proteins were identified in Homo sapiens. Some isoforms of the glutathione S-transferase (spots 25, 26, 27, 28 and 29) and the TEM extended spectrum b-lactamase from Escherichia coli were found in Refacto®. On the contrary, it was not possible to identify any anomalous contaminant in the Advate® preparation, although two anomalous spots (42 and 43) from the 2-DE map were not identified.

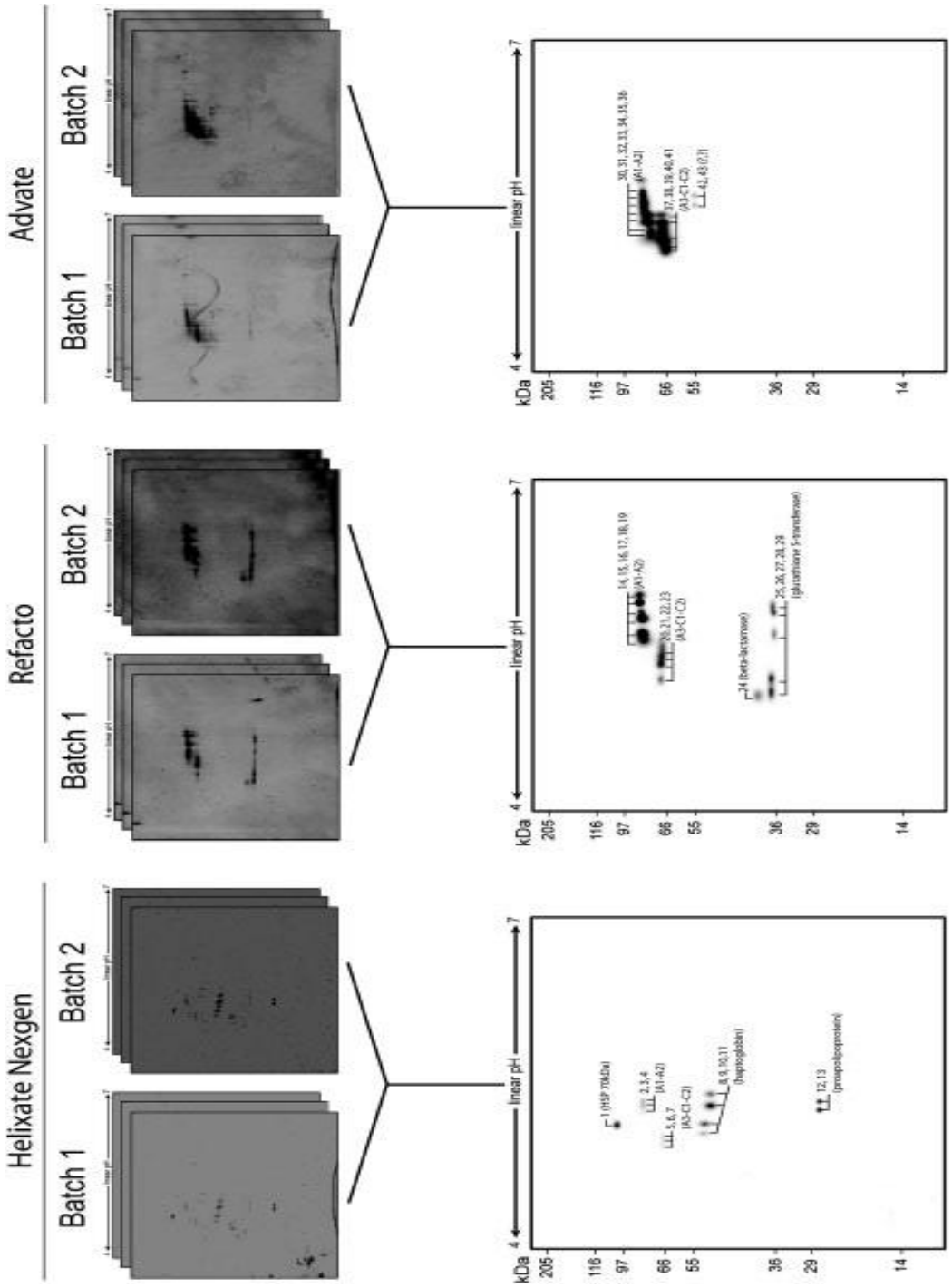


Figure 6: Representative Blue Silver stained 2-D gels of proteins from Helixate NexGens®, Refacto® and Advates®.(A) For each approved rFVIII preparation were analyzed two batches. For each batch were performed three technical replicate. (B) Master maps created matching the proteins that were present in at least five technical replicates on six. The proteins are marked by band number as reported in Table 3.

Spot No.	NEXGEN				Mascot Ion Score	NCBI Accession Number	Protein ID
	m/z	charge state	start-end a	sequence			
1	743.98	2 +	37-49	TTPSYVAFTDTER	50	gi|5729877	heat shock 70kDa protein 8 isoform 1 [Homo sapiens]
	833.04	2 +	57-71	NQVAMNPTNTVFDAKOxidation (M)	52		
	600.06	2 +	160-171	DAGTIAGLNVLI	46		
	495.06	2 +	510-517	LSKEDIER	28		
	659.97	2 +	540-550	NSLESYAFNMKOxidation (M)	49		
2	992.00	2 +	109-126	NMASHPVSLHAVGVSYWK	61	gi|119593052	coagulation factor VIII procoagulant component (hemophilia A) isoform CRA_b [Homo sapiens] A1-A2
	778.87	2 +	127-140	ASEGAEYDDQTSQR	87		
	821.41	3 +	141-161	EKEDDKVFPGGSHTYVWQVLK	50		
	750.96	2 +	186-199	DLNSGLIGALLVCR	57		
	440.12	2 +	233-239	NSLMQDROxidation (M)	42		
	717.89	2 +	233-245	NSLMQDRDAASAR	57		
	555.26	2 +	260-269	SLPGLIGCHR	40		
	665.30	2 +	345-355	VDSCPEEPQLR	47		
	777.38	2 +	379-391	FDDDNSPSFIQIR	60		
	816.96	2 +	561-575	DLASGLIGPLLICYK	34		
	726.86	2 +	591-602	NVILFSVFDENR	64		
3	751	2 +	186-199	DLNSGLIGALLVCR	58	gi|119593052	coagulation factor VIII, procoagulant component (hemophilia A), isoform CRA_b [Homo sapiens] A1-A2
	665	2 +	345-355	VDSCPEEPQLR	45		
	777	2 +	379-391	FDDDNSPSFIQIR	59		
	588	2 +	428-437	SQYLNNGPQR	66		
	634	2 +	447-456	FMAYTDETFK Oxidation (M)	38		
	616	2 +	532-542	WTVTVEDGPTK	56		
	656	2 +	551-560	YYSSFVNMER	57		
	816	2 +	561-575	DLASGLIGPLLICYK	47		
	726	2	591-602	NVILFSVFDENR	58		
4	778.871	2 +	127-140	ASEGAEYDDQTSQR	93	gi|182383	coagulation factor VIII [Homo sapiens] A1-A2
	751.024	2 +	186-199	DLNSGLIGALLVCR	46		
	440.188	2 +	233-239	NSLMQDR Oxidation (M)	39		
	484.208	3 +	233-245	NSLMQDRDAASAR Oxidation (M)	47		
	555.232	2 +	260-269	SLPGLIGCHR	37		
	665.324	2 +	345-355	VDSCPEEPQLR	39		
	777.874	2 +	379-391	FDDDNSPSFIQIR	69		
	588.816	2 +	428-437	SQYLNNGPQR	57		
	509.223	3 +	447-458	FMAYTDETFKTR Oxidation (M)	41		
5	553.219	2 +	135-143	HYFIAAVER	35	gi|183448130	Chain B Crystal Structure Analysis Of Coagulation Factor VIII A3-C1-C2
	555.255	3 +	144-157	LWDYGMSSSPHVLR Oxidation (M)	38		
	588.813	2 +	160-170	AQSGSVPQFK	34		
	720.371	2 +	203-214	AEVEDNIMVTFR Oxidation (M)	63		
	704.925	3 +	339-355	APCNIQMEDPTFKENYR Oxidation (M)	42		
	495.228	2 +	678-687	VTGVTTQGVK	46		
	722.281	2 +	759-770	MEVLGCEAQDLY Oxidation (M)	78		
6	525.15	2 +	1741-1750	AQSGSVPQFK	46	gi|182803	factor VIII [Homo sapiens] A3C1C2
	560.92	3 +	1769-1783	GELNEHLGLLGPYIR	44		
	720.29	2 +	1784-1795	AEVEDNIMVTFROxidation (M)	67		
	722.27	2 +	1853-1864	AWAYFSDVDLEK	54		
	681.75	2 +	1907-1916	SWYFTENMER	36		
	710.27	3 +	1920-1936	APCNIQMEDPTFKENYROxidation (M)	45		
	479.23	3 +	2092-2104	VDLLAPMIIHGIKOxidation (M)	41		
	646.32	2 +	2156-2166	HNIFNPPIIAR	42		
	596.73	2 +	2247-2255	EWLQVDFQK	26		
	683.24	2 +	2256-2268	TMKVTGVTTQGVKOxidation (M)	35		
	495.21	2 +	2259-2268	VTGVTTQGVK	62		
	529.21	2 +	2269-2277	SLLTSMYVKOxidation (M)	44		
	848.95	3 +	2301-2323	VFQGNQDSFTPVVNSLDPPLLTR	41		

7	826.3090 667.353 553.325 832.415 524.748 1.017.527 560.95 720.37 614.952 722.841 689.835 775.893 836.168 778.784 481.895 898.971 732.844 718.463 646.353 495.263 529.27 848.802 792.944 714.23	3+ 3 + 2 + 2 + 2 + 2 + 3 + 2 + 3 + 2 + 2 + 2 + 3 + 3 + 3 + 2 + 2 + 2 + 2 + 2 + 2 + 3 + 2 + 2 +	5-25 26-41 49-57 58-71 74-83 85-101 102-116 117-128 171-185 186-197 240-249 253-265 270-291 325-344 373-385 386-401 405-417 425-437 489-499 592-601 602-610 634-656 660-672 673-684	TTLQSDQEEIDYDDTISVEMK Oxidation (M) KEDFDIYDEDENQSPR HYFIAAVER LWDYGMSSSPHVLR AQSGSVPQFK VVFQEFTDGSFTQPLYR GELNEHLGLLGPYIR AEVEDNIMVTFR VQHHMAPTKDEFDCK AWAYFSDVDLEK SWYFTENMER APCNIQMEDPTF FHAINGYIMDTLPGLVMAQDQR MALYNLYPGVFETVEMLPSK CQTPLGMASGHIR DFQITASGQYGQWAPK LHYSGSINAWSTK VDLLAPMIIHGIK HNIFNPPIIAR VTGVTTQGVK SLLTSMYVK VFQGNQDSFTPVVNSLDPPLLTR IHPQSWVHQIALR MEVLGCEAQDLY	30 58 43 71 32 65 45 70 31 53 42 64 47 50 65 77 67 35 38 52 36 81 48 77	gi|171849088	Chain B Crystal Structure Of Human Factor VIII A3-C1-C2
8	462.10 471.61 490.57 380.52 862.14 637.64 729.81 673.05 602.11	2 + 2 + 2 + 2 + 2 + 2 + 3 + 2 + 2 +	103-111 195-202 219-227 233-238 239-252 253-263 267-286 321-332 333-342	ILGGHLDAK QKVSVNER.VGln->pyro-Glu (N-term Q) VGYVSGWGR FTDHLK YVMLPVADQDQCIROxidation (M) HYEGSTVPEKK SPVGVQPILNEHTFCAGMSKOxidation (M) SCAVAEYGVYVK VTSIQDWVQK	42 28 51 28 72 42 49 77 52	gi|1212947	haptoglobin [Homo sapiens]
9	462.24 673.25 602.32	2 + 2 + 2 +	101-109 319-330 331-340	ILGGHLDAK SCAVAEYGVYVK VTSIQDWVQK	26 41 54	gi|1212947	Haptoglobin [Homo sapiens]
10	462.02 460.46 471.46 617.64 490.48 862.00 573.47 729.68 673.61 601.97	2 + 2 + 2 + 3 + 2 + 2 + 2 + 3 + 2 + 2 +	103-111 112-119 195-202 203-218 219-227 239-252 253-262 267-286 321-332 333-342	ILGGHLDAK GSFPWQAK QKVSVNERGln->pyro-Glu (N-term Q) VMPICLPSKDYAEVGR.Oxidation (M) VGYVSGWGR YVMLPVADQDQCIR.Oxidation (M) HYEGSTVPEK. SPVGVQPILNEHTFCAGMSK. SCAVAEYGVYVK. VTSIQDWVQK.	36 28 30 26 48 77 27 48 38 49	gi|1212947	haptoglobin [Homo sapiens]
11	462.02 460.46 471.46 617.64 490.48 862.00 573.44 729.68 673.61 601.97	2 + 2 + 2 + 3 + 2 + 2 + 2 + 3 + 2 + 2 +	103-111 112-119 195-202 203-218 219-227 239-252 253-262 267-286 321-332 333-342	ILGGHLDAK GSFPWQAK QKVSVNER VMPICLPSKDYAEVGR VGYVSGWGR YVMLPVADQDQCIROxidation (M) HYEGSTVPEK SPVGVQPILNEHTFCAGMSKOxidation (M) SCAVAEYGVYVK VTSIQDWVQK	36 28 30 26 48 77 27 48 38 49	gi|1212947	haptoglobin [Homo sapiens]
12	700.96 434.47 405.834	2 + 3 + 3+	52-64 185-195 220 - 230	DYVSQFEGSALGK THLAPYSDELR ATEHLSTLSEK	31 39 26	gi|178777	proapolipoprotein

13	700.84 434.34	2 + 3 +	52-64 185-195	DYVSQFEGSALGK THLAPYSDELR	27 35	gi|178777	proapolipoprotein

Spot No.	REFACTO						
	m/z	charge state	start-end	sequence	Mascot Ion Score	NCBI Accession Number	Protein ID
14	778.82 440.15 665.34 777.31 588.77 634.79 616.84 656.29 726.85 655.35	2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2	127-140 233-239 345-355 379-391 428-437 447-456 532-542 551-560 591-602 603-612	ASEGAEYDDQTSQR NSLMQDR.DOxidation (M) VDSCPEEPQLR FDDDNSPSFIQIR SQYLNNGPQR FMAYTDETFKOxidation (M) WTVTVEDGPTK YYSSFVNMEROxidation (M) NVILFSVFDENR SWYLTENIQR	87 42 47 47 44 52 56 44 70 45	gi|157863686	B domain-deleted coagulation factor VIII [synthetic construct] A1-A2
15							NI
16	778.84 500.86 484.17 555.25 665.27 777.37 588.74 763.33 827.54 616.83 656.27 816.97 532.69	2 + 3 + 3 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 +	127-140 186-199 233-245 260-269 345-355 379-391 428-437 447-458 516-529 532-542 551-560 561-575 582-590	ASEGAEYDDQTSQR .DLNSGLIGALLVCR NSLMQDRDAASAR.Oxidation (M) .SLPGLIGCHR VDSCPEEPQLR FDDDNSPSFIQIR .SQYLNNGPQR.I FMAYTDETFKTR Oxidation (M) .HLKDFPILPGEIFK .WTVTVEDGPTK YYSSFVNMER.Oxidation (M) DLASGLIGPLLICYK GNQIMSDKR	69 38 52 39 60 75 50 61 55 56 54 56 45	gi|157863686	B domain-deleted coagulation factor VIII [synthetic construct]
17	777.86 588.76 634.75 616.83 648.29 726.87 655.81	2 + 2 + 2 + 2 + 2 + 2 + 2 +	406-418 409-418 428-437 513-523 532-541 572-583 584-593	SYKSQYLNNGPQR SQYLNNGPQR FMAYTDETFKOxidation (M) WTVTVEDGPTK YYSSFVNMEROxidation (M) NVILFSVFDENR SWYLTENIQR	62 40 59 56 42 63 47	gi|157863686	B domain-deleted coagulation factor VIII [synthetic construct]
18	778.81 725.86 665.30 777.41	2 + 2 + 2 + 2 +	108-121 214-226 326-336 360-372	ASEGAEYDDQTSQR NSLMQDRDAASAROxidation (M) VDSCPEEPQLR .FDDDNSPSFIQIR	94 33 42 51	gi|157863686	B domain-deleted coagulation factor VIII [synthetic construct]
19	588.78 634.76 509.23 616.74 656.29 524.73 726.88 655.30	2 + 2 + 3 + 2 + 2 + 2 + 2 + 2 +	409-418 428-437 428-439 513-523 532-541 563-571 572-583 584-593	SQYLNNGPQR FMAYTDETFKOxidation (M) WTVTVEDGPTK YYSSFVNMER GNQIMSDKROxidation (M) NVILFSVFDENR SWYLTENIQR	45 40 55 48 36 72 54	gi|171849087	Chain A Crystal Structure Of Human Factor Viii A1-A2

20	553.30	2 +	822-830	HYFIAAVER	42	gi|157863686	B domain-deleted coagulation factor VIII [synthetic construct] A3-C1-C2
	832.38	2 +	831-844	LWDYGMSSSPHVLROxidation (M)	74		
	588.78	2 +	847-857	AQSGSVPQFKK	31		
	841.02	2 +	875-889	GELNEHLGLLGPYIR	45		
	720.38	2 +	890-901	AEVEDNIMVTFROxidation (M)	74		
	847.45	3 +	902-922	NQASRPYSFYSSLISYEEDQR	37		
	722.39	2 +	959-970	AWAYFSDVDLEK	39		
	722.43	2 +	959-970	AWAYFSDVDLEK	39		
	681.76	2 +	1013-1022	SWYFTENMER	40		
	783.92	2 +	1026-1038	APCNIQMEDPTFKOxidation (M)	41		
	722.38	2 +	1146-1158	CQTPLGMASGHIROxidation (M)	46		
	898.98	2 +	1159-1174	DFQITASGQYGQWAPK	61		
	732.43	2 +	1178-1190	LHYSGSINAWSTK	47		
	646.36	2 +	1262-1272	HNIFNPPIIAR	43		
	705.87	2 +	1341-1352	SNAWRPQVNNPK	36		
	596.82	2 +	1353-1361	EWLQVDFQKGlu->pyro-Glu (N-term E)	48		
	683.84	2 +	1362-1374	TMKVTGVTTQGVKOxidation (M)	65		
	529.25	2 +	1375-1383	SLLTSMYVKOxidation (M)	41		
	722.33	2 +	1446-1457	MEVLGCEAQDLY-Oxidation (M)	72		
21	667.29	3 +	799-814	KEDFDIYDEDENQSPR	29	gi|157863686	B domain-deleted coagulation factor VIII [synthetic construct] A3-C1-C2
	553.23	2 +	822-830	HYFIAAVER	36		
	524.71	2 +	847-856	AQSGSVPQFK	32		
	689.71	2 +	1013-1022	SWYFTENMEROxidation (M)	30		
	710.33	3 +	1026-1042	APCNIQMEDPTFKENYROxidation (M)	34		
	481.83	3 +	1146-1158	CQTPLGMASGHIROxidation(M)	61		
	495.22	2 +	1365-1374	VTGVTTQGVK	47		
	714.31	2 +	1446-1457	MEVLGCEAQDLYOxidation (M)	82		
22	667.26	3 +	-799-814	KEDFDIYDEDENQSPR	36	gi|157863686	B domain-deleted coagulation factor VIII [synthetic construct] A3-C1-C2
	454.90	3 +	820-830	TRHYFIAAVER	38		
	832.43	2 +	831-844	LWDYGMSSSPHVLROxidation (M)	81		
	659.89	2 +	845-856	NRAQSGSVPQFK	39		
	679.02	3 +	858-874	VVFQEFTDGSFTQPLYR	28		
	560.94	3 +	875-889	GELNEHLGLLGPYIR	54		
	712.33	2 +	890-901	AEVEDNIMVTFR	40		
	847.74	3 +	902-922	NQASRPYSFYSSLISYEEDQR	34		
	722.33	2 +	959-970	AWAYFSDVDLEK	66		
	681.79	2 +	1013-1022	SWYFTENMEROxidation (M)	54		
	705.02	3 +	1026-1042	APCNIQMEDPTFKENYR	57		
	481.85	3 +	1146-1158	CQTPLGMASGHIROxidation (M)	59		
	898.94	2 +	1159-1174	DFQITASGQYGQWAPK	79		
	732.37	2 +	1178-1190	LHYSGSINAWSTK	39		
	473.88	3 +	1198-1210	VDLLAPMIIHGIK Oxidation (M)	35		
	646.37	2 +	1262-1272	HNIFNPPIIAR	45		
	937.40	3 +	1285-1308	STLRMELMGCDLNSCSMPLGMESK4 Oxidation (M)	38		
	780.00	3 +	1289-1308	MELMGCDLNSCSMPLGMESK	61		
	587.78	2 +	1353-1361	EWLQVDFQKGlu->pyro-Glu (N-term E)	36		
	683.37	2 +	1362-1374	TMKVTGVTTQGVK	57		
	529.24	2 +	1375-1383	SLLTSMYVK Oxidation (M)	31		
	848.78	3 +	1407-1429	VFQGNQDSFTPVVNSLDPPLLTR	56		
	793.04	2 +	1433-1445	IHPQSWVHQIALR	43		
	722.343	2 +	1446-1457	MEVLGCEAQDLY Oxidation (M)	79		
23	667.366	3 +	799-814	KEDFDIYDEDENQSPR	31	gi|157863686	B domain-deleted coagulation factor VIII [synthetic construct] A3-C1-C2
	553.29	2 +	822-830	HYFIAAVER	38		
	832.407	2 +	831-844	LWDYGMSSSPHVLROxidation (M)	72		
	524.837	2 +	847-856	AQSGSVPQFK	39		
	689.9	2 +	1013-1022	SWYFTENMEROxidation (M)	41		
	710.716	3 +	1026-1042	APCNIQMEDPTFKENYROxidation (M)	39		
	481.899	3 +	1146-1158	CQTPLGMASGHIROxidation (M)	62		
	646.899	2 +	1262-1272	HNIFNPPIIAR	34		
	495.272	2 +	1365-1374	VTGVTTQGVK	54		
	722.335	2 +	1446-1457	MEVLGCEAQDL	82		

24	654.314 544.177 580.686 643.87 638.369 451.144 892.482	2 + 2 + 3 + 2 + 2 + 3 + 2 +	42-53 72-81 145-159 191-202 221-232 242-255 256-271	VGYIELDLNSGK VLLCGAVLSR ELTAFLHNMGDHVTROxidation (M) LLTGELLTLASR SALPAGWFIADK GIIAALGPDGKPSR IVVIYTTGSQATMDER.NOxidation (M)	99 58 29 64 70 36 71	gi|112785110	TEM extended-spectrum beta-lactamase [Escherichia coli]
25	575.43 766.79 668.54 721.34	2 + 2 + 3 + 2 +	19-27 90-103 109-125 114-125	LLLEYLEEK AEISMLEGAVLDIR IAYSKDFETLKVDFLSK DFETLKVDFLSK	46 68 42 54	gi|595706	glutathione S-transferase [unidentified cloning vector]
26	575.35 516.78 766.86 668.62 721.43 343.68	2 + 2 + 2 + 3 + 2 + 2 +	19-27 65-73 90-103 109-125 114-125 219-224	LLLEYLEEK LTQSMAIIR AEISMLEGAVLDIR IAYSKDFETLKVDFLSK DFETLKVDFLSK SDLVPR	46 40 74 37 27 36	gi|595706	glutathione S-transferase [unidentified cloning vector]
27	575.35 516.78 721.43 343.68	2 + 2 + 2 + 2 +	19-27 65-73 114-125 219-224	LLLEYLEEK LTQSMAIIR DFETLKVDFLSK SDLVPR	55 38 29 41	gi|595706	glutathione S-transferase [unidentified cloning vector]
28	668.64 721.44 343.69	3 + 2 + 2 +	109-125 114-125 219-224	IAYSKDFETLKVDFLSK DFETLKVDFLSK SDLVPR	37 27 36	gi|595706	glutathione S-transferase [unidentified cloning vector]
29	575.351 516.781 668.621 721.438 343.683	2 + 2 + 3 + 2 + 2 +	19-27 65-73 109-125 114-125 219-224	LLLEYLEEK LTQSMAIIR IAYSKDFETLKVDFLSK DFETLKVDFLSK SDLVPR	46 40 86 45 32	gi|595706	glutathione S-transferase [unidentified cloning vector]

Spot No.	**ADVATE**						
	m/z	**charge state**	**start-end**	**sequence**	**Mass Ion Score**	**NCBI Accession Number**	**Protein ID**
30	778.795 665.831 777.359 588.748 634.77 616.799 656.307 655.354	2 + 2 + 2 + 2 + 2 + 2 + 2 + 2 +	127-140 345-355 379-391 428-437 447-456 532-542 551-560 603-612	ASEGAEYDDQTSQR VDSCPEEPQLR FDDDNSPSFIQIR SQYLNNGPQR FMAYTDETFK Oxidation(M) WTVTVEDGPTK YYSSFVNMEROxidation (M) SWYLTENIQR	49 54 77 46 39 53 55 55	gi|119593052	coagulation factor VIII procoagulant component (hemophilia A) isoform CRA_b [Homo sapiens]
31	778.81 484.112 665.265 588.72 634.769 616.739 656.236 655.339	2 + 3 + 2 + 2 + 2 + 2 + 2 + 2 +	127-140 233-245 345-355 428-437 447-456 532-542 551-560 603-612	ASEGAEYDDQTSQR NSLMQDRDAASAROxidation (M) VDSCPEEPQLR SQYLNNGPQR FMAYTDETFKOxidation (M) WTVTVEDGPTK YYSSFVNMEROxidation (M) SWYLTENIQR	58 37 37 52 39 47 41 37 30	gi|182383	coagulation factor VIII [Homo sapiens]
32	832.284 898.898 722.304	2 + 2 + 2 +	144-157 472-487 759-770	LWDYGMSSSPHVLROxidation (M) DFQITASGQYGQWAPKOxidation (M) MEVLGCEAQDLYOxidation (M)	31 46 54	gi|183448130	Chain B Crystal Structure Analysis Of Coagulation Factor Viii

32	553.263 555.281 588.851 710.633 494.955 722.334	2 + 3 + 2 + 3 + 2 + 2 +	135-143 144-157 160-170 339-355 678-687 759-770	HYFIAAVER LWDYGMSSSPHVLROxidation (M) AQSGSVPQFKK APCNIQMEDPTFKENYROxidation (M) VTGVTTQGVK MEVLGCEAQDLYOxidation (M)	32 39 31 36 49 70	gi|183448130	Chain B Crystal Structure Analysis Of Coagulation Factor Viii
33	778.812 725.862 665.308 777.418	2 + 2 + 2 + 2 +	108-121 214-226 326-336 360-372	ASEGAEYDDQTSQR NSLMQDRDAASAROxidation (M) VDSCPEEPQLR .FDDDNSPSFIQIR	94 33 42 51	gi|171849087	Chain A Crystal Structure Of Human Factor Viii
34	634.76 509.23 616.748 656.299 524.736 726.888 655.30	2 + 2 + 3 + 2 + 2 + 2 + 2 +	409-418 428-437 428-439 513-523 532-541 563-571 572-583	SQYLNNGPQR FMAYTDETFKOxidation (M) WTVTVEDGPTK YYSSFVNMER GNQIMSDKROxidation (M) NVILFSVFDENR SWYLTENIQR	45 40 55 48 36 72 54	gi|171849087	Chain A Crystal Structure Of Human Factor Viii
35	588.773 509.539 616.819 656.304 726.783 655.33	2 + 3 + 2 + 2 + 2 + 2 +	409-418 428-439 513-523 532-541 572-583 584-593	SQYLNNGPQR FMAYTDETFKTROxidation (M) WTVTVEDGPTK YYSSFVNMEROxidation (M) NVILFSVFDENR SWYLTENIQR	45 45 65 57 75 43	gi|171849087	Chain A Crystal Structure Of Human Factor Viii
36	778.81 484.112 665.265 588.72 634.769 616.739 656.236 655.339	2 + 3 + 2 + 2 + 2 + 2 + 2 + 2 +	127-140 233-245 345-355 428-437 447-456 532-542 551-560 603-612	ASEGAEYDDQTSQR NSLMQDRDAASAROxidation (M) VDSCPEEPQLR SQYLNNGPQR FMAYTDETFKOxidation (M) WTVTVEDGPTK YYSSFVNMEROxidation (M) SWYLTENIQR	58 37 37 52 39 47 41 37	gi|182383	coagulation factor VIII [Homo sapiens]
37	807.444 526.304 578.303 555.266 524.772 720.341 710.344 495.261 722.218	2 + 3 + 2 + 3 + 2 + 2 + 3 + 2 + 2 +	1477-1492 1493-1506 1654-1663 1725-1738 1741-1750 1784-1795 1920-1936 2259-2268 2340-2351	EVGSLGTSATNSVTYK KVENTVLPKPDLPK LCSQNPPVLK LWDYGMSSSPHVLROxidation (M) AQSGSVPQFK AEVEDNIMVTFROxidation (M) APCNIQMEDPTFKENYROxidation (M) VTGVTTQGVK MEVLGCEAQDLY-Oxidation (M)	58 28 37 56 26 70 42 58 59	gi|119593053	coagulation factor VIII procoagulant component (hemophilia A) isoform CRA_c [Homo sapiens]
38	759.342 363.151 807.444 526.30 578.30 555.26 524.77 720.341 710.344 495.261 722.218	2 + 2 + 2 + 3 + 2 + 3 + 2 + 2 + 3 + 2 + 2 +	1382-1395 1405-1411 1477-1492 1493-1506 1654-1663 1725-1738 1741-1750 1784-1795 1920-1936 2259-2268 2340-2351	GAITQSPLSDCLTR SPLPIAK EVGSLGTSATNSVTYK KVENTVLPKPDLPK LCSQNPPVLK LWDYGMSSSPHVLROxidation (M) AQSGSVPQFK AEVEDNIMVTFROxidation (M) APCNIQMEDPTFKENYROxidation (M) VTGVTTQGVK MEVLGCEAQDLY-Oxidation (M)	46 28 58 28 37 56 26 70 42 58 59	gi|119593053	coagulation factor VIII procoagulant component (hemophilia A) isoform CRA_c [Homo sapiens]
39	722.276 710.27 479.239 646.324 596.737 683.243 495.214 529.216	2 + 3 + 3 + 2 + 2 + 2 + 2 + 2 +	1853-1864 1920-1936 2092-2104 2156-2166 2247-2255 2256-2268 2259-2268 2269-2277	AWAYFSDVDLEK APCNIQMEDPTFKENYROxidation (M) VDLLAPMIIHGIKOxidation (M) HNIFNPPIIAR EWLQVDFQK TMKVTGVTTQGVKOxidation (M) VTGVTTQGVK SLLTSMYVKOxidation (M)	54 45 41 42 26 35 62 44	gi|182803	factor VIII [Homo sapiens]

40	555.299 588.795 720.415 683.380 495.325 529.26 722.32	3 + 2 + 2 + 2+ 2 + 2 + 2 +	1725-1738 1741-1751 1784-1795 2256 - 2268 2259-2268 2269-2277 2340-2351	LWDYGMSSSPHVLROxidation (M) AQSGSVPQFKK AEVEDNIMVTFROxidation (M) TMKVTGVTTQGVK Oxidation (M VTGVTTQGVK SLLTSMYVKOxidation (M) MEVLGCEAQDL-Oxidation (M)	47 35 74 38 60 38 53	gi|182383	coagulation factor VIII [Homo sapiens]
41	722.276 681.755 710.27 479.239 646.324 596.737 683.243 495.214 529.21 848.95 714.28	2 + 2 + 3 + 3 + 2 + 2 + 2 + 2 + 2 + 3 + 2 +	1853-1864 1907-1916 1920-1936 2092-2104 2156-2166 2247-2255 2256-2268 2259-2268 2269-2277 2301-2323 2340-2351	AWAYFSDVDLEK SWYFTENMER APCNIQMEDPTFKENYROxidation (M) VDLLAPMIIHGIKOxidation (M) HNIFNPPIIAR EWLQVDFQK TMKVTGVTTQGVKOxidation (M) VTGVTTQGVK SLLTSMYVKOxidation (M) VFQGNQDSFTPVVNSLDPPLLTR MEVLGCEAQDLY	54 36 45 41 42 26 35 62 44 41 62	gi|182803	factor VIII [Homo sapiens]
42							Not identified
43							Not identified

Table II : Spot identification of 2DE in NexGen Helixate®, Refacto® and Advate®

2.4 Discussion

During my research I use proteomics approach to reveal the heterogeneity and the possible presence of contaminants in three commercially rFVIII products, commonly used in the treatment of hemophilia A: Helixate NexGen®, Refacto® and Advate®. Although biochemical, functional and most importantly, clinical studies of rFVIII have been reported (Jankowski 2007; Josic 2006), information regarding eventual contaminants present in commercial products is still limited. From monodimensional gels it's possible showed the presence of multiple bands in Helixate NexGen® and Advate®, but not in Refacto® product. MS/MS analysis of these multiple bands revealed that they represent truncated forms of the full-length protein. Most of the cleavage took place into the B domain. This is in agreement with the previous reports by Hansen 1997. The author of this work determined the origins of the putative protease(s) responsible for the cleavage in the C-terminal part of rFVIII heavy chain. They incubated purified rFVIII with intact cells, lysed cells, or with fresh medium, followed by separation of intact and cleaved rFVIII by F25-affinity chromatography. With this experiment they observed cleavage points in rFVIII heavy chain, i.e. at Glu720 and Tyr729 (in the C-terminal part of the heavy chain). They showed that cleavage took place only after secretion of the protein from the cell and concluded that lyzed cells released proteinase(s) that could be responsible for the cleavage of rFVIII. So, we can asses that this

proteolytic attack is a general phenomenon occurring in any host cell (BHK and CHO cells) used during manufacturing processes and for all rFVIII taken into account in this study. Moreover, for the first time, two forms of A3-C1-C2 domains are observed both before and after thrombin digestion, although further investigations are mandatory to better characterize the sequence of these proteins as reported for A2 domains. However, it is interesting that, upon thrombin digestion, all heterogeneous bands disappeared, giving rise to similar digestion products. In the case of Helixate NexGen® and Advate®, B domain was removed from full-length protein, as previously observed by Jankowski et al. 2007. Nevertheless, the B-domain is not required for FVIII procoagulant activity (Lenting 1998; Pittman 1994) . It normally has a role in intracellular trafficking during the synthesis of FVIII but it is not relevant to the procoagulant function of FVIII (Toole 1996). Another important consideration is that, upon thrombin digestion, all rFVIII produce four fragments, out of which two derive from the A2 domain: one containing the entire A2 domain [FVIII-(372–740)] and the other one truncated A2 in the C-terminal region [FVIII-(372–729)], which has 11 amino acids missing. The latter comes out from a recombinant precursor of FVIII truncated into the C-terminal region. According to Hansen et al. 1997, this cleavage on the precursor could be mediated by cell surface bound proteases, by protease present in the serum used in the production medium or by proteases released from cells upon death and lysis. However, the fact that we found these truncated forms in the C-terminal region, after thrombin digestion, in all the investigated commercial recombinants confirms that this cleavage is independent from the host cell used for gene expression. In this work the authors characterized four different form of recombinant factor VIII B domain deleted. Two of the factor VIII forms had the expected heavy chain consisting of amino acids 1 -740, the two other factor VIII form had C-temini of the heavy chain at Glu720 [FVIII-(l -720) or Tyr729 [FVIII-(1-729)]. Then they measured the specific clotting activity of the different form ad they underlined that in fragment FVIII (1-720) the activity was two fold lower respect the activity shown in fragment (1-729), indicating that amino acids 721-729 are required for full procoagulantactivity. (Kjale 1995) . So we conclude that there is a form less active that can in part compromised the normal activity of factor VIII. By using 1-D gel, only the heterogeneous form of FVIII has been revealed, while eventual contaminants seemed to be absent. To this purpose, a 2-D gel was set up not only to better characterize FVIII concentrates but also and mainly to highlight the presence of impurities and identify eventual unexpected proteins. As a result, by using 2-D IEF-SDS-PAGE, Helixate NexGens®, Refacto® and Advate® contained between 16 and 13 different

protein spots belonging to more than six different proteins (Table 3). Different rFVIII preparations showed a similarity between batches to batches being for: Helixate NexGens®, 98.3%; Refacto®, 99.2%; Advate®, 99.1%, the major differences among the three products might be related to the different production and purification processes applied by the three manufacturers (Brigulla 2006). Helixate NexGen® showed the presence of impurities such as Hsp70 kDa, haptoglobin and proapolipoprotein, whereas glutathione S-transferase and b-lactamase were detected as contaminants in Refacto®. Interestingly, Advate® did not apparently contain any identified contaminant, except for two faint spots that were not identified, even by using MS. However, Basilico et al. (2010), in a recent study, individuated a series of human serum proteins as contaminants in Advate®. The authors claimed that the low number of peptides identified in biotechnological preparations could be due to the presence of albumin. From our analysis, it emerged that all rFVIII batches taken into account did not contain relevant amounts of albumin. The presence of Hsp70kDa in Helixiate NexGen® is not surprising because an established strategy for maximizing recombinant therapeutic protein production is to inhibit apoptosis through the expression of anti-apoptotic genes (Al-Rubeai 1998; Arden 2004). Hsp70kDa is over-expressed in BHK-21 cells engineered to produce rFVIII in order to promote resistance to apoptosis and enhance secretion. One of the major obstacles in the generation of rFVIII is the low expression level, which is about 2–3 orders of magnitude lower than that of comparable-size proteins using similar vectors (Lynch 1993; Soukharev 2002). In conclusion, many authors (Samali 1996; Creagh 2000) supported the necessity of the usefulness of Hsp70 over-expression in genetically engineered cells, but it should be removed in the final purification steps. However, our results showed that traces of the molecular chaperone Hsp70 are still present in the final product of Helixate NexGen®. In Helixate NexGen®, we found also haptoglobin and pro-apolipoprotein. The latter is used to increase the expression levels of rFVIII through the insertion of spliceable nucleotide sequences into FVIII cDNA, whereas haptoglobin, a plasma protein, is revealed in quality control test from fermentation (Jiang 2002). Surprisingly enough, we found in Refacto® a spot corresponding to b-lactamase from E. coli, which probably constitutes a remnant of intermediate plasmid preparation phases. Moreover, several spots of the glutathione S-transferase protein were found. This enzyme could be expressed as a fusion protein in the so-called GST gene fusion system. This system is used to purify and detect proteins of interest. In a GST gene fusion system, the GST sequence is incorporated into an expression vector alongside the gene sequence encoding the protein of interest.

Induction of protein expression from the vector's promoter results in expression of a fusion protein: the protein of interest (rFVIII in this case) fused to the GST protein. This GST-fusion protein can then be purified from cells via its high affinity for glutathione. While glutathione S-transferase is expected to be easily cleaved upon expression, as it contains thrombin cleavage sites, we could detect traces of this protein in the Refacto® preparation.

2.5 Conclusion

In conclusion, from the beginning of antihemophilic treatment with plasma derivatives, major benefits for the patient have been achieved by using recombinant clotting FVIII concentrates. Nonetheless, a number of problems have remained unresolved regarding generation of adverse clinical events. To contribute finding a possible explanation of untoward reactions in the recipients, our investigation revealed that, through the use of full-sequence FVIII cDNA, a heterogeneous mixture of various B-domain-truncated forms of the molecule were obtained in all of the three recombinants taken into account. All of them showed truncated forms in the C-terminal region, probably related to protease activity into host cells. Moreover, NexGen Helixate® and Refacto® preparation contain some impurities coming from the manufacturing process. Although, now we are not in the position to asses that the presence of these contaminants and/or truncated form of rFVIII, related to manufacturing process, could be the reason of the development of inhibitory alloantibodies against factor or any other allergic reactions. With this study we have demonstrated that the bidimensional electrophoresis, followed by mass spectrometry analysis, is a potent tool to reveal the consistency of manufacture process in monitoring batch to batch variations.

Reference:

Al-Rubeai M., Singh R.P. (1998) Apoptosis in cell culture. Curr. Opin. Biotechnol. 2: 152–156.

Andersson L.O., Forsman N., Huang K., Larsen K., Lundin A., Pavlu B., Sandberg H., Sewerin K., Smart J. (1986) Isolation and characterization of human factor VIII: molecular forms in commercial factor VIII concentrate, cryoprecipitate, and plasma. Proc. Natl. Acad. Sci. USA, 83: 2979–2983.

Arden N., Betenbaugh M.J. (2004) Life and death in mammalian cell culture: strategies for apoptosis inhibition. Trends Biotechnol. 22: 174–180.

Basilico F., Nardini I., Mori F., Brambilla E., Benazzi L., De Palma A., Rosti E. (2010) Characterization of factor VIII pharmaceutical preparations by means of MudPIT proteomic approach. J. Pharm. Biomed. Anal. 1: 50–57.

Clifton, J. G., Huang, F., Kovac, S., Yang, X., Hixson, D. C., Josic, C. (2009) Proteomic characterization of plasma-derived clotting factor VIII-von Willebrand factor concentrates. Electrophoresis 30: 3636–3646.

Creagh, E. M., Carmody, R. J., Cotter, T. G. (2000) Heat shock protein 70 inhibits caspase-dependent and independent apoptosis in Jurkat T cells. Cell. Res. 257: 58–66.

Dorner, A. J., Bole, D. G., Kaufman, R. J. (1987) The relationship of N-linked glycosylation and heavy chain-binding protein association with the secretion of glycoproteins. J Cell Biol. 105: 2665-2674.

Eaton D.L, Hass P.E., Riddle L., Mather J., Wiebe M., Gregory T., Vehar G.A. (1987) Characterization of recombinant human factor VIII. J Biol Chem 7: 3285-3290.

Eriksson R.K., Fenge C., Lindner-Olsson E., Liungggvist C., Rosenguist J., Smeds A. L. (2001) The manufacturing process for B-domain deleted recombinant factor VIII. Semin Hematol. 38:24-31.

Fay P.J., Anderson M.T., Chavin S.I., Marder V.J. (1986) Model for the factor VIIIa-dependent decay of the intrinsic factor Xase. Role of subunit dissociation and factor IXa-catalyzed proteolysis.Biochim Biophys. Acta 871: 268–278.

Fay, P.J. Haidaris P.I. Smudzin. T. M. (1991) Human factor VIII., subunit structure. Reconstitution of factor VIIIa, from the isolated A l l A3-CLC2 dimer and A2 subunit, J. Hiol. Chem 266:8057-8962.

Fay P. J. (2004) Activation of factor VIII and mechanisms of cofactor action.Blood 1:1–15.7.

Fulcher C.A., Roberts J.R., Zimmerman T.S. (1983) Thrombin proteolysis of purified factor VIII procoagulant protein: correlation of activation with generation of a specific polypeptide. Blood 4: 807–811.

Kjale M., Hending A., Talbo G., Persson E., Thomsen J., Ezban M. (1995) Amino acid residues 721-729 are required for full factor VIII activity. Eur. J. Biochem. 234: 773–779.

Kim M.R., Kim C.W. (2007) J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 849: 203–210.

Goedert J.J., Chen B.E., Preiss L, Aledort L.M., Rosenberg P.S. (2007) The manufacturing process for B-domain deleted recombinant factor VIII Am J Epidemiol 165:1443–1453.

Gomperts E., Lundblad R., Adamson R. (1992) The manufacturing process of recombinant factor VIII, recombinate. Transfus. Med. Rev. 6: 247–251.

Gorg A., Weiss W., Dunn M. J. (2005) Current two-dimensional electrophoresis technology for proteomics. Proteomics. Dec;4(12):3665-3685. Review. Erratum in: Proteomics. (3):826-827.

Gulla M., Thiele T., Scharf C, Ruddock, S. B., Venz, S., Vo¨ lker, U., Greinacher, A. (2006) Proteomics as a tool for assessment of therapeutics in transfusion medicine: evaluation of prothrombin complex concentrates.Transfution 46: 377–385.

Hansen K., Kjalke M., Rasmussen P. B., Kongerslev P., Ezban M. (1997) Proteolytic cleavage of recombinant two-chain factor VIII during cell culture production is mediated by protease(s) from lysed cells. Cytotechnology, 24: 227–234.

Ikkalat E., Helske G., Myllyla H., Nevanlinna H.R., Pitkänen P., Rasi V. (1982) Changes in the life expectancy of patients with severe haemophilia A. Br. J. Haematol, 52:7–12.

Jankowski M.A., Patel H., Rouse J.C., Marzilli L.A., Weston S.B. Sharpe P.J. (2007) Defining 'full-length' recombinant factor VIII: a comparative structural analysis. Haemophilia. 13:30-37.

Jiang R., Monroe T., McRogers R., Larson P.J. (2002) Manufacturing challenges in the commercial production of recombinant coagulation factor VIII. Haemophilia, 8: 1–5.

Josic D., Brown M.K., Huang F., Lim Y.P., Rucevic M., Clifton, J.G., Hixson D.C. (2006) Proteomic characterization of inter-alpha inhibitor proteins from human plasma. Proteomics, 6: 2874–2885.

Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–686.

Lynch C.M., Israel D.I., Kaufman R.J., Miller A.D. (1993) 4: 259–272. Sequences in the coding region of clotting factor VIII act as dominant inhibitors of RNA accumulation and protein production. Hum Gene Ther. 3:259-272.

Lenting P.J., van Mourik J.A., Mertens K. (1998). The life cycle of coagulation factor VIII in view of its structure and function. Blood, 11: 3983–3996.

Pipe S. W., Morris J. A., Shah J., Kaufman R.J. (1998) Differential interaction of coagulation factor VIII and factor V with protein chaperones calnexin and calreticulin. Biol. Chem. 273: 8537–8544

Pock K., Jungbauer A., Kannicht C., Loster K., Buchacher A., Josic D. (1998) Anal. Chim. Acta 372: 219–229.

Pittman, D. D., Marquette, K. A., Kaufman, R. J. (1994) Blood, 84: 4214–4225. Role of the B domain for factor VIII and factor V expression and function.Blood. 1994 Dec15;84(12):4214-25.

Shevchenko A. Wilm M., Vorm O., Mann M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem 68: 850–858.

Soukharev, S., Hammond, D., Ananyeva, N. M., Anderson, J. A., Hauser, C. A., Pepe, S., Saenko, E. L. (2002) Blood Cells Mol. Dis. 28: 234–248.

Samali, A., Cotter, T. G. (1996) Exp. Cell. Res. 223:163–170. Heat shock proteins increase resistance to apoptosis. Exp Cell Res. **1996** Feb 25;223(1):163-70.

Shen, B.W., Spiegel, P.C., Chang, C.H., Huh, J.W., Lee, J.S., Kim,J.,Kim, Y.H., Stoddard, B.L.(2008)The tertiary structure and domain organization of coagulation factor VIII.Journal Blood 111: 1240-1247

Toole J. J., Pittman, D. D., Orr, E. C., Murtha, P., Wasley, L. C., Kaufman, R. (1986) J., Proc. Natl. Acad. Sci. USA 83: 5939–5942

Triemstra M, Rosendaal FR, Smit C, Van der Ploeg HM, Briët E.(1995) Mortality in patients with hemophilia. Changes in a Dutch population from1 986 to 1992 and 1973 to 1986. Ann Intern Med. 123:823–827

Vehar G. A., Keyt, B., Eaton, D., Rodriguez, H., O'Brien, D. P., Rotblat, F., Oppermann, H., Keck, R., Wood, W. I., Harkins, R. N., Tuddenham, E. G. D., Lawn, R. M., Capon, D. J.,(1984) Nature 312: 337–342.

Chapter 3

Comparison among plasma-derived clotting Factor VIII by using monodimensional gel electrophoresis and mass spectrometry

3.1 Introduction

During the past several decades, hemophilia treatment has advanced considerably as a result of the increased availability and improvement of factor VIII (FVIII) replacement therapy. Technological advances have improved the safety and efficacy of factor replacement products, but patients with haemophilia commonly continue to encounter a wide complication in their treatment due to the emergence of inhibitors, or alloantibodies, in response to factor replacement therapy.Development of factor VIII (FVIII) inhibitors is the most severe and challenging complication of haemophilia treatment. Inhibitors are antibodies that neutralize factor VIII. It is considered as one of the most severe and important complication of hemophilia treatment. Generally the incidence of inhibitor formation is 15% to 20% in Haemophiliac's population. The immune systems in these individuals would be expected to perceive the epitope the missing factor VIII molecule as "foreign" and trigger an immune response against those foreign epitopes (Oldenburg 2006; Boekhors 2008) .

It is not know whether the presence of excess of other plasma proteins, in addition to von Willebrand factor, could stimulate untoward immune responses in the recipient. In plasma prepared FVIII, viral safety is an essential requirement and the purification process should inactivate both enveloped and non-enveloped viruses. Solvent detergent and heat treatment are frequently used to inactivate both viruses. According to literature data, these treatments could result in undesired effect such as: (a) increase of the antigenic properties of the injected material that may induce formation of inhibitory antibodies (Abs) in patients with severe haemophilia A9-11; (b) increase of high molecular weight (HMW) aggregates which may interfere with the complete dissolution of the lyophilised FVIII concentrates, when properly reconstituted. Therefore the European Pharmacopoeia is modifying the FVIII monograph, allowing the use of a filter, provided in the package, to eliminate small flakes or particles. Regardless of the grade of purity, all FVIII concentrates and especially plasma derived concentrates could present formation of HMW aggregates, having variable solubility, even if the formation mechanism and aggregates composition can be different. Nowadays it is not known if virus inactivation processes may induce protein alterations as well, as pdFVIII may

contain impurities or contaminants. To asses the purity of plasma derived factor VIII in this study we characherized through proteomic's approach two important pdFVIII commercially in use : Emoclot® 1000 UI/10mL (Kedrion), and Beriate® (Aventis Behring).

3.2 Materials and method

Commercially-available pdFVIII concentrates were used in this study. Materials consisted of Beriate® 1000UI/10mL (Aventis Behring) batch 81865011E and Emoclot® 1000 UI/10mL batch KA0805 (Kedrion), both products were reconstituted according to the manufacturers' instructions. Protein concentration was estimated by the Bradford method. 40 μg of samples were precipitated with 80% acetone for 2h at –20 °C. The sample was centrifuged at 14 000 rpm for 12 min, the supernatant was removed and the pellet was air dried. The pellets were suspended in 100 μL of sample buffer and were incubated for 20 min in100 mM DTT, 12% sucrose, 0.6% urea, 50 mM Tris-HCl pH 6.8, 0.01% bromophenol blue at room temperature.
SDS-PAGE was performed according to Laemmli's procedure under reducing conditions. Reagents for electrophoretic analyses were obtained from Bio-Rad. The experiment was performed on a 12% , 11%, 9% and 7.5% acrylamide gel, in a chamber with the following dimension 12x16x18 cm. The second electrophoresis gel was performed at 7.5% using a chamber with the following dimension 29x32x17 cm. A Protean II Bio-Rad (Hercules, CA,USA) electrophoresis system (1803160 mm, 0.75 mm thick) was used for the first dimension. Gels were fixed and stained for 30 min in a 5:1:4 (v/v) methanol glacial acetic acid-water mixture. Protein bands were stained with Coomassie Brilliant Blue R-250.

3.2.1 In-gel-digestion

Bands from the monodimensional gel were carefully excised and subjected to in-gel trypsin digestion according to Shevchenko 1996 with minor modifications. The gel pieces were swollen in a digestion buffer containing 50 mM NH4HCO3 and 12.5 ng/μL trypsin (modified porcine trypsin, sequencing grade, Promega, Madison, WI) in an ice bath. After 30 min, the supernatant was removed and discarded, 20 μL of 50 mM NH4HCO3 were added to the gel pieces, and digestion was allowed to proceed at 37 °C overnight. The supernatant containing tryptic peptides was dried by vacuum centrifugation.

3.2.2 RP-nanoHPLC mass spectrometry

Prior to mass spectrometric analysis, the peptide mixtures were re-dissolved in 10 μL of 5% FA (formic acid). Peptide Sequencing was performed by Nano-RP-HPLC-ESI-MS/MS. Peptide mixtures were separated using a nano flow-HPLC system (Ultimate;Switchos; Famos; LC Packings, Amsterdam, The Netherlands). A sample volume of 10 μL was loadedthrough the autosampler onto a homemade 2 cm fused silica precolumn (75 μm i.d.; 375 μm o.d.; Reprosil C18-AQ, 3 μm, Ammerbuch- Entringen, Germany) at a flow rate of 2 μL/min. Sequential elution ofpeptides was accomplished using a flow rate of 200nL/min and a linear gradient from Solution A (2% acetonitrile; 0.1% formic acid) to 50% of Solution B (98% acetonitrile; 0.1%formic acid) in 40 min over the precolumn in-line with a homemade 10-15 cm resolving column (75 μm i.d.; 375 μmo.d.; Reprosil C18-AQ, 3 μm, Ammerbuch-Entringen). Peptideswere eluted directly into a high capacity ion trap(model HCTplus Bruker-Daltonik, Germany). Capillary voltage was 1.5-2 kV and a dry gas flowrate of 10 L/min was used with a temperature of 200 °C. The scan range used was from 300 to 1800 m/z. Protein identification was performed by searching inthe National Center for Biotechnology Information non redundant (NCBInr) database using the MASCOT program (http://www. matrixscience.com). The following parameters were adopted for data base searches: complete carbamido methylation of cysteines, partial oxidation of methionines, peptide mass tolerance (1.2 Da, fragment mass tolerance (0.9 Da and missed cleavages 2. For positive identification, the score of the result of [-10 × Log(P)] had to be over the significance threshold level ($P < 0.05$).

3.3 Results

Two commercially-available preparations of plasma FVIII concentrates Beriate® and Emoclot® were examined to determine eventual impurities and/or contaminants. To this regard, SDS-PAGE was used for protein separation, and mass spectrometryfor protein identification. Figure 7 compares the minigel electrophoretic separation pattern of pdFVIII from two different manufacturers: Emoclot® and Beriate®. We use different concentrations of acrylamide, ranging from 7.5% to 12 % (see Figure 1 A-D) and mini gel to set up the best condition. As shown in fig. 1, best results were obtained by using acrylamide at 7,5%. Table II reports the identification spots numbered as in Figure 7. The most abundant proteins are in the spot 1 and 2 for both Emoclot ® and Beriate®. They contain vWF as a predominant protein

with 49 recognized peptides covering the 17% of sequence. The percentage excess of vWF over pdFVIII is more than 90% in pdFVIII preparations. In both cases we found a co-eluted fibronectin 1 protein which is one of the most abundant components in the blood plasma. Besides the presence of fibronectin, both samples showed a significant amount of inter-alpha inhibitor (IaI), pre-alpha inhibitor (PaI) which represented the major contaminant, together to fibrinogen and Kininogen. Interestingly Emoclot® contained human IgM heavy chain (spot 6) and in the same band we found FVIII, which is very weak comparing to fibrinogen and vWF. In Beriate®, it is interesting to underline that we detected the presence galectin 3 binding protein (G3BP), which co-eluted with FVIII. In this case we are in presence of truncated form of FVIII revealed by mass spectrometry as "Chain A, Crystal Structure Analysis Of Coagulation FVIII" with accession number: gi|183448129.

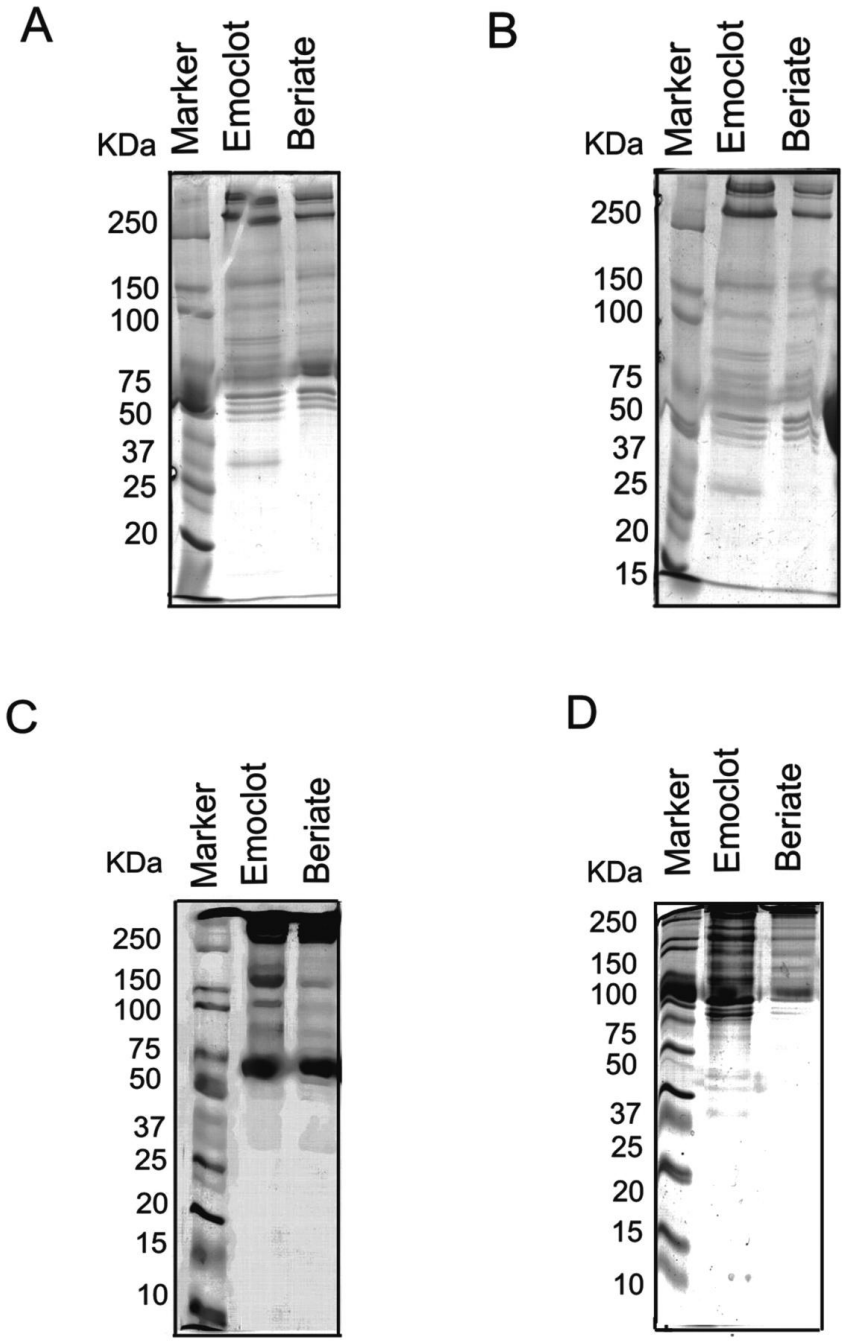


Figure 7 : Mini gel SDS PAGE of two pdVIII preparations: Emoclot ® 1000 UI/10 mL batch KA0805 Kedrion and Beriate® P1000 batch 81865011E CLS Behring . Samples are reduced according to Laemmli's17 method and are analysed at different concentration of acrilamide a) 7,5% b) 9% c) 11% d) 12%

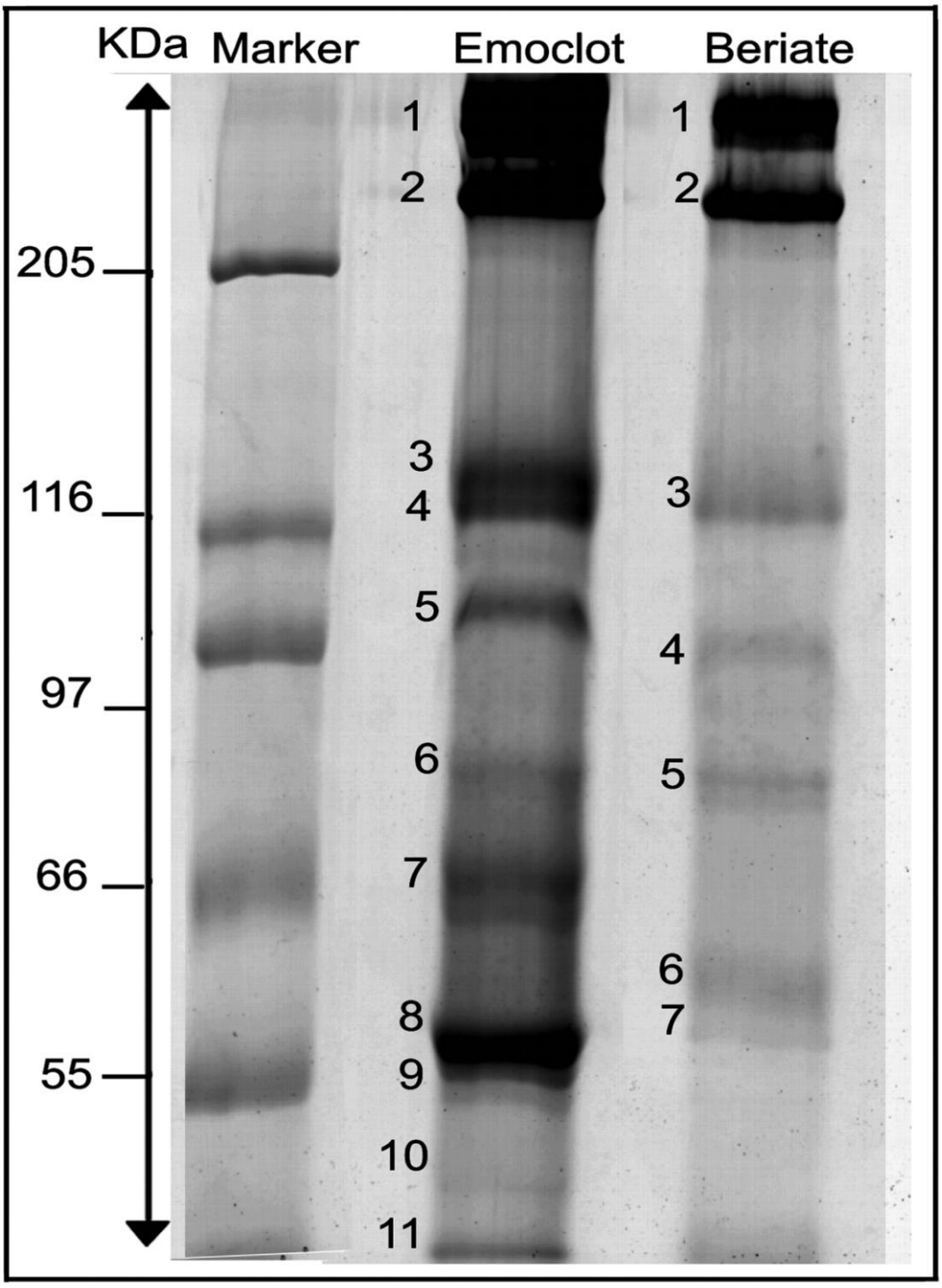


Figure 8 : SDS-PAGE (7.5%) of reduced samples (pdFVIII) according to Laemmli's method.Protein are detected by Coomassie Brilliant Blue. Lane 1 HMW standard , Lane 2: 40 ug of pdFVIII Emoclot ® 1000 UI/10 mL batch KA0805 Kedrion, Lane 3: 40 ug of pdFVIII Beriate® P1000 batch 81865011E CLS Behring

BERIATE						
Spot number	**Protein name**	**ID**	**Mascot score %**	**p/M r**	**Number mathced peptides**	**NCBI Accession number**
1 BE	von Willebrand factor prepropeptide	5.44	1756	237661	53	gi|340361
2 BE	inter-alpha (globulin) inhibitor H2 [Homo sapiens]	6.59	849	99813	29	gi|119585664
2 BE	alpha-1-microglobulin; bikunin [Homo sapiens]	5.58	251	17215	5	gi|579676
3 BE	coagulation factor VIII [Homo sapiens]		175	268250	4	gi|182383
3 BE	inter-alpha (globulin) inhibitor H2 [Homo sapiens]	6.56	816	105606	20	gi|55958063
4 BE	galectin 3 binding protein [Homo sapiens]	5.13	285	66202	5	gi|5031863
5 BE	beta-fibrinogen precursor	8.31	819	55545	20	gi|182430
6 BE	kininogen 1 isoform 2 [Homo sapiens]	6.29	147	48936	3	gi|4504893
7 BE	fibrinogen gamma chain [Homo sapiens]	5.61	293	50077	6	gi|182439
EMOCLOT						
Spot number	**Protein name**	**ID**	**Mascot score %**	**p/M r**	**Number mathced peptides**	**NCBI Accession number**
1 EMO	von Willebrand factor preproprotein [Homo sapiens]	5.30	1665	322429	55	gi|89191868
2 EMO	von Willebrand factor prepropeptide	5.44	703	237661	20	gi|340361
3 EMO	alpha-1-microglobulin; bikunin [Homo sapiens]	5.58	172	17215	3	gi|579676
3 EMO	von Willebrand factor preproprotein [Homo sapiens]	5.30	782	322429	19	gi|89191868
4 EMO	coagulation factor VIII, procoagulant component (hemophilia A), isoform CRA_b [Homo sapiens]	6.51	73	239233	4	gi|119593052
4 EMO	alpha-1-microglobulin; bikunin [Homo sapiens]	5.58	154	17215	3	gi|579676
5 EMO	kininogen 1 isoform 2 [Homo sapiens]	6.29	486	48936	16	gi|4504893
6 EMO	IGHM protein [Homo sapiens]	5.82	292	66032	7	gi|12804853
6 EMO	alpha-fibrinogen precursor	8.26	250	70223	7	gi|182424
6 EMO	B domain-deleted coagulation factor VIII [synthetic construct]	6.36	600	168754	16	gi|157863686
7 EMO	alpha-fibrinogen precursor	8.26	1166	70223	39	gi|182424
8 EMO	beta-fibrinogen precursor	8.31	209	55545	6	gi|182430
9 EMO	fibrinogen, beta chain preproprotein [Homo sapiens]		1311	56577	53	gi|70906435
10 EMO	fibrinogen gamma chain [Homo sapiens]	5.61	480	50077	12	gi|182439
11 EMO	fibrinogen gamma chain [Homo sapiens]	5.61	657	50077	17	gi|182439
11 EMO	Chain B, Structure Of Complement C3b: Insights Into Complement Activation And Regulation	5.18	325	104912	8	gi|118137965

Table III.: Identification trought HPLC-ESI-MS-MS of band obtained from monodimensional gel of plasma derived factor VIII Beriate® and Emoclot®

3.4 Discussion

From our analysis, it came out that both Beriate® and Emoclot® contained a large amount of vWF, as well as of other plasma proteins such as fibrinogen, fibronectin, and IgM heavy chain. This is not surprising, since the production of pdFVIII concentrates requires large plasma volumes and includes several consecutive concentration and purification steps using cryoprecipitation techniques, and chromatographic steps, which reduce but not eliminate all components of plasma. Several in vitro studies have indicated that the presence of excess of these plasma proteins, in addition to vWF, could influence the function of cellular components of the immune system (Mannucci 1996). In particular, in agreement with Josic (2006) in Beriate® and Emoclot® we have isolated IaIp as a by-product of clotting of FVIII from human plasma. IaIp are a family of structurally-related serine protease inhibitors which are found in relatively high concentrations in human plasma. Recent studies have implicated a role for IaIp in sepsis, and have demonstrated their potential as biomarkers in sepsis and cancer, as IaIp likely exert an important role in tumour invasion and metastasis and in inflammation (Kobayashi 1991). Recently, a significant decrease of IaIp levels has been shown in plasma of adult patients and new borns with clinically-proven sepsis (Lim 2003; Baek 2003).Moreover, experimental animal studies have demonstrated the beneficial effects of IaIp in improving survival and halting the progression of sepsis in adult animals (Yang 2002; Wu 2004). These results highlight the potential of IaIp as a diagnostic and therapeutic agent (Lim 2003; Yang 2002; Wu 2004; Himmelfarb 2004; Wang 2004; Zhang 2004) . However in this specific case, during pdFVIII purification, they could represent a protein burden to the patient and, in improved pdFVIII preparations, protein contaminations should be reduced as they may cause an immunogenic response, (Burnouf 1992) as reported also by Josic 2006 . Hereby we showed that pdFVIII concentrates contain a significant amount of foreign proteins which, have been shown to be related to severe side reaction sin the recipient (Josic 1999). In Beriate®, we found G3BP which (Blostein 2007) claimed to be a potential contaminant of recombinantly produced factor IX . G3BP is a secreted glycoprotein present in the extracellular matrix of many tissues. Galectins and their binding proteins have primarily been described in cell-cell and cell-matrix interactions involved in autoimmunity, inflammation and cancer biology (Barondes 1994; Perillo 1998).

In the present study we found truncated form of pdFVIII in Beriate® because of proteolytic process that could be occurred during purification procedure Josic 1999 .FVIII is sensitive to proteolysis (Stadler 2006) and, therefore, it has to be stabilized during the production process

and in the final formulation. Besides proteolysis, most serious complication is the development of inhibitor antibodies most frequently at an early stage of therapy, caused from truncated form of FVIII. These antibodies are capable of blocking FVIII pro coagulant activity (Kasper 2002; Gomperts 2006). There have been dangerous outbreaks of inhibitors in multi transfused patients in the past, and they seem to be due to the creation of neo epitopes in the FVIII molecule during the manufacturing process (Josic 1999; Peerlinck 1997). In the present study, G3BP appears in the same spot of pdFVII due to the similar apparent molecular weight, G3BP plays a role in cell apoptosis and have relevant immune modulatory activities given its role in autoimmunity, tumour metastasis and inflammation. This may be especially relevant in haemophiliacs as these patients are often chronically infected with hepatitis C and/or HIV and maybe immune compromised. Elevated G3BP levels in AIDS patients seem to correlate with higher viral loads, especially in haemophiliac HIV-positive patients (Cherayil 1989; Diaz 2009; Groschel 2000) . Therefore, the presence of G3BP may have untoward effects on the health of these patients. In conclusion most pdFVIII concentrates contain excess of plasma proteins which could influence the function of cellular components of the immune system and other contaminants, whose broad biological effects could result in unforeseen consequences in the recipients when infused for therapeutic purposes

3.5 References:

Baek Y.W., Brokat S., Padbury J.F., Pinar H., Hixson D.C., Lim Y.P. (2003) Inter-alpha Inhibitor Proteins in infants and Decreased Levels in Neonatal Sepsis. J Pediatr 143: 11-5.

Barondes S.H., Cooper D.N., Gitt M.A. Barondes S.H., Castronovo V., Cooper D.N., Cummings R.D., Drickamer K., Feizi T., Gitt M.A., Hirabayashi J., Hughes C., Kasai K. (1994) Galectins. Structure and function of a large family of animal lectins. J Biol Chem 269: 20807-20810.

Boekhorst J., Lari G.R., D'Oiron R., Costa JM, Nováková IR, Ala FA, Lavergne JM, VAN Heerde W.L. (2008) Factor VIII genotype and inhibitor development in patients with haemophilia A: highest risk in patients with splice site mutations. Haemophilia.14(4):729–735.

Blostein M., Cuerquis J., Galipeau J. (2007) Galectin 3-binding protein is a potential contaminant of recombinantly produced factor IX. Haemophilia.13:701-706.

Burnouf T. (1992) Safety aspects in the manufacturing of plasma-derived coagulation factor concentrates. Biologicals 20: 91-100.

Cherayil B.J., Weiner S.J., Pillai S. (1989) The Mac-2 antigen is a galactose-specific lectin that binds IgE. J Exp Med. 170: 1959-1972.

Diaz J.A., Ramacciotti E., Wakefield T.W. (2010) Do galectins play a role in venous thrombosis? A review. Throm Res. 125:373-376.

Gomperts E.D. (2006) The need for previously untreated patient population studies in understanding the development of factor VIII inhibitors. Haemophilia 12: 573-578

Groschel B., Braner J.J., Funk M., Linde R., Doerr H.W., Cinatl J. Jr, Iacobelli S. (2000) Elevated plasma levels of 90 K (Mac-2 BP) immunostimulatory glycoprotein in HIV-1-infected children. J Clin Immunol 20: 117-122.

Himmelfarb M., Klopocki E., Grube S., Staub E., Klaman I., Hinzmann B., Kristiansen G., Rosenthal A., Dürst M., Dahl E. (2004) ITIH5, a novel member of the inter-alpha-trypsin inhibitor heavy chain family is downregulated in breast cancer. Cancer Lett; 204: 69-77.

Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–686.

Lim Y.P., Bendelja K., Opal S.M., Siryaporn E., Hixson D.C., Palardy J.E. (2003) Correlation between mortality and the levels of inter-alpha inhibitors in the plasma of patients with severe sepsis. J Infectì Dis 188: 919-926.

Kasper C.K. (2002) Concentrate safety and efficacy. Haemophilia; 8: 161-165.

Kobayashi H. Shinohara H, Oi H.Sugimura M, Terao T, Fujie M. (1994) Urinary trypsin inhibitor (UTI) and fragments derived from UTI by limted proteolysts efficiently inhibit tumor cell invasion. Clin Exp Metastasis 12: 117-128.

Josić D., Buchacher A., Kannicht C., Lim Y.P., Löster K., Pock K., Robinson S., Schwinn H., Stadler M. (1999) Degradation Products of Factor VIII Which Can Lead to Increased Immunogenicity. Vox Sanguinis, 77: 90-99.

Josic D., Brown M.K., Huang F., Lim Y.P., Rucevic M., Clifton J.G., Hixson D.C. (2006) Proteomic characterization of inter-alpha inhibitor proteins from human plasma. Proteomics 2006; 6: 2874-2885.

Mannucci P.M. (1996) The choice of plasma-derived clotting factor concentrates. Bailliere's Clinical Haematology 9: 273-290

Oldenburg J., Pavlova A. (2006) Genetic risk factors for inhibitors to factors VIII and IX. Haemophilia;12 Suppl 6:15–22.

Peerlinck K., Arnout J., Di Giambattista M., Gilles J.G., Laub R., Jacquemin M., Saint-Remy J.M., Vermylen J. (1997) Factor VIII inhibitors in previouslytreated hemophilia A patients with a double virus-inactivated plasma derived factor VIII concentrate. Thromb Haemost; 77: 80-86.

Perillo N.L., Marcus M.E., Baum L.G. (1998) Galectins: versatile modulators of cell adhesion, cell proliferation, and cell death. J Mol Med 76: 402-412.

Shevchenko A., Wilm M., Vorm O., Mann M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem 68: 850-858

Stadler M., Gruber G., Kannicht C., Biesert L., Radomski K.U., Suhartono H., Pock K., Neisser-Svae A., Weinberger J., Römisch J., Svae T.E. (2006) Characterisation of a novel high-purity, double virus inactivated von Willebrand Factor and Factor VIII concentrate (Wilate®). Biologicals 34 : 281-288.

Wang Z., Yip C., Ying Y., Wang J., Meng X.Y., Lomas L., Yip T.T., Fung E.T. (2004) Mass spectrometric analysis of protein markers for ovarian cancer. Clin Chem 50: 1939-1942.

Wu R., Cui X., Lim Y.P., Bendelja K., Zhou M., Simms H.H., Wang P. (2004) Delayed administration of human inter-alpha inhibitor proteins reduces mortality in sepsis. Crit Care Med 2004; 32: 1747-1752.

Yang S., Lim Y.P., Zhou M., Salvemini P., Schwinn H., Josic D., Koo D.J., Chaudry I.H., Wang P. (2002) Administration of human inter-alpha inhibitors maintains hemodynamic stability and improves survival during sepsis. Crit Care Med 30: 617-622.

Zhang Z., Bast R.C., Yinhua Y. (2004) Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. Chan1 Cancer Research 64: 5882-5890.

Chapter 4

Metabolomics: a new Rapid Resolution Reverse-Phase HPLC strategy to investigate various metabolic species in different biological models

4.1 Introduction: Metabolomics

Metabolomics is the study of metabolites and their role in various physiological states; it is a novel methodology arising from the post-genomics era and has extensive biomedical application. It can be expressed as "the quantitative measurement of the dynamic multiparametric metabolic response of living systems to patho-physiological stimuli or genetic modifications " (Nicholson 1999). The beginning of metabolomics traces back all the way to 2000-1500 B.C. when traditional Chinese doctors began using ants in order to evaluate the urine of patients to determine if the urine contained the high glucose of diabetics. At this time, others tasted the urine for sweetness in order to check for the same thing. Urine was also a factor in determining diabetes in Ancient Egypt where it was determined by frequent urination. In 1905 J.J. Thomson of the University of Cambridge developed the first mass spectrometer. Also in this year there was more work in determining what other things were in urine and Otto Knut Olof Folin reported that methods for analysis of urine for urea, ammonia, creatine, uric acid. His findings were all published in one issue of Physical Review. The next step in the path to modern metabolomics came by 1946 when Felix Botch of Stanford and Edward Purcell of Harvard simultaneously published the first NMR in the same issue of Physical Review. The separation of metabolites through chromatography also made the study of metabolomics possible. As chromatographic separations were discovered and made possible in the 1960's the ability to study individual metabolites was made possible with the necessary instruments in place there was a small gap of time until 1971 when Mamer and Horning performed the first mass-based metabolomic experiments. Shortly after they began their work Modern Metabolomics began to form when Arthur B Robinson and Linus Pauling investigated biological variability being explained by ranges of nutritional requirements. By studying early chromotagraphic separations in urine he found that the chemical constituents of the urine were loaded with useful information. The first paper on Metabolomics, though not called metabolomics at the time, was by Robinson and Pauling in 1971. It was titled "Quantitative Analysis of Urine Vapor and Breath by Gas-Liquid Partition Chromatography" and was published in Proceedings of the National Academy of Sciences. Robinson went on to publish more papers, and along with colleagues they identified diseases, conditions, and physiological age based on the data that they found. This research was another ground

breaking finding, and it led the way for a new discovery in 1990, when hydrophilic interaction chromatography was introduced for the separation of peptides, nucleic acids, and other polar compounds, which was then used in the research of metabolism as well.

4.1.2 Coining the Term

At this point there was a good basis of what metabolomics would become, but it still was not called metabolomics. In 1998 the term metabolomics finally came to be when it was used by S.G. Oliver and his colleagues in their published literature in Trends in Biotechnology Oliver, S. G (1998). Modern metabolomics began about ten years ago and yet many continue to question the relative performance of this area of technology in advancing biology.

Metabolome	Metabolome refers to the complete set of small-molecule metabolites (such as metabolic intermediates, hormones and other signaling molecules, and secondary metabolites) to be found within a biological sample, such as a single organism. The word was coined in analogy with transcriptomics and proteomics; like the transcriptome and the proteome, the metabolome is dynamic, changing from second to second. Although the metabolome can be defined readily enough, it is not currently possible to analyze the entire range of metabolites by a single analytical method.
Metabolites	Metabolites are the intermediates and products of metabolism. Within the context of metabolomics, a metabolite is usually defined as any molecule less than 1 kDa in size 10.
Metabolomics	Metabolomics is defined as "the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification". The word origin is from the Greek meta meaning change and nomos meaning a rule set or set or laws.

Table IV: concept of metabolome metabolites and metabolomics**.**

Metabolomics, the comprehensive analysis of metabolites present in a biological sample, has emerged as the third major path of functional genomics besides transcriptomics (mRNA profiling) and proteomics (Fiehn 2002). Information obtained from the metabolome is particularly useful given that small molecules represent the downstream outcome of cellular machinery and can provide a metabolic phenotype of a biological system. Another advantage of studying the metabolome (versus the proteome or genome) is that the analysis can be performed independently of a genome sequence or large expressed sequence tag databases and therefore can be applied to virtually any biological system. Combined with the information obtained on transcriptome and proteome, this would lead to a nearly complete molecular picture of the state of a particular biological system at a given time. Moreover, metabolomics is being used to uncover the biochemical bases of cells quality traits, and how genetic and environmental factors biochemically influence these traits. Dramatic technological advances in the biological sciences over the past few years have forged a new era of research including the emerging field of systems biology. Although the understanding of living organisms at the molecular system level is still in its infancy, it is evident that comprehensive investigations of the ‘‘omics cascade’’ with genomics, transcriptomics, proteomics, and metabolomics are important building blocks and will play a central role in this new science (Fig 9) In the future combination with genomics, transcriptomics and proteomics, metabonomic analysis is being increasingly used in the discovery and development of new medicines.

4.1.3 Metabolomic’s approaches: target analisys and metabolomic profiling.

The genome, transcriptome, and proteome elucidations are based on target chemical analyses of biopolymers composed of 4 different nucleotides (genome and transcriptome) or 22 aminoacids (proteome). Those compounds are highly similar chemically, and facilitate high-throughput analytical approaches. Within the metabolome, there is, however, a large variance in chemical structures and properties. Thus, the metabolome consists of extremely diverse chemical compounds from ionicinorganic species to hydrophilic carbohydrates, volatile alcohol sand ketones, amino and non-amino organic acids, hydrophobic lipids, and complex natural products. That complexity makes it virtually impossible to simultaneously determine the complete metabolome. Therefore, the metabolome has been studied with efficient sample

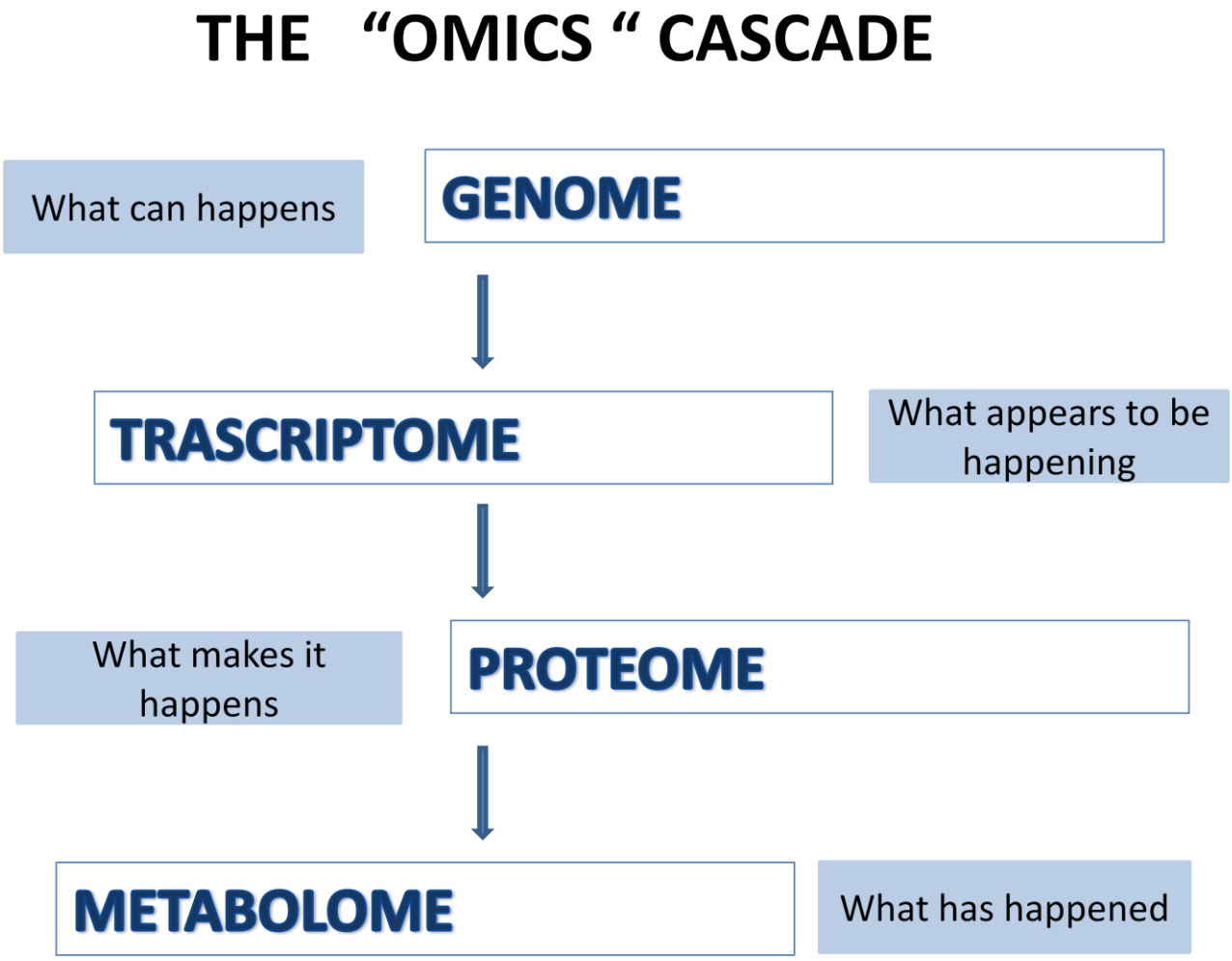


Figure 9: The omics cascade. Complementary to genomics or proteomics, the aim of metabolome analysis is to describe qualitatively and quantitatively the final products of cellular regulatory pathways and can be seen as the ultimate response of a biological system to genetic factors and/or environmental changes.

preparation and with selective extraction coupled to a combination of different analytical techniques to achieve as much information as possible. Fiehn (2002) defined different approaches for analysis of the metabolome: target analysis, metabolite profiling, metabolomics, and metabolic fingerprinting. (Fig. 10) However from the methodological point of view, there are basically only two different strategies: a) target analysis and b) metabolite profiling. Targeted analysis is very useful for the study of the primary effect of a genetic alteration, and the analytical procedures must include identification and absolute quantification of the selected metabolites in the sample. Metabolite profiling on the other hand, involves rapid analysis, often not quantitative, of a large number of different metabolites with the objective to identify a specific metabolite profile that characterizes a

given sample. This approach can be sub-divided into (a) metabolic fingerprinting and (b) metabolic footprinting. Fingerprinting covers the scanning of a large number of intracellular metabolites detected by a selected analytical technique or by a combination of different techniques in a defined situation metabolic footprinting is a more recently proposed approach (Allen 2003), which is technically similar to fingerprinting, but is focused on the measurement of all extracellular metabolites present in a spent culture medium. The compounds determined are metabolites secreted by the cells into the medium and the medium components biochemically transformed by the organism. Both approaches of metabolite profiling can be used to distinguish between different physiological states of wild-type strains, In a targeted metabolomics strategy, predefined metabolite-specific signals (by selected reaction monitoring (SRM) by tandem mass spectrometry (MS/MS), or selected ion recording (SIR) by GC-MS) are often used to determine precisely and accurately relative abundancies and concentrations of a limited number of preknown and expected endogenous metabolites. In the latter case, it is implicit that only limited number of known and expected endogenous metabolitescould be investigated, although it is possible to precisely and accurately determine their relative abundances and concentrations. Non targeted metabolomics is used for global metabolome analysis, that is, comprehensive analysis of all the measurable analytes in a sample (including analyte identification of unknown signals) and must be coupledt o chemometric methods to compress the data into a small set of signals,

4.1.4 Technical evolution of metabolomics: from NMR to MS

There are several analytical strategies that can be used to analyse the metabolome , such as nuclear magnetic resonance (NMR), fourier transformation infrared spectroscopy (FT-IR) , and mass spectrometry (MS) coupled to separation techniques such as high performance liquid chromatography (HPLC), gas chromatography (GC), or capillary electrophoresis (CE). Mass spectrometry (MS) and NMR are the most frequently employed methods of detection in the analysis of the metabolome. NMR in particular, is very useful for structure characterization of unknown compounds, and has been applied to the analysis of metabolites

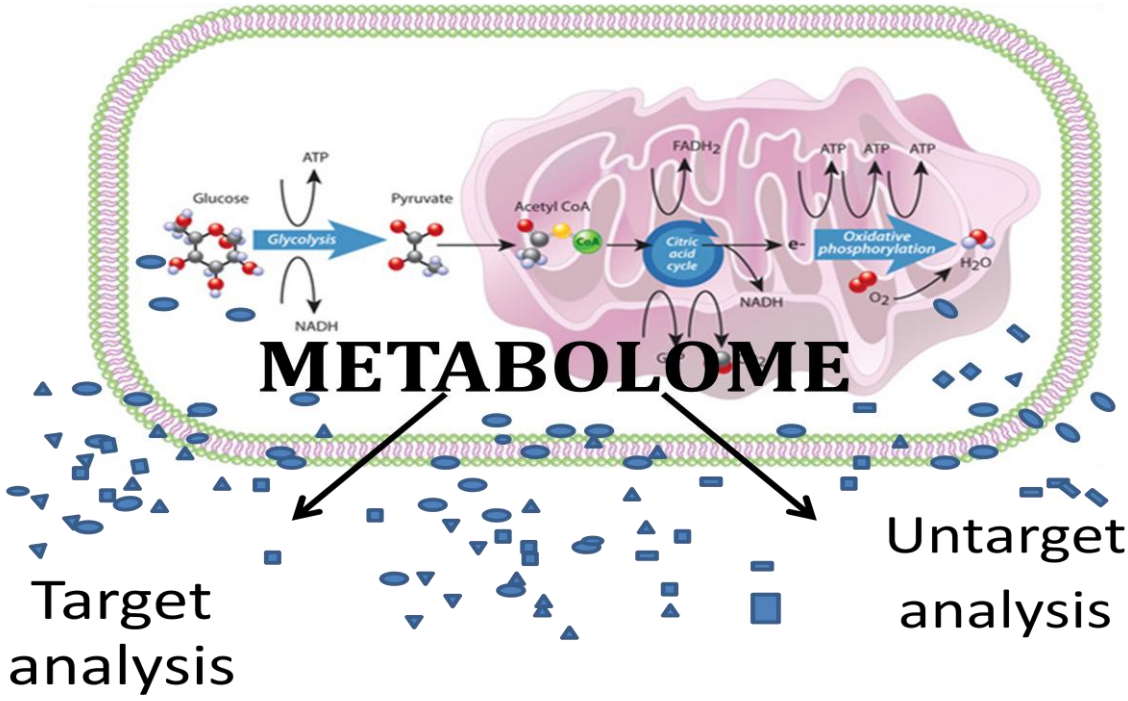


Figure 10: Metabolomics approachs consist in target analysis and untarget analysis

in biological fluids and cells extracts (Shockcor 1996). NMR was favoured by machine accessibility, established data handling, and the nondestructive nature of the analysis. In certain circumstances, the 1H NMR spectrum is insufficient on its own to provide information that will fully characterize a metabolite. This limitation is obviously the case where analytes contain functional groups that are deficient in protons or where the protons can readily chemically exchange with the solvent. However, one of the main limitations of NMR is the poor sensitivity. (Fig 11) Nonetheless, MS has gradually replaced NMR due to the higher sensitivity, that improved metabolite discrimination, coverage of the metabolome space, and modularity to perform compound-class-specific analyses. The most important advantages of MS are its high sensitivity, and high-throughput in combination with the possibility to confirm the identity of the components present in the complex biological samples as well as the detection and in most of the cases, the identification of unknown and unexpected compounds. Modern MS provides highly specific chemical information that is directly related to the chemical structure, such as accurate mass, isotope distribution pattern for elemental formula determination, and characteristic fragment-ions for structural elucidation or identification via spectral matching to authentic compound data.

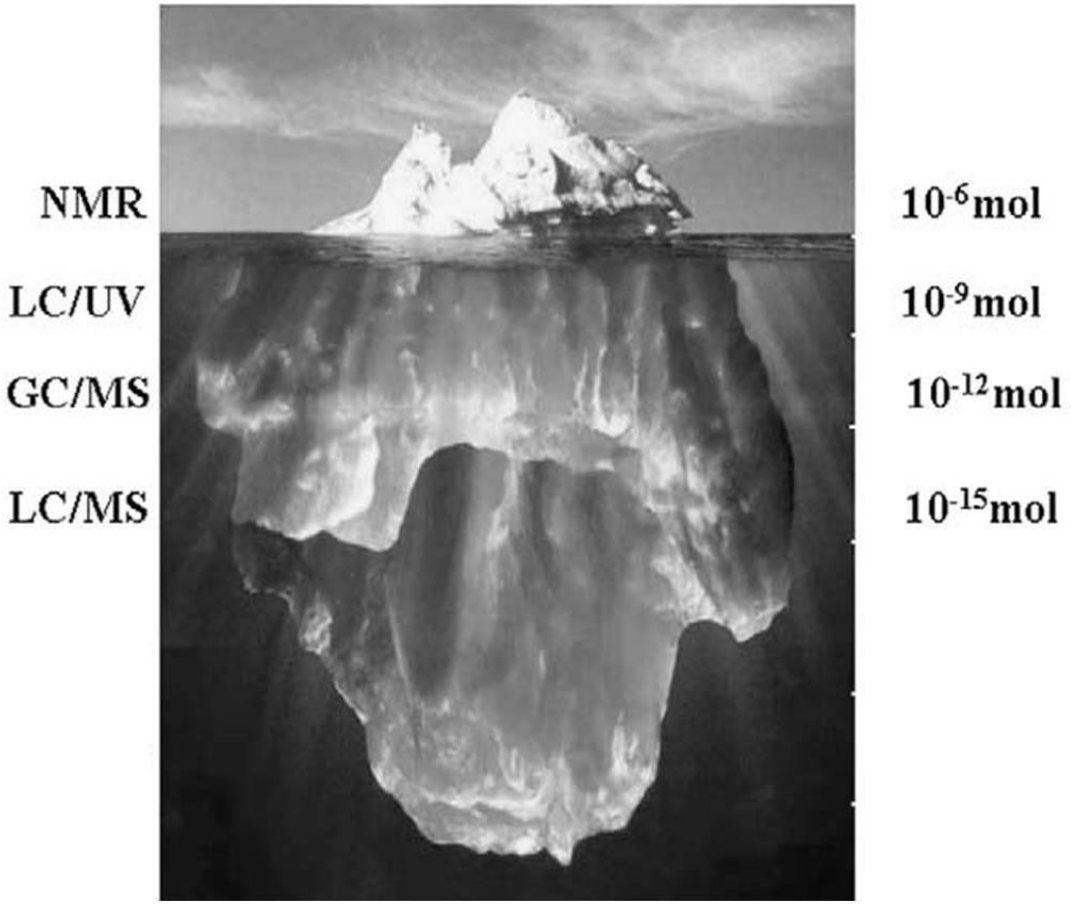


Figure 11 : a comparison of the relative sensitivities of various metabolomic tools.NMR has rapid analysis times but suffers from lower sensitivity thus allowing visualization only of the more concentrated metabolites (i.e. the tip of the iceberg). GC/MS and HPLC/MS provide good selectivity and sensitivity.

Moreover, the high sensitivity of MS allows detection and measurement of picomole to femtomole levels of many primary and secondary metabolites. These unique advantages make MS an important tool in metabolomics. Furthermore, the combination of separation techniques (e.g., chromatography) with MS tremendously expands the capability of the chemical analysis of highly complex biological samples. The basic information of mass spectra is characterized by its simplicity. The spectrum displays masses of the ionized molecule and its fragments, and those masses are simply the sums of the masses of the component atoms. In some cases, a mass spectrum contains a wealth of specific analytical and structural information, much more information than the expert in the field currently can utilize.

	NMR	MS
Detection limits	micromolare	picomolare
Universality of metabolite detection	If metabolite contains hydrogen can be detected	Problem with chromatography separation with or with ion suppression
Amount of sample used	Typically 200 - 400 ul	Low microlitres range
Sample recovery	Technique is not destructive	Technique is destructive
Analitical reproducibility	Very high	Fair
Sensitivity	Fair	Very high
Price	Very expensive	Not very expensive
Time to collect data	5 min	30 min
Avalaibility of database	Not yet comprehensive	Comprehensive database

Table V : advantages, and disadvantagesof two most important techniques used in metabolomics.

4.1.5 Mass spectrometry

Mass spectrometry is a well-known analytical tool for measuring the molecular weight of a sample or distinguishing molecules by their mass-to-charge ratios (Feng 2008). Mass spectrometers, that are routinely used for purposes such as drug discovery, diagnostics and bio-analyses, can measure the mass of a molecule only after it converts the molecule to a gas-phase ion. To do so, imparts an electrical charge to 17 molecules and converts the resultant

flux of electrically charged ions into a proportional electrical current that a data system then reads. The data system converts the current to digital information, displaying a mass spectrum (Feng 2008). Mass spectrometers are generally composed of three fundamental parts, namely the ionization sources, the mass analyzer and the detector (Fig.11). Ionization source is the part of mass spectrometer that ionizes the target materials. The ionization methods employed by the source usually determine what types of samples that can be analyzed. Ionization of the sample molecules were largely produced by electron impact generating radical ions and more could easily break covalent bonds in the molecules generating extensive and reproducible fragmentation. As MS technology advanced, "soft" ionization methods started to emerge especially the electrospray ionization (ESI) (Fenn 1989) and matrix assisted laser desorption/ionization (MALDI) (Karas 1988; Tanaka 1988), which revolutionized MS techniques. In both methods, sample molecules are ionized with minimal fragmentation. In MALDI, the analyte is mixed with a solution of matrix which strongly absorbs a UV light and allowed to dry and crystallize.

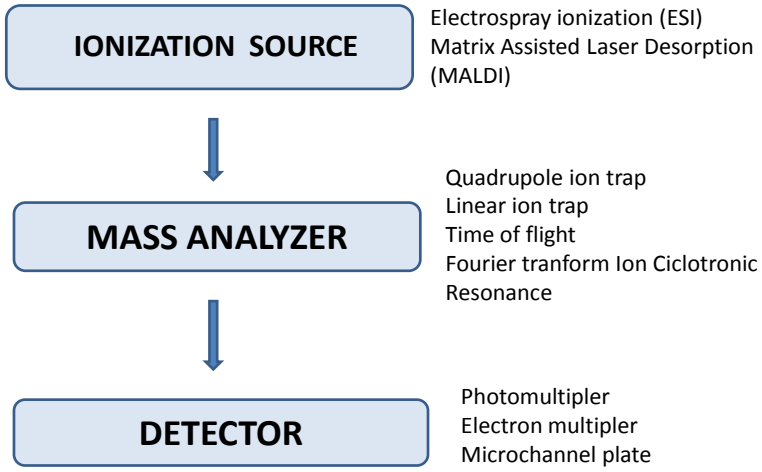


Figure 12: General components of mass spectrometers

The matrix compound is excited via absorption of a pulsed UV laser and imparts the energy to the analyte, resulting in desorption and ionization. However, increased laser fluence often results in deposition of "excessive" energy, leading to analyte ion fragmentation. The extent of the fragmentation can often be controlled by modulating the power of the laser beam, allowing this phenomenon to be used analytically as a means of producing structurally

diagnostic fragment ions. Generally, MALDI surpasses ESI in terms of sensitivity and is more tolerant to salts. Superior sensitivity, relative simplicity of operation and ease of automation made it a top choice as an analytical technique for a variety of proteomics-related applications (Tanaka 1987 ; Karas 1988). ESI, that is considered to be the "softest" ionization methods, uses electrical energy to assist the transfer of ions from solution into the gaseous phase before they are subjected to mass spectrometric analysis. The widespread use of LC-MS for global metabolic profiling is relatively recent, but there has been a rapid and continuing increase in the number of publications based onthis approach (Fig 13,14)

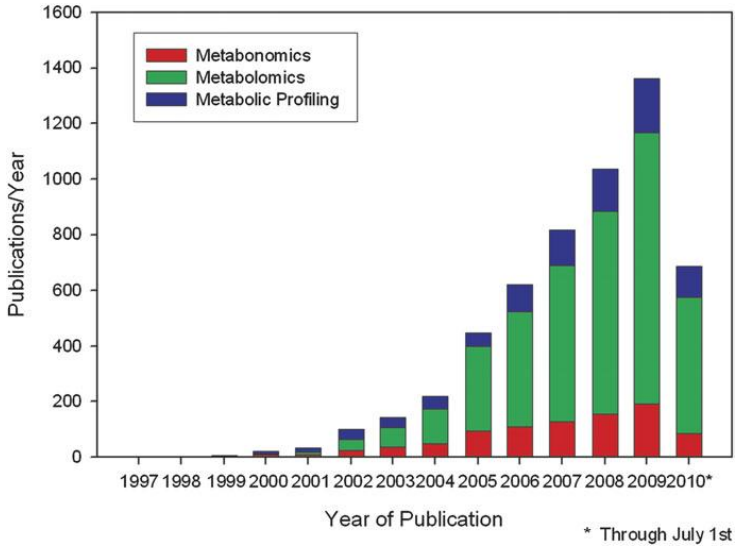


Figure 13: Number of publications/year including metabolomics, metabonomics, or metabolic profiling in the title, keywords, or abstract as determined by Scopus Note that papers using all three terms would be counted in each segment so the total number of publications may be actually somewhat lower than indicated,

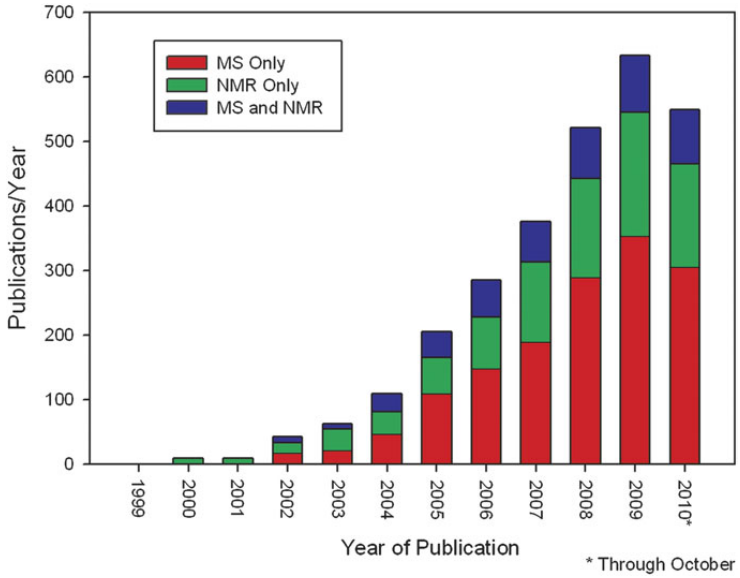


Figure 14: number of publications/year containing metabonomics or metabolomics in the title, abstract, or keywords that also contain MS exclusively (red); NMR exclusively (green), or both (blue).

4.1.6 Work flow of metabolomics.

The application of MS for analysis of cellular metabolites has grown dramatically over the last two decades, and today MS is the single most important detector method in biotechnology. Especially the introduction of ESI techniques in the late 80s has revolutionized bio analysis of small molecules. This techniques offer an excellent combination of sensitivity and selectivity, and they have become an indispensable platform in biology and metabolomics. Chromatography coupled to mass spectrometry coupling chromatography to MS offers an excellent solution to complex mixture analyses and has been extensively used in metabolomics. Chromatographic separation of metabolites prior to MS analyses has several advantages: (1) reduction of matrix effects and ionization suppression, (2) separation of isomers, (3) provides additional and orthogonal data (i.e., retention time/factor/index) valuable for metabolite annotation, and (4) allows for more accurate quantification of individual metabolites. Currently, three predominant chromatographic techniques have been incorporated in MS based metabolomics . i.e., gas chromatography (GC), liquid chromatography (LC) and capillary electrophoresis (CE). Multidimensional separation techniques such as two-dimensional GC and LC have further enabled the separation of even more complex biological mixtures, but are less widely employed. LC-MS is an important tool in metabolomics and can be tailored for targeted or non-targeted metabolomics. While both normal phase (NP) and reversed phase (RP) columns have been employed in metabolomics, RP columns such as C18 and C8 are by far the most utilized. Basing on these considerations, we present a robust, rapid and simple targeted metabolomic method, which exploits rapid resolution reverse phase (RR-RP-HPLC)/ESI-MS, for MRM-based quantitation of glutamate. Besides, we optimize an extraction protocol and test this RR-RP-HPLC/ESI-MS analytical strategy to determine and quantify a list of compounds from the main metabolic pathways (glycolysis, Kreb's cycle, pentose phosphate pathway). To test the validity of the extraction and quantitation methods, we performed multiple analyses on different biological samples.

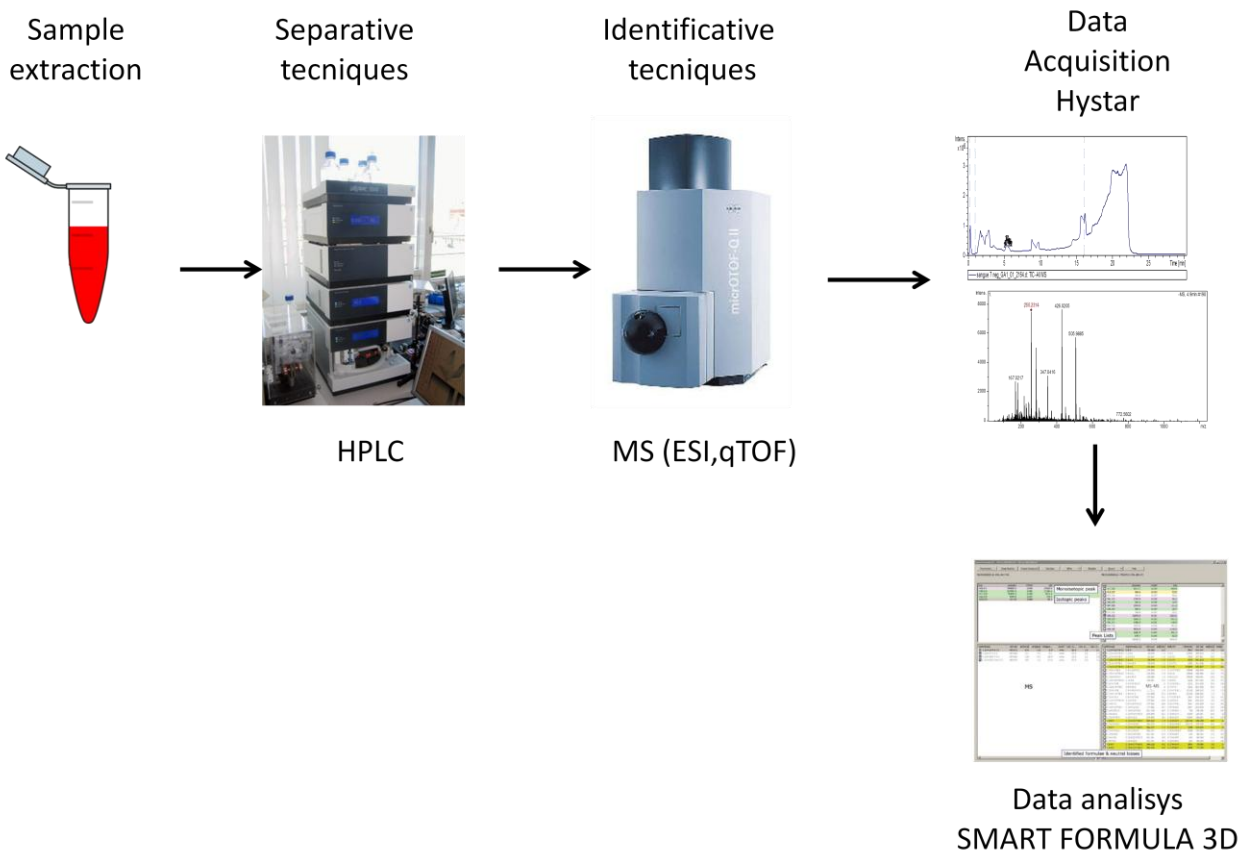


Figure 15 : Workflow used in metabolomics

4.2 Materials and methods

Acetonitrile, formic acid, and HPLC-grade water, purchased from Sigma Aldrich (Milano, Italy). Standards (equal or greater than 98% chemical purity) ATP, L-lactic acid, phosphogluconic acid, NADH, D-fructose 1,6 biphosphate, D-fructose 6-phosphate, glyceraldehyde phosphate, phosphoenolpyruvic acid, L-malic acid, L-glutamic acid, oxidized glutathione, α-ketoglutarate were purchased from Sigma Aldrich (Milan). Standards were stored either at -25°C, 4°C or room temperature, following manufacturer's instructions. Each standard compound was weighted and dissolved in nanopure water. Starting at a concentration of 1 mg/ml of the original standard solution, a dilution series of steps (in 18 MΩ, 5% formic acid) was performed for each of the standards in order to reach the limit of detection (LOD) and limit of quantification (LOQ).

4.2.1 Sample preparation

Red blood cells. Red blood cells units were drawn from healthy human volunteers according to the policy of the Italy Blood Transfusion Service for donated blood and all the volunteers

provided their informed consent in accordance with the declartion of Helsinki. . We studied RBC units collected from 4 donors [male=2, female=2, age 48±11.5 (mean ± S.D.)] in the middle region of Italy. RBC units were stored for 42 days under standard conditions (CDP-SAGM, 4°) and samples were removed aseptically for the analysis after 0, 7, 14, 21, 28 and 42 days of storage. For each sample, 0.5mL from the pooled erythrocyte stock was transferred into a microcentrifuge tube and processed for metabolite extraction. Erythrocyte samples were then centrifuged at 1000g for 2 minutes at 4°C. Tubes were then placed on ice while supernatants were carefully aspirated, paying attention not to remove any erythrocyte at the interface. Samples were further processed for metabolite extraction. SAOS cell lines. Transformed Saos-2 cells, either induced and non-induced, were kindly provided by Prof. Melino Gerry (University of "Tor Vergata ", Rome - Italy) and prepared as reported. (Gressner et al 2005) Human osteosarcoma cell line Saos-2 (p53, p63 and pRb null) was purchased from ATCC (Rockville, MD). Cells were grown in monolayer cultures in Dulbecco-MEM media supplemented with 10% heat-inactivated FBS, L-glutamine (2 mM), penicillin (100 IU/ml), and streptomycin (100 mg/ml) at 37°C in a humidified atmosphere of 5% CO2 and 95% air. The cells were split every 3 days using 0.25% (v/v) trypsin (Gibco BRL, Gaithersburg, MD) in versene buffer. Cells were routinely checked to be mycoplasma free. Saos-2 cells with doxycycline (Dox)-inducible expression of HA-TAp63a were generated as described previously (Gressner 2005). For the experiments 8 x 105 cells were plated and allowed to attach on plastic dishes (100 x 20 mm) prior to treatment. Proliferation was determined by a count of cells in a Neubauer cytometer chamber. Viability was assessed by 0.4% (w/v) trypan blue dye (Gibco BRL) exclusion. 106 cells (either non-induced controls or induced with 2.5 μg/ml doxycycline at 24h from induction) were further processed for metabolite extraction and glutamate level assessment through mass spectrometry, as described below.Blastocoel fluid. The method for blastocyst micropuncturing and aspiration of blastocoel fluid was adapted from Brison (1993) In brief, expanded day 8 blastocysts were removed from culture and transferred to a 10 nl droplet of pre-warmed Hepes-SOFaaBSA under a mineral oil underlay. The blastocysts were immobilized by a holding pipette connected to an air-filled syringe and mounted on a micromanipulator. The medium surrounding the embryo was gently removed to ensure that the site of micropuncture was not contaminated by external culture medium. A finely pulled, oil-filled pipette was introduced through the mural trophectoderm to avoid damaging the ICM cells, and blastocoel fluid was aspirated gently until the blastocyst had fully collapsed around the pipette. The retrieved fluids were expelled into a dish under oil and frozen at −80°C alongside 0.5 nl control

droplets of SOFBSA. Blastocoel fluid samples were then thawed and directly processed for metabolic extraction.

4.2.2 Sample extraction and determination of extraction efficiency

Samples were extracted following the protocol by Sana 2008., with minor modifications. A schematized version of the protocol is described in Table VI. Cell-free samples (blastocoel fluid) were treated following the same protocol, exception made for step 1 and 13, in which no cell lysates were present and thus no cells were visible at the interface.Finally, the dried samples were re-suspended in 1 mL of water, 5% formic acid and transferred to glass autosampler vials for LC/MS analysis.The efficiency of the extraction protocol was determined using L-malic acid as an internal standard in Krebs' cycle-devoid red blood cell extracts at step 4 at different concentrations (0; 1; 5; 10 mg). Detected L-malic acid concentrations were calculated upon independent determination of the L-malic acid standard curve (five points curve, three technical replicates each). In order to verify the linearity and reproducibility of the extraction method 4 technical replicates were performed for each extraction.

4.2.3 Rapid Resolution Reverse-Phase HPLC

An Ultimate 3000 Rapid Resolution HPLC system (LC Packings, DIONEX, Sunnyvale, USA) was used to perform metabolite separation. The system featured a binary pump and vacuum degasser, well-plate autosampler with a six-port micro-switching valve, a thermostated column compartment. A Dionex Acclaim RSLC 120 C18 column 2.1mm×150mm, 2.2 µm was used to separate the extracted metabolites. LC parameters: injection volume, 20 µL; column temperature, 25°C; and flowrate of 0.2 mL/min. The LC solvent gradient and timetable were identical during the whole period of the analyses. A 0–95% linear gradient of solvent A (0.1% formic acid in water) to B (0.1% formic acid in acetonitrile) was employed over 15 min followed by a solvent B hold of 2 min, returning to 100% A in 2 minutes and a 6-min post-time solvent A hold.

1	Place 0,5 mL of the erythrocyte samples in a 1,7 ml microcentrifuge tube
2	Centrifuge at 1000 x g for 2 minutes at 4 °C
3	Place the tubes on ice
4	Aspirate the supernatant being careful not to remove any eritrocytes
5	Resuspended the erythrocytes by addind 0,15 mL of ice cold pH adjusted ultra-pure water (Milli-Q) to lise cells.
6	Plunge the tubes into dry ice or a circulating bath at -25°C for 0.5 min.
7	Plunge the tubes into a water bath at 37 °C for 0.5 min.
8	Add 0.6 mL of -20°C methanol containing internal standards (malic acid 0,1,5,10 mg /ml) *
9	Vortex the tubes to ensure complete mixing
10	Transfer the tubes to room temperature
11	Add 0.45 mL chloroform to each tube
12	Vortex the tube to briefly every 5 min for 30 minutes returning the tubes to the cold to the cold bath between vortexing
13	Transfer the tubes to room temperature
14	Add 0,15 mL of ice cold pH adjusted ultra-pure water (Milli-Q) to the tubes
15	Centrifuge the tube at 1000 x g for 1 min at 4 °C
16	Transfer the tube to -20°C freezer for 2-8 h
17	Transfer the top and the bottom phases together
18	Add an equivalent volume of acetonitrile to precipitate any proteins and transfer to refrigerator (4°C) for 20 min
19	Centrifuge at 10000 x g for 10 min at 4 °C
20	Aspirate the supernatant
21	Dry the tube under vacuum
22	Resuspend the content of each tube by adding water (Milli -Q) with 0,5% formic acid

Table VI : step for metabolites extraction

4.2.4 ESI Mass Spectrometry

Metabolites were directly eluted into a High Capacity ion Trap HCTplus (Bruker-Daltonik, Bremen, Germany). Mass spectra for metabolite extracted samples were acquired in positive ion mode. ESI capillary voltage was set at 3000V (+) ion mode. The liquid nebulizer was set to 30 psi and the nitrogen drying gas was set to a flow rate of 9 L/min. Dry gas temperature was maintained at 300°C. Data was stored in centroid mode. Internal reference ions were used

to continuously maintain mass accuracy. Data was acquired at a rate of 5 spectra/s with a stored mass range of *m/z* 50–1500. Data were collected using Bruker Esquire Control (v. 5.3 – build 11) data acquisition software. In MRM analysis, m/z of interest were isolated, fragmented and monitored (either the parental and fragment ions) throughout the whole RT range. Validation of HPLC on-line MS-eluted metabolites was performed by comparing transitions fingerprint, upon fragmentation and matching against the standards metabolites through direct infusion with a syringe pump syringe pump (infusion rate 4 µl/min). Standard curve calibration were performed either on precursor and fragment ion signals. Only the former were adopted for quantitation, as precursor ion signals guaranteed higher intensity and thus improved LOQ and LOD.

4.2.5 Data elaboration and statistical analysis

LC/MS data files were processed by Bruker DataAnalysis 4.0 (build 234) software. Files from each run were either analyzed as .d files or exported as mzXML files, to be further elaborated for spectra alignment, peak picking and quantitation with In Silicos Viewer 1.5.4 (Insilicos LLC; Seattle, USA). For Total Ion Current (TIC) analyses, all compounds and compound-related components (i.e. features) in a spectrum were considered for quantitation. In positive-ion mode this included adducts (H+, Na+ and K+), isotopes and dimers. These related ions were treated as a single compound or feature for preliminary quali-quantitative analysis of metabolites of interest. Quantitative analyses of standard compounds were performed on SRM data. Each standard metabolite was run in triplicate, at incremental dilution until LOD and LOQ were reached. The limit of detection for each compound was calculated as the minimum amount injected which gave a detector response higher than three times the signal-to-noise ratio (S/N). To evaluate the potential of the method for quantitative analysis of selected metabolites, intra- and inter-day repeatability of retention times, and linearity of the RR-RP-HPLC-ESI-MS method were tested. Intra-day repeatability was measured by injecting the same standard solution (glutamic acid 1 µg/ml) three times in a single day. Inter-day repeatability was measured by analysing the same standard solution over 5 different days. Five-point standard curves were established by plotting integrated peak areas versus concentration. Each point on the calibration curve is the mean value of three independent measurements using the RR-RP-HPLC-ESI-MS method. Linearity of the observed quantities, slope, intercept and linear correlation values were all calculated via Microsoft Excel (Microsoft, Redmond,WA, USA). Glutamate levels in SAOS-2 cell lines, prior to or at 24h from doxycycline induced expression of TAp63 were assayed in five independent runs each

and quantified according to calibration curves obtained as reported above. Data were further refined and plotted with GraphPad Prism 5.0 (GraphPad Software Inc.)

4.3 Results

4.3.1 Extraction efficiency and technical reproducibility

An efficient and effective metabolomic analytical strategy should primarily take into account a series of crucial parameters, including extraction efficiency and technical reproducibility of the extraction protocol. In order to assess the efficiency of the extraction protocol and thus provide a reliable quantification of the monitored metabolites, we introduced L-malic acid at different concentrations (0; 1; 5; 10 mg), as internal standard in red blood cell extracts at step 4 of the extraction protocol (Table VI). Four technical replicates were performed in order to assess the reproducibility of the extraction process. Because of the lack of nuclei and mitochondria, mature red blood cells are incapable of generating energy via the (oxidative) Krebs cycle. Therefore erythrocytes mainly rely on 4 main metabolic pathways: the Embden-Meyerhof pathway (glycolysis), in which most of the RBC adenosine triphosphate (ATP) is generated through the anaerobic breakdown of glucose; the hexose monophosphate shunt (HMS), which produces NADPH to protect Red blood cells from oxidative injury; the Rapoport-Lubering shunt, responsible for the production of 2,3-diphosphoglycerate (DPG) for the control of Hb oxygen affinity; and finally, the methemoglobin (met-Hb) reduction pathway, which reduces ferric heme iron to the ferrous form to prevent Hb denaturation. (Wiback 2002; Schmaier 2003). Red blood cells do maintain a number of proteins which have been demonstrated to be potentially enzymatically active, such as malate dehydrogenase, although they represent but a functionless remainder after the dedifferentiation of reticulocytes into the mature red blood cells. (Leskovac 1975) The exogenously introduced malic acid has been thus used as an internal standard to test the efficiency of the extraction protocol and thus calculate a coefficient to derivitize the absolute concentration of the monitored metabolite in the original sample. MS-detected L-malic acid counts were used to calculate its concentration, basing on an independently calculated standard curve ($r = 0{,}999952$ – Figure 16). As a result, we confirmed the efficiency of the extraction method (99,662% ± 2,978 – mean + SD) in all the replicates at each tested concentration of the internal standard.

Malic acid internal standard during RBC extraction: extraction efficiency. * Below LOD and LOQ – minimum counts detected	
Concentration [mg/ml]	**Dilution and concentration [mg/ml] (averages)**
0	0*
1	0,947
5	4,987
10	9,997
Linearity	Y= 326299,1 X + 6713,113
r	0,999952
Extraction efficiency (±Sd)	99,662% ± 2,978

Figure 16 : Malic acid internal standard during RBC extraction: extraction efficiency

4.3.2 Simoultaneously testing of multiple metabolites from the main metabolic pathways through a single strategy on different biological samples

Targeted multiple reaction monitoring (MRM) was performed to quantify a series of metabolites involved in glycolysis, Krebs' cycle, pentose phosphate pathway, redox homeostasis and nucleotide metabolism, as reported in **Table VII**. The linearity of the RR-RP-HPLC-ESI-MS response was measured for each compound by recording the responses at different concentrations. Intra-and inter-day variability were tested with positive results (reproducibility >98%) also for MS analyses. Up to 10 metabolites were simoultaneously monitored through MRM in samples from SAOS-2 cell lines, red blood cells, diffusion medium of retina neural cells and blastocoels fluid. Reproducibility of retention times and peak elution was very high (>98%) in each tested sample. The elevated sensitivity of the method and the possibility to perform multiple metabolite qualitative and quantitative identification (Figure 17), allowed to overcome the hurdles deriving from low sample availability, as in the case of the blastocoel fluid (Figure 18).

	Metabolite	PubChem ID	Monoisotopic mass	MS/MS	Retention time (min)	Standard curves	Linearity: orders of magnitude	Linear correlation coefficient	Number in Figure 4
Glycolysis	Fructose 6 Phospate (F6P)	69507	260,0297	99	2,6	Y = 437467,7 X + 3383,886	4	0,993721	**1**
	Fructose 1,6 biphosphate (FBP)	10267	339,9960	99	2,8	Y = 139519,1 X + 6106,018	5	0,999991	**2**
	Glyceraldheyde 3 phosphate (G3PD)	729	169,9980	99	3	Y= 168431,94 X + 6446,423	5	0,996182	**3**
	Phosphoenolopyruvate (PEP)	1005	167,9824	151	2,8	Y= 1667498 X + 4095,535	5	0,999987	**4**
	Lactic acid	612	90,0317	63	11,8	Y= 563010,1X+ 1423,73	4	0,998704	**5**
Krebs	α-ketoglutaric acid	164533	144,0822	55	3	Y= 441223,6X + 4870,91	4	0,995342	**6**
	Malic acid	525	134,0215	73	8,7	Y= 326299,1 X + 6713,113	4	0,999952	**7**
Pentose phosphate pathway	6-phospgluconic acid	91493	276,0246	259	3,1	Y=792357X + 5951,28	4	0,994114	**8**
Redox defenses	Glutamic acid	611	147,0532	128	2,5	Y= 26772,17X+ 1450,912	6	0,999327	**9**
	Glutathione (oxidized)	65359	612,1520	355	5,4	Y= 44122,36 X+ 4870,91	5	0,996585	**10**
Nucleotides	ATP	5957	506,9957	410	2,9	Y= 93015,19 X +6734,24	4	0,997755	**11**
NADH	NADH	928	665,1248	524	3,8	Y= 43921X+ 1306,532	5	0,997223	**12**

Table VII : Metabolites identified through RR-RP-HPLC – MRM-ESI/MS

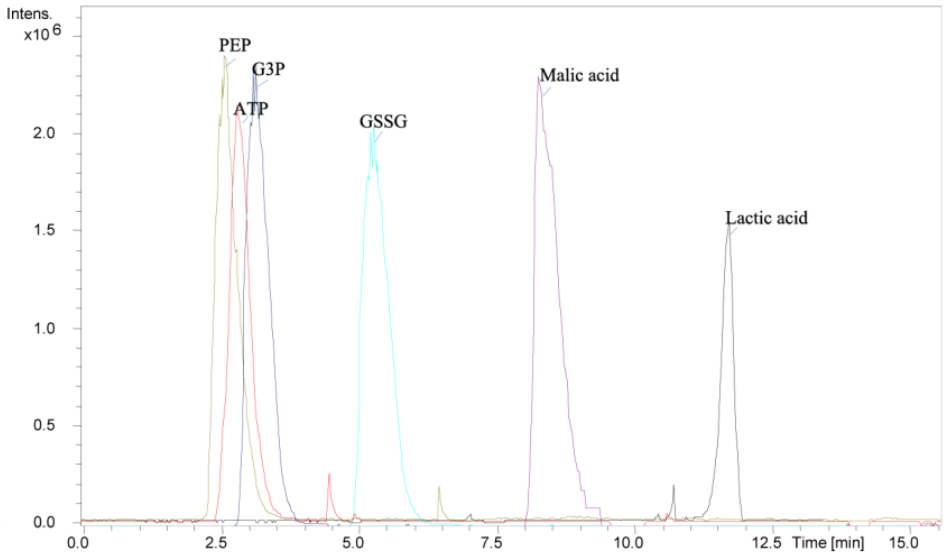


Figure 17 :Multiple reaction monitoring MRM analysis of six different (standard) metabolites including phosphoenolpyruvate (PEP), adenosine triphosphate (ATP) glyceraldehyde-3-phosphate (G3P), oxidized glutathione (GSSG), malic and lactic acids

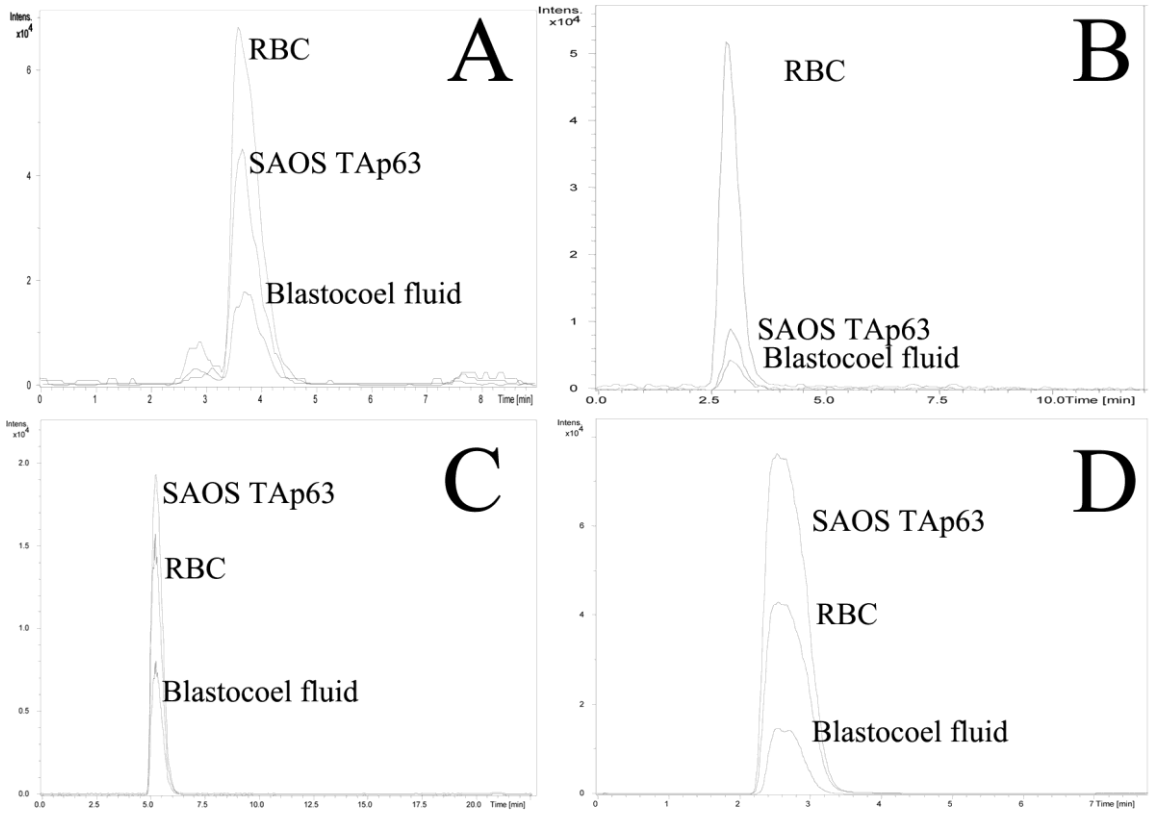


Figure 18 : MRM spectra of four representative metabolites (NADH (A); ATP (B); GSSG (C) and glyceraldehyde-3-phosphate (D)) in three differentbiological samples (red blood cells—RBC; SAOS induced to express TAp63 at 24 h from induction with doxycycline; blastocoel fluid).

4.4. Discussions

Classical work in tumor cell metabolism focused on bioenergetics, particularly enhanced glycolysis and suppressed oxidative phosphorylation (the 'Warburg effect'). (Warburg 1956) Appreciation of the generality of the Warburg effect stimulated the broader concept that a 'metabolic transformation' is required for tumorigenesis. In parallel, glutamine/glutamate metabolism can allow cells to meet both the anaplerotic and NADPH demands of growth. (Deberardinis 2008) In cells, glutamate can be further converted into α-ketoglutarate, which is an important substrate for the citric acid cycle (TCA) to produce ATP in mitochondria. Furthermore, glutamate is a precursor of reduced GSH, one of the most important antioxidant molecules and a scavenger for ROS. Alterations in glutamate levels have been observed in tumour cells upon induced expression of p53 and subsequent activation of the glutamine synthase 2 gene, an enzyme which catalyzes the hydrolysis of glutamine to glutamate. (Deberardinis 2008). Upon p53-induced glutaminase 2 level increase, either improvement of anti-oxidant stress responses and activation of pro-apoptotic cascades have been observed in tumour cells, hinting at a likely association of glutamate levels with impaired proliferation and survival capacities in tumour cells. p63 and p73 give rise to proteins that have p53-agonistic as well as p53-antagonistic functions and new functions. (Brison 1993) p53 and p63 share a conserved transactivation domain (TA), which promotes the transcription triggering activity of these proteins and hints at a likely overlap between the downstream targets of p53 and p63. (Melino 2003) Glutaminase 2 is supposed to be one of these shared downstream targets, although supporting evidence for p63 has not been hitherto provided. Glutamate levels were thus assayed in p53 and p63-null SAOS-2 cell lines transformed to express the TAp63 isoform under doxycycline induction, as an indirect marker of glutaminase 2 expression levels. Metabolite extraction was performed at day 0 and at 24h upon induction with doxycycline, resulting in a significant increase (1.5 fold change) in glutamate concentrations in the latter case These partial results are only representative of the likely applicability of the technique in this biological issue and, theoretically, for glutamate level determination in most various samples.

4.5 Conclusions

We hereby presented a rapid, efficient and robust metabolite extraction strategy in different biological samples. In parallel, we optimized RR-RP-HPLC/MS parameters in order to qualitatively and quantitatively identify several metabolite classes in MRM mode, simoultaneously. Finally, we evidenced the advantages of the present method over a previously-proposed strategy for one representative metabolite, glutamic acid, which holds relevant biological implications in most various physiological activities, including nervous system functioning and cancer cell proliferation and apoptosis. These advantages include a reduced sample handling during sample preparation, through the elimination of the chemical derivatization steps, increased sample extraction efficiency, sensitivity and linearity over an increased dynamic range of concentrations. The advantages of such analytical approach include high sensitivity and versatility for a broad range of metabolites, without the need for metabolite-specific chemical derivatization. This kind of methodology study described the feasibility and ease of a fast HPLC-MS-based metabolomic approach which enables detection, validation and quantitation of metabolites of biological interest for embryo development from blastocoele fluid down to the picomole quantities and in red blood cell during storage.

4.6 References:

Allen J.K., Davey H.M., Broadhurst Dheald l Rowland JOliver S.Kell D. (2003). Highthroughput characterisation of yeast mutants for fu4nctional genomics using metabolic footprinting. Nat. Biotechnol. 21, 692–696.

Brison D.R., Hewitson L.C., Leese H.J. (1993) Glucose, pyruvate, and lactate concentrations in the blastocoel cavity of rat and mouse embryos. Mol Reprod Dev. Jul;35(3):227-232.

Deberardinis R.J., Sayed N., Ditsworth D., Thompson C.B., (2008) Is cancer a disease of abnormal cellular metabolism? New angles on an old idea. Genet. Med. 767-777.

Feng X., Liu X., Luo Q., Liu B.F. (2008) Mass spectrometry in systems biology: an overview. Mass Spectrom 27: 636-660.

Fenn J.B., Mann M., Meng C.K., Wong S.F., Whitehouse C.M. (1989) Electrospray ionization for mass spectrometry of large molecules. Science 246: 64-71.

Fiehn, O. (2002). Metabolomics: the link between genotypes and phenotypes. Plant Mol. Biol. 48, 155–171.

Gressner O., Schilling T., Lorenz K., Schulze Schleithof E., Koch A., Schulze-Bergkamen H., Lena, E. Candi A.M., Terrinoni A., Catani M.V., Oren M., Melino G., Krammer P.H., Stremmel W., Müller M.,(2005) TAp63alpha induces apoptosis by activating signaling via death receptors and mitochondria. EMBO J., 24(13), 2458-2471.

Karas M., Hillenkamp F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal chem. 60: 2299-2301.

Melino G., Lu X., Gasco M., Crook T. Knight R.A. (2003) Functional regulation of p73 and p63: development and cancer..Trends Biochem Sci.28:663-670. Review.

Nicholson J.K., Lindon J.C., Holmes E. (1999) "Metabonomics": understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data" Xenobiotica 29:1181–1189.

Oliver S.G., Winson M.K., Kell D.B. Baganz, F. (1998). Systematic functional analysis of the yeast genome. Trends Biotechnol. 16: 373–378.

Pauling L., Robinson A.B., Teranishi R., Cary P. (1971) Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. Proc Natl Acad Sci U S A. 68(10):2374-2376.

Sana T.R., Waddell K., Fischer S.M. (2008) A sample extraction and chromatographic strategy for increasing LC/MS detection coverage of the erythrocyte metabolome. J Chromatogr B Analyt Technol Biomed Life Sci.871: 314-321.

Shockcor J.P., Unger S.E., Wilson I.D., Foxall P.J.D, Nicholson J.K., Lindon J.C. (1996) Combined HPLC, NMR spectroscopy, and ion-trap mass spectrometry with application to the detection and characterization of xenobiotic and endogenous metabolites in human urine. Anal. Chem. 68:4431–4435.

Tanaka K., Waki H., Ido Y., Akita S., Yoshida T., Matsuo T. (1988) Protein and polymer analyses up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 2: 151-153.

Warburg O. (1956) On the origin of cancer cells, Science. 123: 309–314.

Wiback S.J., Palsson B.O. (2002) Extreme pathway analysis of human red blood cell metabolism. Biophys J. 83: 808-818.

Chapter 5

A mass spectrometry-based targeted metabolomics strategy of human Blastocoele fluid: a promising tool in fertility research

5.1 Introduction

The HPLC-MS metabolic approach that we have development can be utilised for many application . In particular in the last part of my study I used this method to find a correlation between quali-quantitative profiles of small molecules of metabolic interest and the outcome of embryo transfer in blastocoels fluid for the tequi FIVET

5.1.1 FIVET : in vitro fertilization and embryo transfer

IVF (in vitro fertilization) is the most common form of ART (Assisted Reproductive Technology). It is also usually the most effective treatment for most other types of infertility as well. Once fertilised, the resulting embryos are placed back in the woman's uterus in the hope that a successful pregnancy will follow. Human *in vitro* fertilisation (IVF) protocols aim at enabling patients to achieve an acceptable singleton pregnancy rate. This is usually - achieved either by the extended culture of human embryos to the blastocyst stage (Gardner 1998; Gardner 1999) IVF is performed in treatment cycles. Pregnancy may be achieved after one treatment cycle or it may take up to four treatment cycles to achieve a successful pregnancy. Each treatment cycle of IVF consists of the following steps:

a) Evaluation for IVF - Infertility evaluation prior to IVF is done by blood tests to monitor your hormone levels and ultrasound examination to determine the optimum time to harvest your egg from the ovary. You will be administered an injection of human Chorionic Gonadotrophin (hCG) 35 hours before your egg is retrieved from the ovary. Human Chorionic Gonadotrophin (hCG) triggers final stages of ovulation.

b) Egg Retrieval – During this step in IVF procedure, the eggs from the ovary are aspirated transvaginally by an ultrasound guided probe. A fine needle pierces the ovarian follicle and the follicular fluid, which contains the egg, is extracted. This step of IVF procedure is performed under local anesthesia and intravenous sedation to help you relax and minimize discomfort during the procedure.

c) IVF Lab- The eggs are washed and then cultured in a fluid containing specialized chemicals and nutrients. A couple of hours after egg collection, the husband / partner provides a semen specimen. It is extremely important that both egg and sperm are maintained at body temperature. The embryologist adds a droplet of semen to each test tube containing the eggs and places it in the incubator. After 16–20 hours, the eggs are examined under the microscope and checked for the first signs of fertilization. A fertilized ovum looks like a small ball with two eyes under the microscope. After two to three days, if the embryos are growing normally, they are ready to be transferred to the woman's uterus.

d) Embryo Transfer of IVF (FIVET) – The embryos are placed in a tiny plastic tube which is introduced into the uterus through the cervix, the embryo(s) are transferred into the endometrial cavity (uterus). When more eggs are retrieved than needed, we can freeze them for future use (cryopreservation). Embryos can be frozen at any stage between day one and day six after egg retrieval and can be stored for years. We offer frozen embryo transfer at our offices to allow you the chance of conception when the time is right for you. During in vitro fertilization treatment (IVF), a surplus of oocytes is produced which results in a surplus of embryos which must be preserved. Due to the recent tendency in ART treatment to transfer fewer embryos in order to avoid multiple pregnancies, partly a result of new laws and guidelines aimed at controlling the number of embryos to be transferred and the number of oocytes to be fertilized, the need for simpler, more effective methods for the cryopreservation of embryos and oocytes is increasing. Currently, zygotes and embryos are typically cryopreserved by means of the traditional slow-rate freezingprotocol, while oocyte freezing is still in the experimental phase. The aim of cryopreservation is to preserve fertilized dividing embryos or unfertilized eggs (oocytes) for use in the future. This is helpful in many ways. There are those who conceive with a fresh cycle and return years later to attempt another pregnancy with their frozen embryos. Others who have not been successful in their fresh IVF cycle can try to conceive through a frozen embryo transfer (FET) cycle. This allows for a second or third attempt at conception without exposure to fertility drugs and the high cost of a fresh IVF cycle. We can also freeze a patient's eggs for future use when someone is diagnosed with cancer and treatment (cancer treatment with chemotherapy or radiation) that might affect future fertility.

5.1.2 Italian law about assisted reproduction technology

In February 2004, the Italian Parliament approved a law (namely 40/2004) regulating assisted reproduction technology (Benagiano 2010). This law imposed many limitations on Italian reproductive specialists. The most important of these restrictions was the provision that no more than three oocytes could be fertilized at one time during an IVF treatment, since all embryos obtained had to be transferred simultaneously. On May 2009, the Italian Constitutional Court outlawed some restrictions set out in the 40/2004 law (Benagiano 2010). The most important point of the ruling is that embryo protection is limited by the imperative to ensure a concrete possibility to achieve a successful pregnancy. However, the Constitutional Court, with its Decision n. 151/2009 stated that some of the provisions of Law n. 40/2004 were constitutionally illegitimate. More specifically, the Court annulled the sentence "for a single and simultaneous implantation, in any case not superior to three", thereby allowing for the possibility to also create, for each procreative cycle, more than three embryos, which not necessarily have to be implanted together. This because for some women the implantation of three embryos could be insufficient to obtain a pregnancy, whereas for others it could lead to multiple-twins pregnancies. Consequently, it will be up to the medical doctor who follows the procreation to determine how many embryos should be created. The remaining embryos are to be cryoconserved for a future implantation, if the woman's health conditions allow it.

5.1.3 Blastocyst

A blastocyst is a highly developed embryo that has divided many times to a point where it is nearly ready to implant on the walls of the uterus. A blastocyst has come a long way from its beginning as a single cell. During maturation, an embryo rests inside a protective shell called a zona pellucida. You can think of this protective shell as being much like a chicken egg. But, unlike chicken eggs, human embryos do not remain inside a shell. Instead, the embryo hatches (breaks out of the shell) on the fifth or sixth day so it can attach to the uterine wall (implantation). Just prior to hatching, an embryo becomes a blastocyst. Embryos developing to the critical blastocyst stage have a much greater chance of implanting successfully and resulting in an ongoing pregnancy. That is because these embryos have passed an important test. During the first few days, the embryo relies on the mother's egg for all its nutrients. However, in order to 15 survive past day three or four, the embryo must activate its own

genes. Not all embryos are successful. In fact, only about one-third of the embryos become blastocysts. Yet these embryos are more highly-developed, healthier, and stronger, and have a higher rate of implantation when compared to day three embryos. Due to the higher probability of survival, we transfer fewer back into the uterus.

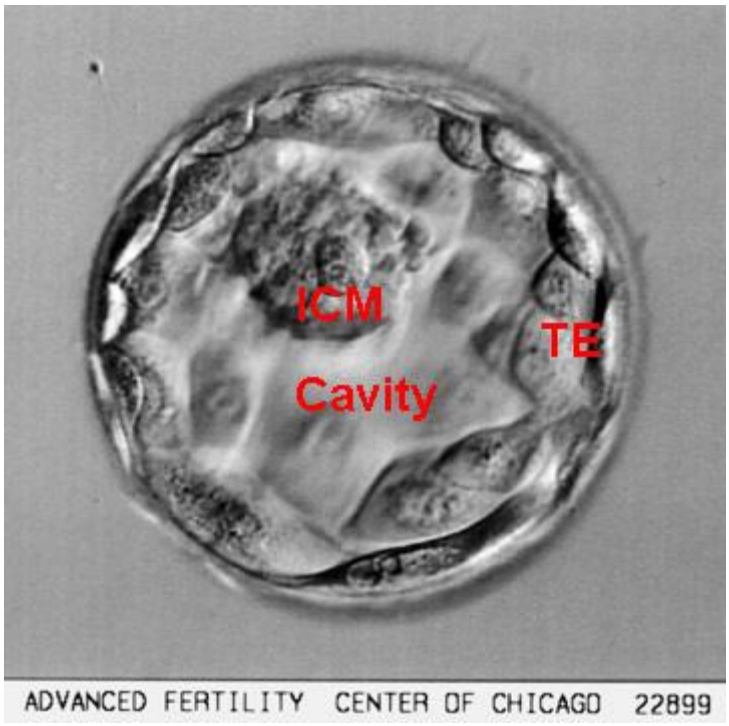


Figure 19: Picture of a high quality human blastocyst embryo 5 days after fertilization The clump of cells in the 10 to 12 o'clock area is the inner cell mass (ICM) which become the fetus. The trophectoderm cells (TE) that will form the placenta surround the fluid cavity. The fluid-filled blastocoels cavity is in the center.

5.1.4 Blastocyst grading system

The quality of the embryo, or fertilized egg, is one of a number of important components in the success of an in vitro fertilization (IVF) pregnancy. There are several types of information used to evaluate the health of embryos, including the number of cells present and the grade, or quality of the embryo. Assigning a grade is one way to identify the best quality embryos for the selection of embryo transfer during an IVF procedure. Using this information, determinations are made about how many embryos to transfer, how many to freeze, and what to do with embryos that are not developing well. Embryo grading has been proposed as one of the external quality assurance schemes (Matson 1998) and is principally based on the assessment of morphological characteristics in a fast, easy and non-invasive way (Baczkowski 2004). As stated by Racowsky 2010, a grading system must be simple,

containing characteristics with a proven predictive value and easily to adopt in different lab (Racowsky 2010). Therefore, the SART committee developed a three point grading system based on the evaluation of the number and size of blastomeres and the degree of fragmentation Racowsky et al. evaluated the system and reported a significant association with life birth (Racowsky 2011) This blastocyst grading system assigns 3 separate quality scores to each blastocyst embryo:

1. Blastocyst development stage - expansion and hatching status
2. Inner cell mass (ICM) score, or quality
3. Trophectoderm (TE) score, or quality

Expansion grade	Blastocyst development and stage status
1	Blastocoel cavity less than half the volume of the embryo
2	Blastocoel cavity more than half the volume of the embryo
3	Full blastocyst, cavity completely filling the embryo
4	Expanded blastocyst, cavity larger than the embryo, with thinning of the shell
5	Hatching out of the shell
6	Hatched out of the shell

ICM grade	Inner cell mass quality
A	Many cells, tightly packed
B	Several cells, loosely grouped
C	Very few cells

TE grade	Trophectoderm quality
A	Many cells, forming a cohesive layer
B	Few cells, forming a loose epithelium
C	Very few large cell

Table VIII: Blastocysts are given a quality grade for each of the 3 components and the score is expressed with the expansion grade listed first, the inner cell mass grade listed second and the trophectoderm grade third.

A wide array of grading systems have been developed over the years, mostly based on morphological features and most recently, on metabolism fluctuations with encouraging

results (Gardner 1987; Scott R 2008; Sturmey 2008). The role of metabolism in the developing embryo has been widely investigated in the literature (Leese 2008). Early metabolic data on embryos nearly always related to substrate uptake or catabolism, but have rarely been placed in context with energy production (ATP content and turnover) (Botros 2008). This also implied the monitoring of catabolic substrate precursors and intermediates, especially of the main metabolic pathways, such as glycolysis and Kreb's cycle. Metabolomics approaches to developing embryos have been proposed in recent years to discriminate eligible profiles to be related with embryo viability (Gardner 1987 ;- Scott R 2008; Sturmey 2008). However, since the approaches performed so far often relied on rather in sensitive techniques such as Raman spectroscopy (Scott 2008), only indirect embryo metabolism investigations have been performed through the analysis of extended culture media of human embryos to the blastocyst stage. In my study , we show how the sensitivity of the hereby proposed mass spectrometry (MS)-based approach allows for direct measurement in blastocoele fluid of a handful of metabolites relevant for embryo development, as reported in the literature, with no harm to the blastocyst and without disrupting the clinical pipeline prior to cryostorage of pre-implantation balstocysts. Indeed, samples as low as 0.5 nl, which routinely represent a waste product of the vitrification process of blastocyst prior to cryostorage, can be processed for metabolite detection through rapid resolution reversed phase (RR-RP) high performance liquid chromatography (HPLC)-MS.

The role of metabolism in the developing embryo has been widely investigated in the literature. Early metabolic data on embryos nearly always related to substrate uptake or catabolism, but have rarely been placed in context with energy production (ATP content and turnover). This also implied the monitoring of catabolic substrate precursors and intermediates, especially of the main metabolic pathways, such as glycolysis and Kreb's cycle. Metabolomics approaches to developing embryos have been proposed in recent years to discriminate eligible profiles to be related with embryo viability. However, since the approaches performed so far often relied on rather insensitive techniques such as Raman spectroscopy, only indirect embryo metabolism investigations have been performed through the analysis of extended culture media of human embryos to the blastocyst stage.

5.2 Materials and methods

Acetonitrile, formic acid, and HPLC-grade water, purchased from Sigma Aldrich (Milano, Italy). Standards (equal or greater than 98% chemical purity) ATP, glucose-6-phosphate, L-lactic acid, 6-phosphogluconic acid, NADH, NADPH, L-glutamic acid, α-ketoglutarate were purchased from Sigma Aldrich (Milan). Standards were stored either at -25°C, 4°C or room temperature, following manufacturer's instructions. Each standard compound was weighted and dissolved in nanopure water. Starting at a concentration of 1 mg ml^{-1} of the original standard solution, a dilution series of steps (in 18 MΩ, 5% formic acid) was performed for each of the standards in order to reach the limit of detection (LOD) and limit of quantification (LOQ). For each metabolite, LOD was calculated as 3 * standard deviation of 20 blank runs, while the LOQ as 10 * standard deviation of the blank.

5.2.1 Blastocoele fluid collection

Blastocoele fluid collection was performed at the Cervesi Hospital (Cattolica, Italy), upon approval of the local Ethic Committee. The method for blastocyst micropuncturing and aspiration of blastocoel fluid was adapted from Brison (1993), with minor modifications In brief, expanded day 5 blastocysts were removed from culture and transferred to a 10 nl droplet of pre-warmed Hepes-SOFaaBSA under a mineral oil underlay. The blastocysts were immobilized by a holding pipette connected to an oil-filled syringe and mounted on a micromanipulator. The medium surrounding the embryo was gently removed to ensure that the site of micropuncture was not contaminated by external culture medium. A finely pulled, oil-filled pipette was introduced through the mural trophectoderm to avoid damaging the ICM cells, and blastocoel fluid was aspirated gently until the blastocyst had fully collapsed around the pipette. The retrieved fluids were expelled into new purified water drops and frozen at −80°C alongside 0.5 nl control droplets of purified water. Blastocoele fluid samples were then thawed and directly processed for metabolic extraction.

5.2.2 Metabolite extraction

Metabolites were extracted using the protocol by D'Alessandro2011, with minor modifiactions. Samples were resuspended in 80 µl of -20°C methanol. 100 µl of -20 °C chloroform were added and tubes were then mixed every 5 min for 30 minutes. Subsequently, 20 µl of ice cold pH adjusted ultra-pure water (18 MΩ) was added to each tube and the tubes were centrifuged at 1000 x g for 1 min at 4 °C, before being transferred to -20° C for 2-8 h.

After thawing, liquid phases were recovered and an equivalent volume of acetonitrile was added to precipitate any residual protein. The tubes were then transferred to refrigerator (4°C) for 20 min, centrifuged at 10000 x g for 10 min at 4 °C and the supernatants were recovered into a 2 ml tube. Collected supernatants were dried as to obtain visible pellets. Finally the droed samples were re suspended in 200 ul of water, 5% formic acid and trasferrred to glass autosampler vials for LC/MS analysis.

5.2.3 Rapid resolution reversed-phase HPLC

An Ultimate 3000 rapid resolution fast HPLC system (LC Packings, DIONEX, Sunnyvale, USA) was used to perform metabolite separation. The system featured a binary pump and vacuum degasser, well-plate autosampler with a six-port micro-switching valve, a thermostated column compartment. A Dionex Acclaim RSLC 120 C18 column "2.1 mm x 150 mm, 2.2 um" was used to separate the extracted metabolites. LC parameters: injection volume, 20 µL (ten runs were possible with each extracted and re-suspended sample); column temperature, 30°C; and flow-rate of 0.2 ml min^{-1}. The LC solvent gradient and timetable were identical during the whole period of the analyses. A 0–95% linear gradient of solvent A (0.1% formic acid in water) to B (0.1% formic acid in acetonitrile) was employed over 15 min followed by a solvent B hold of 2 min, returning to 100% A in 2 minutes and a 6-min post-time solvent A hold.

5.2.4 ESI mass spectrometry

Metabolites were directly eluted into a High Capacity ion Trap HCTplus (Bruker-Daltonik, Bremen, Germany). Mass spectra for metabolite extracted samples were acquired in positive ion mode. ESI capillary voltage was set at 3000V (+) ion mode. The liquid nebulizer was set to 30 psig and the nitrogen drying gas was set to a flow rate of 9 L/min. Dry gas temperature was maintained at 300°C. Data was stored in centroid mode. Internal reference ions were used to continuously maintain mass accuracy. Data were acquired at a rate of 5 spectra/s with a stored mass range of *m/z* 50–1500. Data were collected using Bruker Esquire Control (v. 5.3 – build 11) data acquisition software. In MRM analysis, m/z of interest were isolated, fragmented and monitored (either the parental and fragment ions) throughout the whole RT range. Validation of HPLC on-line MS-eluted metabolites was performed by comparing transition fingerprints, upon fragmentation and matching against the standards metabolites through direct infusion with a syringe pump syringe pump (infusion rate 4 µl/min). Standard

calibration curves were performed either on precursor and fragment ion signals. Only the former were adopted for quantitation, as precursor ion signals guaranteed higher intensity and thus improved LOQ and LOD. Transitions were monitored to validate each detected metabolite.

5.2.5 Metabolite analysis and data elaboration: Post-hoc bioinformatic elaboration

LC/MS data files were processed by Bruker DataAnalysis 4.0 (build 234) software. Quantitative analyses of standard compounds were performed on MRM data. Each standard metabolite was run in triplicate, To evaluate the potential of the method for quantitative analysis of selected metabolites, intra- and inter-day repeatability of retention times, and linearity of the RR-RP-HPLC-ESI-MS method were tested. Intra-day repeatability was measured by injecting the same standard solution (1 μg ml^{-1}) three times in a single day. Inter-day repeatability was measured by analysing the same standard solution over 5 different days. Five-point standard curves were established by plotting integrated peak areas versus concentration. Each point on the calibration curve is the mean value of three independent measurements using the RR-RP-HPLC-ESI-MS method. Linearity of the observed quantities, slope, intercept and linear correlation values were all calculated via Microsoft Excel (Microsoft, Redmond,WA, USA).

5.3 Results and discussion

The correlation between early developmental metabolism and blastocyst implantation likelihood has been growingly investigated over the last decades. While morphology currently represents a widely accepted criterion for embryo assessment (Figure 20), recently optical and non-optical spectroscopy approaches to blastocyst metabolism have been proposed to monitor fluctuations of metabolic intermediates in culture media of developing blastocysts. Though these approaches yielded promising results (Scott 2008) and preliminary data suggested that they might be useful for embryo assessment prior to single embryo transfer in adjunct to morphology parameters (Seli 2010), the intrinsic lack in sensitivity of these analytical methods might result in incomplete or partially-reliable outputs especially in borderline cases. Although more sensitive approaches have already been proposed and applied in this field, such as microfluorimetric assays (Chi 2002), these alternative analytical methods are hampered by the need for chemical derivatization strategies which vary from metabolite to metabolite, both requiring time-consuming efforts for method set up from time to time and a

strong biochemical background which often lacks in the routine laboratory end-user. Thus, a valid alternative to the already available approaches should be both sensitive and versatile. In this view, we hereby propose a complementary MS-based strategy which might represent a further stride towards embryo assessment through targeted monitoring of metabolism fluctuations, not only indirectly from culture media, but also directly through the analysis of blastocoele fluid, an IVF-waste product prior to vitrification for cryostorage. From sample volumes as low as 0.5 nl we could detect and quantify against external standards a group of metabolites (**Table IXa**), whose roles in blastocyst development and embryo metabolism have long been postulated (Leese 2008). The list included i) ATP (Figure 18.A); ii) glucose-6-phosphate; iii) lactate; iv) NAD+ (Figure 18.B) and v) NADH and vi) NADPH; vii) 6-phosphogluconic acid (Figure 18.C); viii) glutamic acid and ix) α-ketoglutarate. Theoretically, this strategy could be also applied to every other metabolite the concentration of which would have lied within the limits of detection (for detection only) and limits of quantifications, as calculated through repeated measurements of a high-purity external standard (see also the Materials and Methods section and reference (D'Alessandro 2011) for further details). While we had already documented the metabolome-wide versatility and matrix-independent robustness of the hereby proposed method we had not yet extensively applied this method to the study of blastocoele fluid in order to determine the extent of metabolites that could be both of biological interest in this matrix and rapidly monitored through this approach. Of note, ATP, lactate, NAD^+/NADH, NADPH, 6-phosphogluconic acid, glutamic acid and ketoglutarate could be monitored and validated through MS/MS (details of fragment transitions are provided in **Table VIII**, along with PubChem ID, chromatographic retention times) quantified down to the picomole range (over 4-5 orders of magnitude), in line with our previous set up of the method for other biological matrices, such as red blood cells, SAOS cell lines.

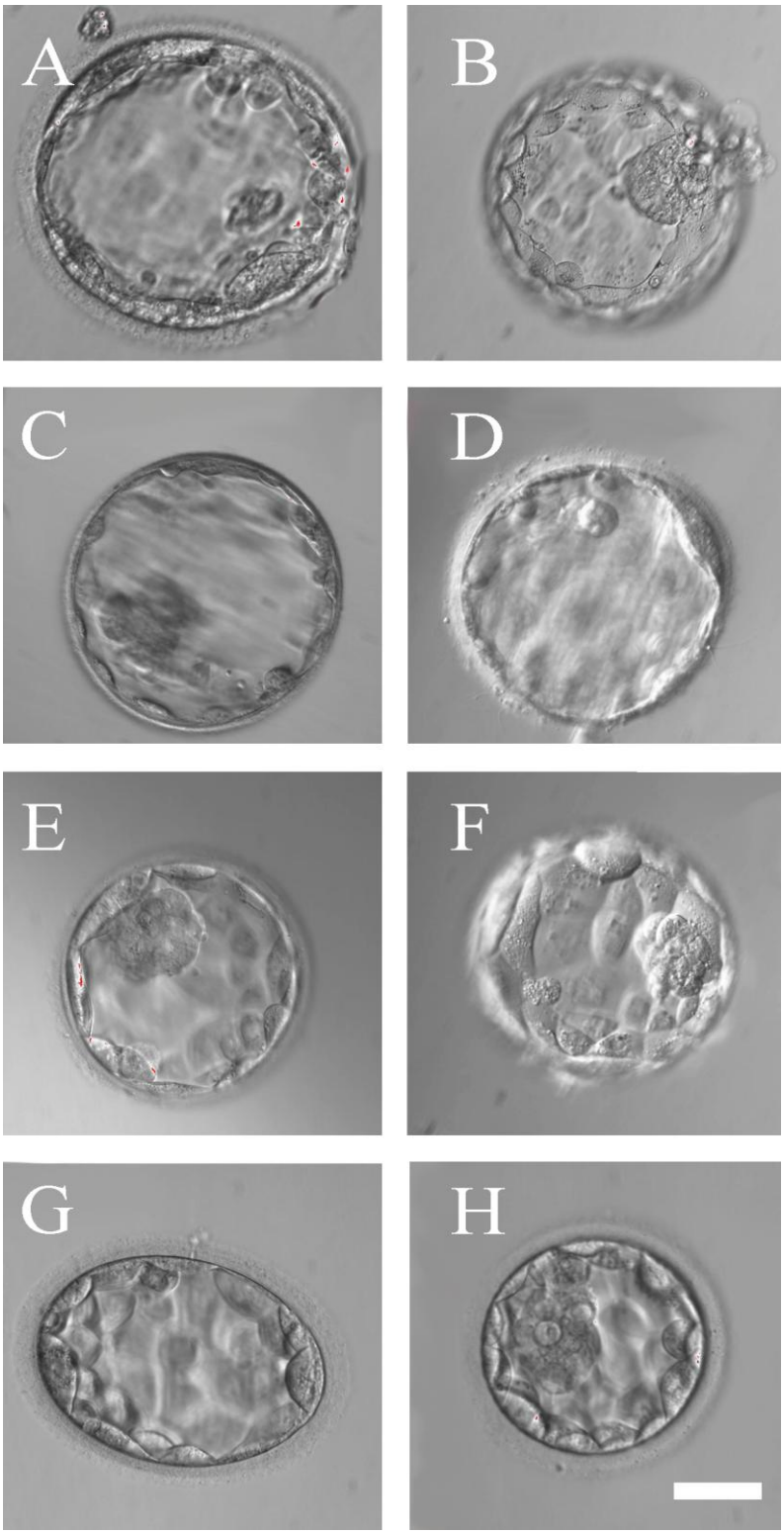


Figure 20 : Microscopic images of human blastocysts, according to classification based on "Grading criteria for human blastocyst" .

Metabolite	PubChem ID	Monoisotopic mass	MS/MS	Retention time (min)	Standard curves	Linearity: orders of magnitude	Linear correlation coefficient	Ion mode
Lactic acid	612	90,0317	43	2,8	Y= 563,0101X+ 14,2373	4	0,998704	-
Glucose-6-phosphate	69507	260,0297	97	2,6	Y = 416,5324 X + 22,43876	4	0,993721	-
α-ketoglutaric acid	164533	146,0215	57	3,0	Y= 441,2236X + 48,7091	4	0,995342	-
6-phospgluconic acid	91493	276,0246	79	3,1	Y=864,213X + 39,5128	4	0,994114	-
Glutamic acid	611	147,0532	128	2,5	Y= 267,7217X+ 14,50912	5	0,999327	+
ATP	5957	506,9957	410	3,5	Y= 930,1519 X +37,3424	4	0,997755	+
NAD+	5893	663,1091	524	3,9	Y= 524,16X+ 23,06532	4	0,996154	+
NADH	3687	665,1248	649	4,1	Y= 439,21X+ 13,06532	5	0,997223	+
NADPH	5884	745,4209	729	5,4	Y=735,33X+32,41	4	0,995472	+

Table VII : Metabolites identified through RR-RP-HPLC- MRM-ESI/MS

All of these compounds early eluted upon rapid resolution chromatographic separation, allowing for extremely low time-consuming runs (15 minutes including post-gradient solvent hold for cleaning the column). The choice of monitoring the 8 compounds (**Table VII**) stemmed from information gathered from literature (Leese 2007) ATP (Figure 20.A), for example, is the primary energy token in all living cells and the ATP/ADP ratio in rabbit blastocyst has been extensively studied since 1971 (Brooks 1971). ATP levels have long been known to reflect viability of the embryo, its survival/implant, necrosis or apoptosis rate (the latter still requiring ATP) (Enos 1983). Therefore, the attention of metabolic studies in this field has soon moved towards understanding whether healthy blastocysts tended to produce ATP through glycolysis, or rather relied on Kreb's cycle, as it emerged upon early studies (Enos 1983). Indeed, glucose was the first energy substrate shown to support development of mammalian embryos outside the female reproductive tract (Hammond 1949). Glucose supported mouse blastocyst development from eight-cell embryos, but not from earlier stages,

when supplemented as the sole substrate in the medium (Whitten 1956). On the other hand, glucose resulted to be partly or largely responsible for the "blocks to development", leading to the formulation of the so-called "quiet metabolism" theory (Leese 2007). Glucose-6-phosphate is the first metabolite of glucose in cell cytoplasm, through phosphorylation of the oxidril group in carbon 6 by the enzyme hexokinase. Therefore, glucose-6-phosphate levels might represent a solid indirect marker of glucose uptake by the developing embryo. While it might be controversial its presence in the blastocoele fluid, the understanding of the biological bases for our observations is beyond the scope of the present communication. The inability of glucose to support in vitro development of early cleavage stage mouse embryos stimulated interest in evaluating other intermediates of carbohydrate metabolism. In this view, pyruvate (uptake) and lactate (release) have emerged as key regulatory substrates. Lactate for example, when substituted as calcium lactate for calcium chloride, sustained two cell mouse development up to the blastocyst stage, revealing a critical role for lactate in mammalian embryo development in vitro (Whitten 1956). Moreover, since there is a marked increase in lactate formation from pyruvate when the cytosolic NAD+/NADH ratio is low, NADH levels could be directly proportional to pyruvate formation and subsequent entrance to Krebs cycle (Gardner 2000). In line with this, it has been reported a decrease in developmental ability of porcine embryos cultured by use of glucose instead of pyruvate and lactate after the fertilization, which may be due to the rise in ROS generation in Day 1 embryos (Karja 2006). Generation of ROS induced by glucose utilization was assumed to be caused by the activation of NADPH oxidase, an enzyme that catalyzes the oxidation of NADPH, generates NADP that serves as a coenzyme of the oxidative arm of the pentose phosphate pathway (PPP) (Takahashi 1992). Production of superoxide anion and H2O2 via NADPH oxidase has been described on a rabbit blastocyst surface (Manes 1995,) and the incubation of mouse embryos with an inhibitor of NADPH oxidase induces a dose-dependent reduction in H2O2 production (Nasr-Esfahani 1991). Since activity of PPP was higher at zygotes and embryos at the 2-cell stage compared with later developmental stages (Swain et al 2002), thus enhanced ROS production in embryos cultured with glucose in this study may be associated with the glucose utilization by the early developmental stage of porcine embryos through the PPP. The first step of PPP leads to the production of 6-phosphogluconolactone through reduction of NADP to NADPH by the enzyme glucose-6-phosphate dehydrogenase. 6-phosphogluconolactone is then converted to 6-phosphogluconic acid by a gluconolactonase. Therefore, monitoring of 6-phosphogluconic acid (Figure 18.C) and NADPH could provide precious hints on eventual shifts towards the PPP in the developing embryo through the analysis of the blastocoele fluid.

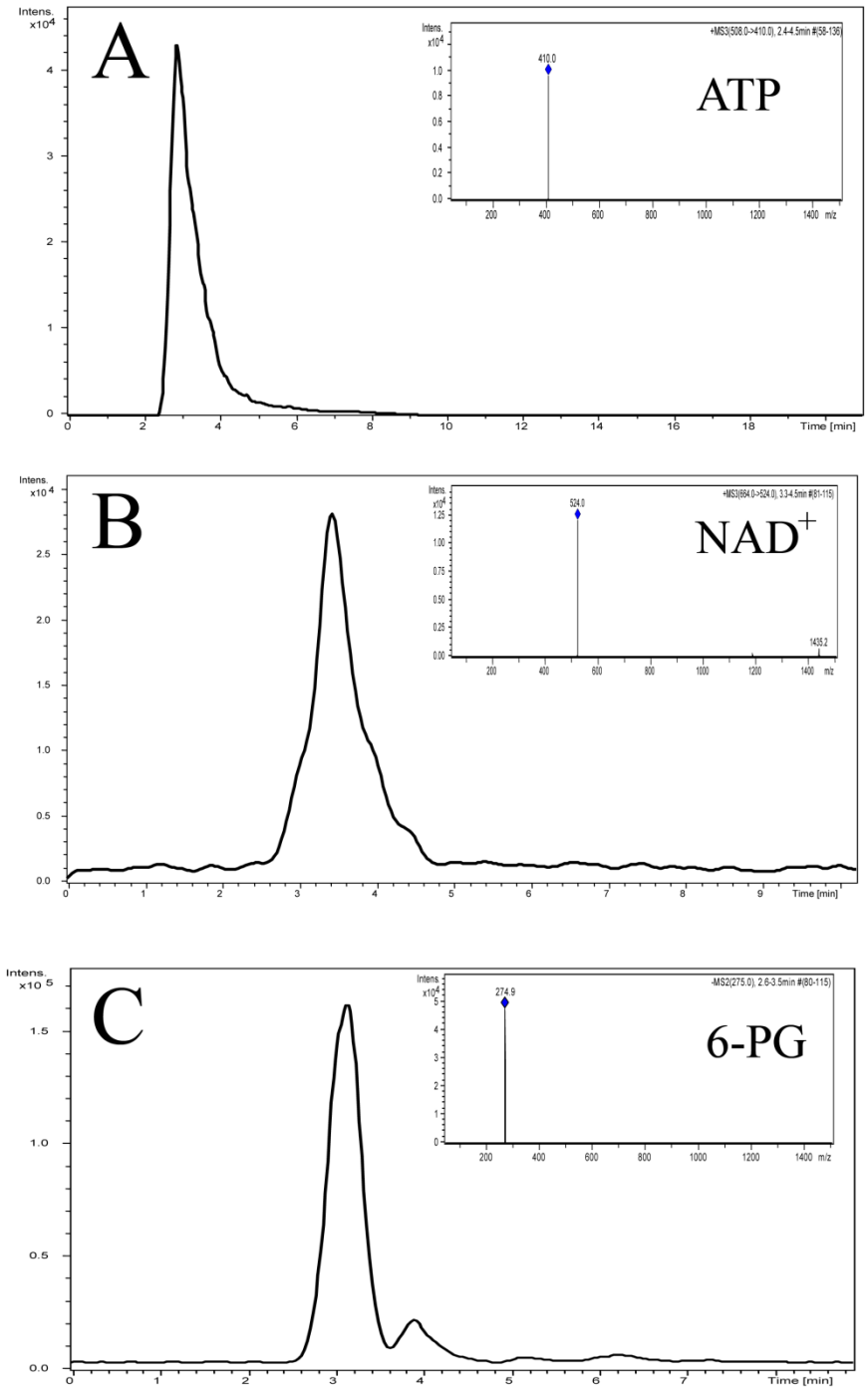


Figure 21: Three representative chromatograms of metabolites detected in blastocoele fluid through multiple reaction monitoring (MRM) in MS-based targeted analysis: ATP (A), NAD^+ (B) and 6-phosphogluconic acid (6PG) (C). In the right side of each graph the MS andMS/MS spectra are indicated for 6PG and ATP, NAD^+, respectively.

As alternative catabolic substrates and fundamental anabolic substrates for protein build-up, amino acids also profoundly enhance the preimplantation development of embryos and

strongly influence viability of mbryos after transfer (Zhang 1990, Lane 1994).Glutamine, for example, is utilized as an energy substrate in pre-implantation mouse embryos (Chatot 1997). Glutaminase is the enzyme responsible for the conversion of glutamine to glutamic acid, which then enters the Krebs cycle as α-ketoglutarate. Our method allowed detection and quantitation of glutamate and ketoglutarate down to picomole quantities (1.35 and 13.70 injected pmol, respectively).

5.4 Conclusion

Whit this study I can demonstrate that a fast HPLC-MS-based metabolomic approach which enables detection, validation and quantitation of metabolites of biological interest for embryo development from blastocoele fluid down to the picomole quantities.The advantages of such analytical approach include high sensitivity and versatility for a broad range of metabolites, without the need for metabolite-specific chemical derivatization. The high sensitivity of the MS analysis on the blastocoele fluid enables to gather direct information on embryo metabolism, in contrast with the approaches adopted so far which aim at indirectly determining metabolite levels from the culture media or rather directly through disruptive analysis on the whole blastocyst. Blastocoele fluid is a routine waste product prior to pre-implantation blastocyst vitrification. Further clinical studies will be directed at exploiting this approach in adjunct to morphology scores in order to determine whether an actual correlation exists between metabolite levels in the blastocoele fluid and the implantation efficiency.

5.5 References:

Baczkowski T., Kurzawa R., Gabowski W: (2004) Methods of embryo scoring in vitro fertilization. Reprod Biol 4:5-22.

Benagiano G., Gianaroli L. (2010) The Italian Constitutional Court modifies Italian legislation on assisted reproduction technology. Reprod Biomed Online. Mar;20(3):398-402. Epub 2009 Dec

Botros L., Sakkas D., Seli E. (2008) Metabolomics and its application for non-invasive embryo assessment in IVF. Mol Hum Reprod. 14(12):679-690.

Brison D.R., Hewitson L.C., Leese H.J. (1993) Glucose, pyruvate, and lactate concentrations in the blastocoel cavity of rat and mouse embryos. Mol Reprod Dev 35(3):227-232.

Brooks D.E., Lutwak-Mann C. (1971) Content of ATP and ADP in Rabbit Blastocysts. Nature ; 229: 202-203.

Chatot C.L., Lawry J.R., Germain B., Ziomek C.A.(1997) Analysis of glutaminase activity and RNA expression in preimplantation mouse embryos. Mol Reprod Dev. ; 47(3):248-254.

Chi M.M., Hoehn A., Moley K.H. (2002) Metabolic changes in the glucose-induced apoptotic blastocyst suggest alterations in mitochondrial physiology. Am J Physiol Endocrinol Metab. ; 283(2): 226-232.

D'Alessandro A., Gevi F., Zolla L. (2011) A robust high resolution reversed-phase HPLC strategy to investigate various metabolic species in different biological models. Mol Biosyst.; 7(4):1024-1032.

Enos D.J., Balaban R.S. (1983) Energy metabolism of preimplantation mammalian blastocysts. Am J Physiol. 245(1):40-45.

Gardner D.K., Leese H.J. (1987) Assessment of embryo viability prior to transfer by the non invasive measurement of glucose uptake. J Exp Zool 1987; 242:103–105.

Gardner D.K., Vella P., Lane M., Wagley L., Schlenker T., Schoolcraft W. (1998) Culture of human blastocysts increases implantation rates and reduces the need for multiple embryo transfers. Fertil Steril 69:84–88.

Gardner D.K., Schoolcraft W. (1999) Elimination of higher order multiple gestations by blastocyst culture and transfer. In: Shoham Z, Howles C, Jacobs H, editors. Elimination of higher order multiple gestations by blastocyst culture and transfer. London: Martin Dunitz; . 267–274.

Gleicher N., Barad D. (2006) The relative myth of elective single embryotransfer. Hum Reprod 21:1337–44.)

Hammond J. (1949) Jr. Recovery and culture of tuba1 mouse ova. Nature 163:28-29.

Lane M., Gardner D.K. (1994) Increase in postimplantation development of cultured mouse embryos by amino acids and induction of fetal retardation and exencephaly by ammonium ions. J Reprod Fertil 102:305312.

Leese H.J., Baumann C.G., Brison D.R., McEvoy T.G., Sturmey R.G.(2008) Metabolism of the viable mammalian embryo: quietness revisited. Mol Hum Reprod. 14(12):667-672.

Leese H.J., Sturmey R.G., Baumann C.G., McEvoy T.G. (2007) Embryo viability and metabolism: obeying the quiet rules. Hum Reprod.; 22(12):3047-3050.

Matson P: (1998) Internal quality control and external quality assurance in the IVF laboratory. Hum Reprod, 13:156-165.

Karja N.W., Kikuchi .K., Fahrudin M., Ozawa M., Somfai T., Ohnuma K., Noguchi J., Kaneko H., Nagai T. (2006) Development to the blastocyst stage, the oxidative state, and the quality of early developmental stage of porcine embryos cultured in alteration of glucose concentrations in vitro under different oxygen tensions. Reprod Biol Endocrinol 4:54.

Manes C. and Lai N.C. (1995) Nonmitochondrial oxygen utilization by rabbit blastocysts and surface production of superoxide radicals. J Reprod Fertil 104, 69-75.

Nasr-Esfahani M.M., Johnson M.H (1991) The origin of reactive oxygen species in mouse embryos cultured in vitro. Development ; 113:551-560.

Pennings G.(2000) Multiple pregnancies: a test case for the moral quality of assisted reproduction. Hum Reprod ;15:2466–2469.

Racowsky C., Vernon M., Mayer J., Ball G.D., Behr B., Pomeroy K.O., Wininger D., Gibbons W., Conaghan J., Stern J.E. (2010) Standardization of grading embryo morphology. J Assist Reprod Genet , 27:437-439.

Racowsky C., Stern J.E., Gibbons W.E., Behr B., Pomeroy K.O., Biggers J.D. (2011) National collection of embryo morphology data into Society for Assisted Reproductive Technology Clinic Outcomes Reporting System: associations among day 3 cell number, fragmentation and blastomere symmetry and live birth rate. Fertil Steril, 95:1985-1989.

Scott R., Seli E., Miller K., Sakkas D., Scott K., Burns D.H.(2008) Non invasive metabolomic profiling of human embryo culture media using Raman spectroscopy predicts embryonic reproductive potential: a prospective blinded pilot study. Fertil Steril; 90(1):77-83

Seli E., Vergouw C.G., Morita H., Botros L., Roos P., Lambalk C.B., Yamashita N., Kato O., Sakkas D.(2010) Non invasive metabolomic profiling as an adjunct to morphology for noninvasive embryo assessment in women undergoing single embryo transfer. Fertil Steril. 94(2):535-542.

Sturmey R.G., Brison D.R., Leese H.J. (2008) Symposium: innovative techniques in human embryo viability assessment. Assessing embryo viability by measurement of amino acid turnover. Reprod Biomed Online. 17(4):486-496.

Takahashi Y., First N.L (1992). In vitro development of bovine one-cell embryos: Influence of glucose, lactate, pyruvate, amino acids and vitamins. Theriogenology ; 37:963-978.

Whitten W.K.(1956) Culture of tubal mouse ova. Nature 177:96.

Zhang X., Armstrong D.T. (1990)Presence of amino acids and insulin in a chemically defined medium improves development of 8-cell rat embryos in vitro and subsequent implantation in vivo. Biol Reprod 42:662-668.

CHAPTER 6

Alterations of Red Blood Cell metabolome during cold liquid storage of erythrocyte concentrates in CPD-SAGM

6.1 Introduction

Red blood cell: are the most common type of blood cell. In humans, mature red blood cells are flexible biconcave disks that lack a cell nucleus and most organelles. 2.4 million new erythrocytes are produced per second. (Sackmann 1995) The cells develop in the bone marrow and circulate for about 100–120 days in the body before their components are recycled by macrophages. Each circulation takes about 20 seconds. Approximately a quarter of the cells in the human body are red blood cells. (Dean Pierigè 2008)

6.1.2 Red blood cell storage.

Red blood cells are still the most widely transfused blood component at present, the most widely used protocol for the storage of red blood cells (for up to 42 days) is the collection of blood into anticoagulant solutions (typically citrate-dextrose-phosphate); red cell concentrates are prepared by the removal of plasma and, in some cases, also leukoreduction. The product is stored at 4 ± 2° C in a slightly hypertonic additive solution, generally SAGM (sodium, adenine, glucose, mannitol, 376 mOsm/L)1 The current standard requirements for patenting new additive solutions in the USA, and also suggested in the recommendations of the European Counci l37, are essentially based on two parameters: the level of haemolysis (below the threshold of 0.8% at the end of the storage period, following the introduction of the "95/95" rule38) and a survival rate of the transfused cells of more than 75% at 24 hours after transfusion. The major metabolic function of the erythrocyte is to produce the necessary cofactors (ATP, NADPH and NADH) by energy and redox metabolisms for maintaining its osmotic balance and electroneutrality and fighting oxidative stresses (Bossi 1996; Joshi 1989). These cofactors are also necessary for the bioconcave shape of the cell as well as for the specific intracellular cation concentrations. Indeed, while current guidelines allow storing RBC concentrates for as long as 42 days under cold (1-6°C) liquid blood-bank conditions, it is still matter of debate whether blood stored longer than 14 days end up resulting in impaired safety and effectiveness of the blood therapeutic to the transfused recipient. A brief list of the elements of the so-called "red blood cell storage lesion" includes (Bennett 2007): morphological changes, slowed metabolism with a decrease in the concentration of adenosine

triphosphate (ATP), acidosis with a decrease in the concentration of 2,3- diphosphoglycerate (2,3-DPG), loss of function (usually transient) of cation pumps and consequent loss of intracellular potassium and accumulation of sodium within the cytoplasm, oxidative damage with changes to the structure of band and lipid peroxidation, apoptotic changes with racemisation of membrane phospholipids and loss of parts of the membrane through vesiculation. Some of these changes occur within the first few hours of storage, for example, the decrease in pH or the increases in potassium and lactate; others, however, take days or weeks. Together, these events risk compromising the safety and efficacy of long-stored red blood cells, reducing their capacity to carry and release oxygen, promoting the release of potentially toxic intermediates (for example, free haemoglobin can act as a source of reactive oxygen species) and negatively influencing physiological rheology (through the increased capacity of the red blood cells to adhere to the endothelium (Anniss 2005; Koshkaryev 2009). The present study provides the first evidence that over storage duration metabolic fluxes in red blood cells proceed from pentose phosphate pathway towards purine salvage pathway, instead of massively re-entering glycolysis via the nonoxidative phase.

6.2 Materials and methods

6.2.1 Sample collection

Red blood cells units were drawn from healthy human volunteers according to the policy of the Italian National Blood Centre guidelines (Blood Transfusion Service for donated blood) and all thevolunteers provided their informed consent in accordance with the declaration of Helsinki. We studied RBC units collected from 10 healthy donor volunteers [male=5, female=5, age 42.3 ± 10.5 (mean ± S.D.)]. RBC units were stored for up to 42 days under standard conditions (CDP-SAGM,4°), while samples were removed aseptically for the analysis on a weekly basis (at 0, 7, 14, 21, 28,35 and 42 days of storage).

6.2.2 Metabolite extraction

For each sample, 0.5mL from the pooled erythrocyte stock was transferred into a microcentrifuge tube (Eppendorf ® Germany). Erythrocyte samples were then centrifuged at 1000g for 2 minutes at 4°C. Tubes were then placed on ice while supernatants were carefully aspirated, paying attention not to remove any erythrocyte at the interface. Samples were extracted following the protocol by D'Alessandro.. The erythrocytes were resuspended in 0.15 mL of ice cold ultra-pure water (18 MΩ) to lyse cell, then the tubes were plunged into a

water bath at 37°C for 0.5 min. Samples were mixed with 0.6 mL of -20°C methanol and then with 0.45 mL chloroform. Subsequently, 0.15ml of ice cold ultra-pure water were added to each tube and they were transferred to -20°C freezer for 2-8 h. An equivalent volume of acetonitrile was added to the tube and transferred to refrigerator (4°C) for 20 min. Samples with precipitated proteins were thus centrifuged for 10000 x g for 10 min at 4 °C. Finally, samples were dried in a rotational vacuum concentrator (RVC 2-18 - Christ Gmbh; Osterode am Harz, Germany) and resuspended in 200 µl of water, 5% formic acid and transferredto glass auto-sampler vials for LC/MS analysis.

6.2.3 Rapid Resolution Reversed-Phase HPLC

An Ultimate 3000 Rapid Resolution HPLC system (LC Packings, DIONEX, Sunnyvale, USA) was used to perform metabolite separation. The system featured a binary pump and vacuum degasser,well-plate autosampler with a six-port micro-switching valve, a thermostated column compartment. Samples were loaded onto a Reprosil C18 column (2.0mm×150mm, 2.5 µm - Dr Maisch, Germany) for metabolite separation. Chromatographic separations were achieved at a column temperature of 30°C; and flow rate of 0.2 mL/min. For downstream negative ion mode (-) MS analyses, A 0–100% linear gradient of solvent A (10mM tributylamine aqueous solution adjusted with 15mM acetic acid, pH 4.95) to B (methanol mixed with 10 mM TBA and with 15 mM acetic acid, pH 4.95) was employed over 30 min, returning to 100% A in 2 minutes and a 6-min post-time solvent A hold. For downstream positiveion mode (+) MS analyses, a 0–100% linear gradient of solvent A (ddH_2O, 0.1% formic acid) to B(acetonitrile, 0.1% formic acid) was employed over 30 min, returning to 100% A in 2 minutes and a 6-min post-time solvent A hold.

6.2.4 Mass Spectrometry: Q-TOF settings

Due to the use of linear ion counting for direct comparisons against naturally expected isotopic ratios, time-of-flight instruments are most often the best choice for molecular formula determination. Thus, mass spectrometry analysis was carried out on an electrospray hybrid quadrupole time-of flight mass spectrometer MicroTOF-Q (Bruker-Daltonik, Bremen, Germany) equipped with an ESI-ion source. Mass spectra for metabolite extracted samples were acquired both in positive and in negative ion mode. ESI capillary voltage was set at 4500V (+) (-) ion mode. The liquid nebulizer was set to 27 psi and the nitrogen drying gas was set to a flow rate of 6 L/min. Dry gas temperature was maintained at 200°C. Data were

stored in centroid mode. Data were acquired with a stored mass range of m/z 50–1200. Automatic isolation and fragmentation (AutoMS$_n$ mode) was performed on the 4 most intense ions simultaneously throughout the whole scanning period (30 min per run). Calibration of the mass analyzer is essential in order to maintain an high level of mass accuracy. Instrument calibration was performed externally every day with a sodium formate solutionconsisting of 10 mM sodium hydroxide in 50% isopropanol: water, 0.1 % formic acid. Automatedinternal mass scale calibration was performed through direct automated injection of the calibration solution at the beginning and at the end of each run by a 6-port divert-valve.

6.2.5 Data elaboration and analysis

In order to reduce the number of possible hits in molecular formula generation, we exploited theSmartFormula3D TM software (Bruker Daltonics, Bremen, Germany), which directly calculatesmolecular formulae based upon the MS spectrum (isotopic patterns) and transition fingerprints (fragmentation patterns). This software generates a confidence-based list of chemical formulae on the basis of the precursor ions and all fragment ions, and the significance of their deviations to the predicted intact mass and fragmentation pattern (within a predefined window range of 5 ppm). Triplicate runs for each one of the 10 biological replicate over storage duration were exported as mzXML files and processed through XCMS data analysis software (Scripps Centre for Metabolomics) Tautenhahn 2011. Mass spectrometry chromatograms were elaborated for peak alignment, matching and comparison of parent and fragment ions, and tentative metabolite identification (within a 20 ppm mass-deviation range between observed and expected results against the internaldatabase – METLIN (Smith 2005). XCMS is an open-source software and is freely available from the website (http://metlin.scripps.edu/download/).

6.3 Results and Discussions.

6.3.1 Glycolytic intermediates were rapidly depleted over the first two weeks, while end-products accumulated

Rapid pH drop over storage duration is long known to relate to active glycolysis and lactate accumulation, other than cation homeostasis dysregulation. In our previous targeted investigation we could only monitor six distinct glycolytic metabolites, including hexoses 6-phosphate (either glucose or fructose 6-phosphate, as it is not possible to discriminate these

molecules from MS analysis of intact mass or fragmentation patterns), fructose 1,6-biphosphate, glyceraldehyde 3-phosphate, DPG, phosphoenolpyruvate and lactate (D'Alessandro 2011). Through shifting from a targeted (MRM) to an untargeted (MicroTOFQ-based) platform, we were hereby able to confirm trends for the glycolytic metabolites and complete the list with dihydroxyacetone phosphate, phosphoglycerate and pyruvate (Figure 22). While technical advantages of actual metabolome-wide rather than targeted approaches have been already reported in the frame of RBCs (i.e. hereditary stomatocytosis and sickle cell disease) (Darghouth 2011), no actual metabolomics study has so far addressed the storage issue. In biological terms, we could confirm (Nishino 2009) increases of early glycolytic intermediates (hexose 6- phosphate, fructose 1,6-biphosphate, glyceraldehyde 3-phosphate and dihydroxyacetone phosphate) within the first week of storage, while a rapid decrease of all glycolytic metabolites was observed soon afterwards. On the other hand, late glycolytic intermediates, such as phosphoenolpyruvate, pyruvate and lactate followed different trends, with pyruvate and lactate slowly increasing throughout the whole storage period (Figure 22), partly confirming and partly expanding available data from literature for SAGM-stored erythrocyte concentrates (Burger 2010; Högman 2006). Glycerol 3-phosphate was observed to increase within the first two weeks of storage and thus rapidly decrease. The trend for this metabolite is interesting since glycerol 3-phosphate, a precursor to glycerol which is synthesized from glyceraldehyde 3-phosphate, in human cells is exploited in lipidogenesis for biosynthesis of triglycerids. However, phospholipid synthesis is known to be active in retyculocytes and suppressed in mature RBCs (Percy 1973). Alteration of membrane lipid homeostasis has been widely documented in the frame of RBC storage, due to lipid peroxidation leading to membrane-targeting shape alterations, which are characterized by progressive loss of lipids (and membrane-associated proteins) through vesiculation. The progressive drop of glycerol 3-phosphate levels from day 14 onwards might reflect the extent of vesiculation phenomena. In parallel with this statement, a significant increase within the first two weeks of storage and a subsequent rapid decrease was hereby observed for a wide series of fatty acids and lipids over storage duration .

6.3.2 Metabolix fluxes towards the Pentose PhosphatePathway were altered

An indirect parameter to assess an increase in oxidative stress in RBCs is the ratio of glycolysis/pentose phosphate pathway (PPP) fluxes. Under normal steady-state conditions, 92% of glucose is metabolized along glycolysis (Embden Meyerhoff) and 8% along PPP. Under oxidant conditions up to 90% of glucose can be metabolized along PPP (Messana 1999

Messana 2000). The main purpose of the PPP is to regenerate NADPH from NADP through an oxidation/ reduction reaction. This reaction is coupled to the formation of ribose 5-phosphate from glucose 6-phosphate. In RBCs, the major role of NADPH is to reduce the disulfide form of glutathione to the sulfhydryl form. The reduced glutathione is pertinent for maintaining the normal structure of RBCs and for keeping hemoglobin in the ferrous state [Fe(II)]. The nonoxidative portion of the pathway creates carbon chain molecules ranging from 3 to 7 carbons. These compounds are intermediates in glycolysis and gluconeogenesis or other biosynthetic processes. The oxidative phase of PPP primarily produces NADPH and ribose 5-phosphate, while the nonoxidative phase yields fructose 6-phosphate, and glyceraldehyde 3-phosphate In our previous investigation, we concluded that PPP appeared to be over-activated at the oxidativephase level. However, we could only postulate that some blockade might have existed at the nonoxidative phase, since it had been reported by NMR that no glyceraldehyde 3-phosphate was produced via PPP in long stored RBCs (Messana 1999). Hereby, we could confirm and expand previous observation about the increase of metabolic intermediates and byproducts of the oxidative phase of PPP (6-phosphogluconolactone, 6- phosphogluconic acid, NADPH) (Figure 23). Additionally, we could further delve into our previous hypothesis about a blockade at the nonoxidative phase level, through monitoring trends for sedoeptulose 7-phosphate (increasing its levels over storage duration) and other metabolic intermediates of the nonoxidative phase (erythrose 4-phosphate, ribose 5-phosphate, ribose 1- phosphate, xilulose 5-phosphate), which decreased progressively over storage duration (Figure 23). A net increase was observed from PPP towards the Purine Salvage Pathway rather than reentering glycolysis.Since glyceraldehyde 3-phosphate was confirmed not to increase over storage duration, in agreement with literature (Messana 2000), the increase of oxidative phase intermediates might be explained through divert to Purine Salvage Pathway (PSP). Purine nucleotides may be synthesizedin cells de novo or reconstructed from already existing free purine bases through the salvagereactions (reutilization) (Schuster 2005) .

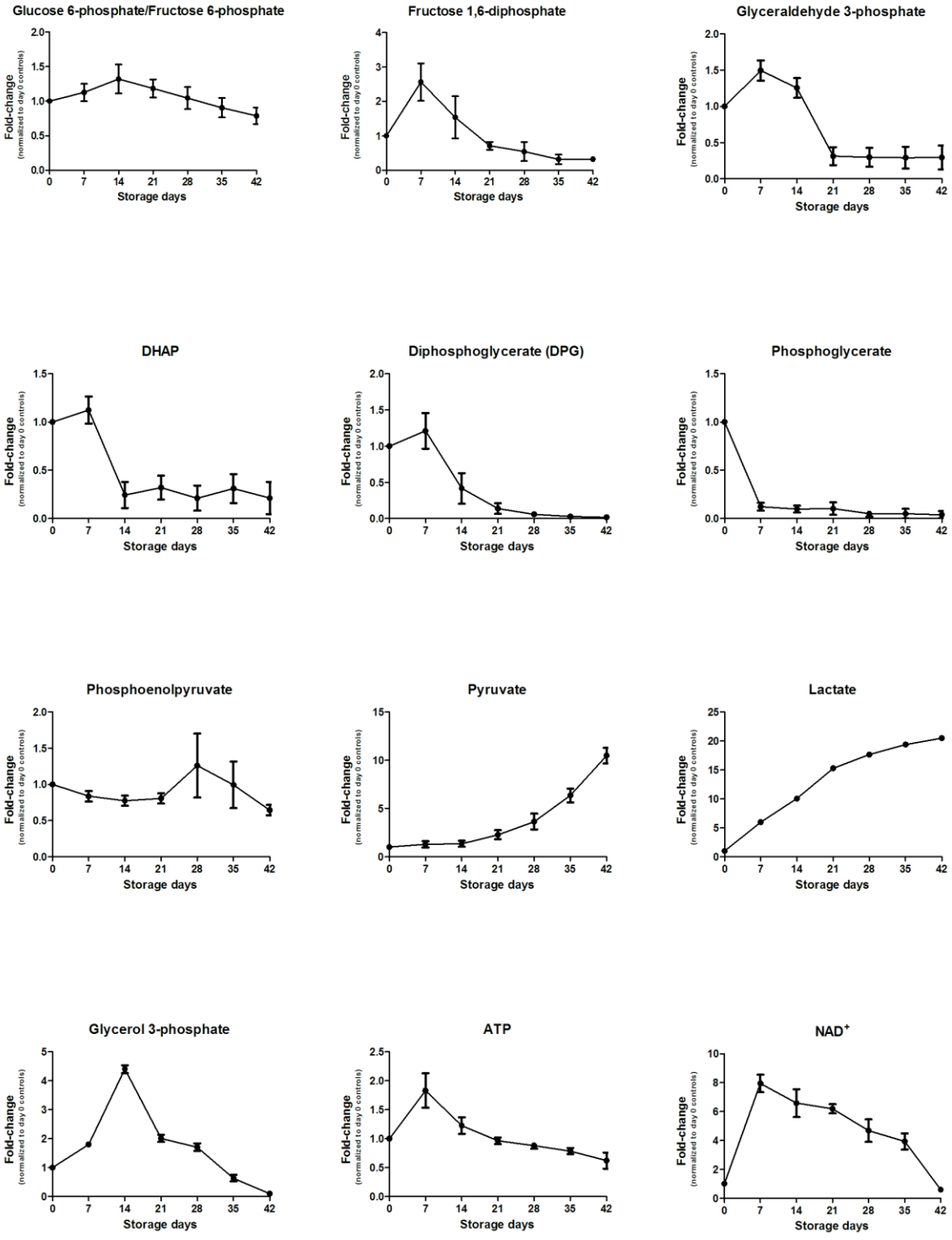


Figure 22: Time course metabolomic analyses of SAGM-stored RBCs, upon normalization against day 0 controls, through plotting of fold-change variations in agreement with analogous studies of untargeted metabolomics applied to RBC investigations.

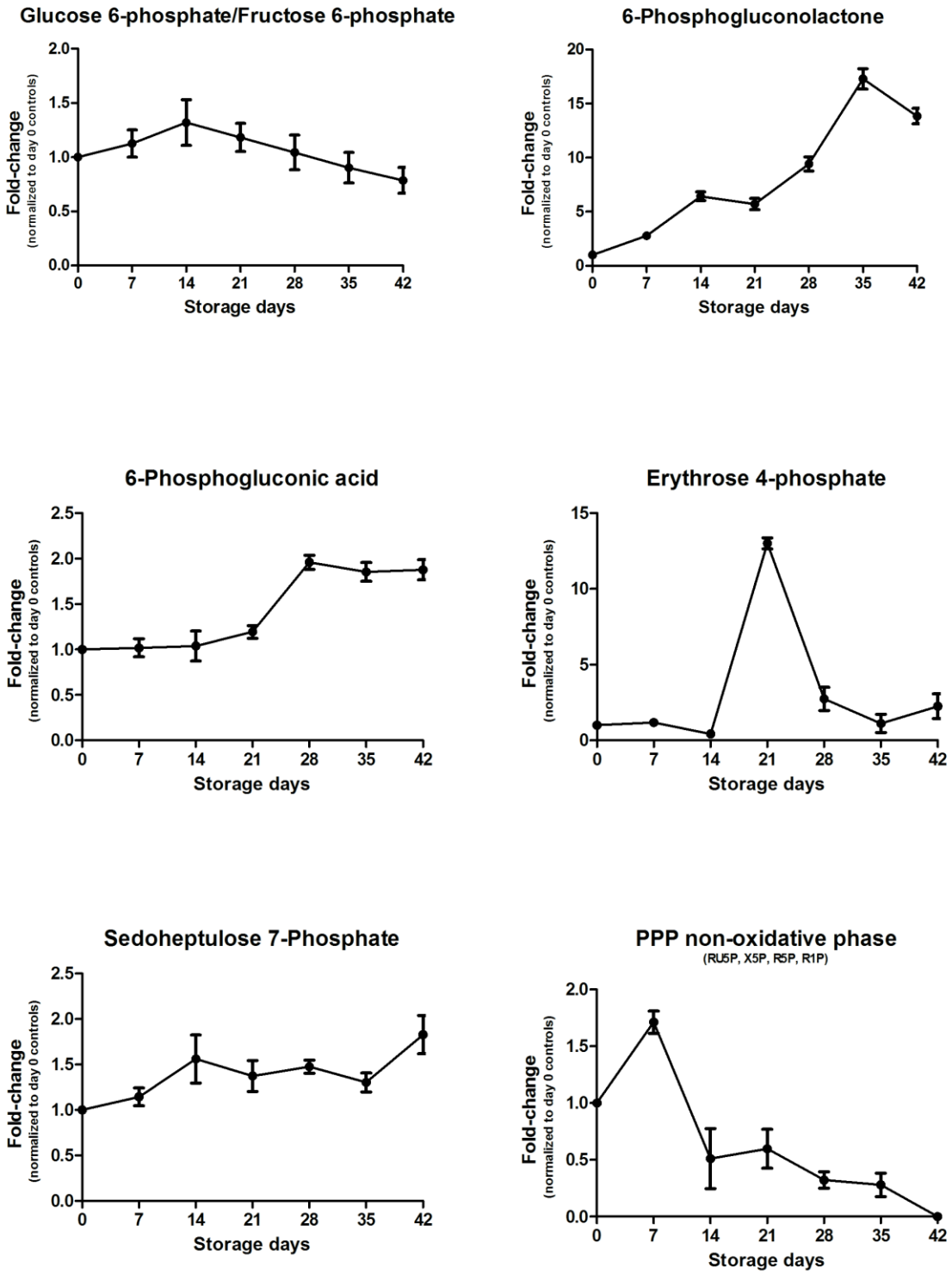


Figure 23: An overviewof trends for Pentose Phosphate Pathway metabolites.

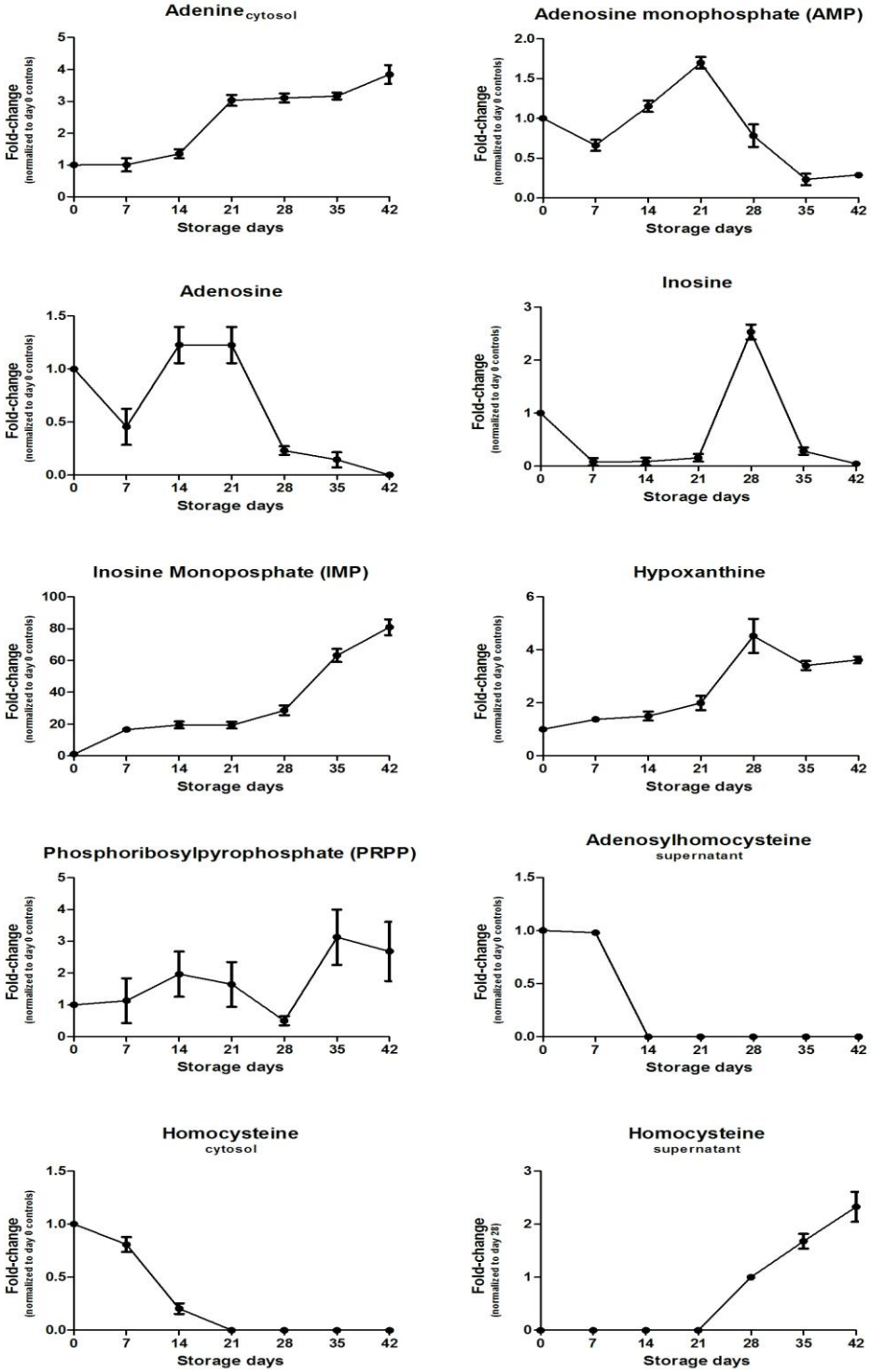


Figure 24: An overview of trends for Purine Salvage Pathway metabolites.

Mature erythrocytes cannot synthesize 5-phosphoribosylamine andthat is why the synthesis of nucleotides de novo is not possible in these cells (Schuster 2005). However, RBCsrely on alternate routes by which nucleosides or bases can be recycled to give nucleotide triphosphates, hence the name of salvage pathway. Adenine nucleotides are 70-80% of all freeerythrocyte nucleotides, and their precursors in RBCs are adenine and adenosine, the transport ofwhich through the erythrocyte membranes takes place through the facilitated diffusion. OverRBC storage duration, we observed a decrease of phosphorylated forms of adenine (ATP – Figure22; ADP, AMP and adenosine – Figure 24), while adenine it self increased throughout storage duration (Figure24). Hypoxanthine, phosphorybosylpyrophosphate and inosine monophosphate, which are major substrates for salvage reactions increased constantly as storage progressed (Figure 24). Adenosinehomocysteine, which serves as a substrate to produce adenosine and homocysteine, was rapidly depleted from the supernatant along with both adenosine and homocysteine, the latter accumulating in the supernatant (Figure 24). Anomalies to homocysteine fine-tuning are known to be related to oxidative stress and glutathione homeostasis in RBCs as indices for middle-aged untreated essential hypertension patients (Muda 2003).

6.4 Conclusion

In the present study, we performed a statistically-robust metabolome-wide analysis via MS on RBC samples over storage duration under cold liquid blood bank conditions. As a result, we could confirm and expand existing literature about the rapid fall of glycolytic rate and accumulation of glycolysis end products. A shift was observed towards the oxidative phase of PPP, in response to an exacerbation of oxidative stress. However, metabolic fluxes proceeded from PPP towards PSP, instead of massively re-entering glycolysis via the nonoxidative phase. The present study will pave the way for future investigations aiming to assess the validity of newly proposed additive solutions or alternative storage strategies through monitoring of metabolism via a broader array of metabolic parameters.

6.5 References:

Anniss A.M., Glenister K.M., Killian J.J., Sparrow R.L. Proteomic analysis of supernatants of stored red blood cell products. Transfusion. 2005;45:1426–1433.

Bennett-Guerrero E.,Veldman T.H., Doctor A., Bennett-Guerrero E., Veldman T.H., Doctor A., Telen M.J., Ortel T., Reid T.S. 2007) Evolution of adverse changes in stored RBCs. Proc Nat Acad Sci USA.104:17063–17068.

Bossi, D., Giardina, B., (1996) . Red cell physiology. Mol. Aspects Med. 17, 117–128.

Burger P,. Korsten H., De Korte D., Rombout E., Van Bruggen R., Verhoeven A.J.(2010) An improved red blood cell additive solution maintains 2,3-diphosphoglycerate and adenosine triphosphate levels by an enhancing effect on phosphofructokinase activity during cold storage. Transfusion. (11):2386-2392.

D'Alessandro A., Gevi F., Zolla L.(2011) A robust high resolution reversed-phase HPLC strategy to investigate various metabolic species in different biological models. Mol Biosyst. 7(4):1024-1032.

D'Alessandro A., D'Amici G.M., Vaglio S., Zolla L.(2010) Time-course Investigation of SAGM-Stored Erythrocyte Concentrates: from Metabolism to Proteomics. Hematologica; 125-129.

Darghouth D., Koehl B., Madalinski G., Heilier J.F., Bovee P., Xu Y., (2011) et al. Pathophysiology of sickle cell diseaseis mirrored by the red blood cell metabolome. Blood. 117(6):e57-66.

Högman C.F., Löf H., Meryman H.T.(2006) Storage of red blood cells with improved maintenance of 2,3- bisphosphoglycerate. Transfusion.;46(9):1543-1552.

Koshkaryev A, Zelig O, Manny N, Yedgar S., Barshteinet R. Rejuvenation treatment of stored red blood cells reverses storage-induced adhesion to vascular endothelial cells. Transfusion. 2009;49:2136–21.43.

Joshi A., Palsson, BO., (1989). Metabolic dynamics in the human red cell. Part I. A comprehensive kinetic model. J. Theoret. Biol. 141, 515–528

Messana I., Misiti F., Sherbini S., Giardina B., Castagnola M.(1999) Quantitative determination of the main glucose metabolic fluxes in human erythrocytes by 13C- and 1H-MR spectroscopy. J Biochem Biophys Methods.;39(1-2):63-84.

Messana I., Ferroni L., Misiti F., Girelli G., Pupella S., Castagnola M., Zappacosta B.,Giardina B..(2000) Blood bank conditions and RBCs: the progressive loss of metabolic modulation. Transfusion. ;40(3):353-360.

Muda P., Kampus P., Zilmer M., Zilmer K., Kairane C., Ristimäe T., Fischer K., Kampus P, Zilmer Teesalu R. (2003). Homocysteine and red blood cell glutathione as indices for middle-aged untreated essential hypertension patients. J Hypertens.;21(12):2329-3233.

Nishino T., Yachie-Kinoshita A., Hirayama A., Soga T., Suematsu M., Tomita M. (2009)In silico modeling and metabolome analysis of long-stored erythrocytes to improve blood storage methods. J Biotechnol.;144(3):212-223.

Percy A.K., Schmell E., Earles B.J., Lennarz W.J. (1973) Phospholipid biosynthesis in the membranes of immature and mature red blood cells. Biochemistry 12(13):2456-2461.

Pierigè F., Serafini S., Rossi L., Magnani M. January 2008. "Cell-based drug delivery". Advanced Drug Delivery Reviews 60 (2): 286–295.

.

Sackmann E. (1995) Biological Membranes Architecture and Function., Handbook of Biological Physics, (ed. R.Lipowsky and E.Sackmann, vol.1, Elsevier, 1995 Laura Dean. Blood Groups and Red Cell Antigens 1457-1465.

Schuster S., Kenanov D.(2005) Adenine and adenosine salvage pathways in erythrocytes and the role of Sadenosylhomocysteine hydrolase. A theoretical study using elementary flux modes. FEBS J 272(20):5278-5290.

Smith C.A., O'Maille G., Want E.J., Qin C., Trauger S.A., Brandon T.R., Custodio D.E., et al. METLIN: a metabolite mass spectral database. Ther Drug Monit. 2005;27(6):747-51.

Tautenhahn R., Patti G.J., Kalisiak E., Miyamoto T., Schmidt M., Lo F.Y., et al. MetaXCMS: second-order analysis of untargeted metabolomics data. Anal Chem. 2011;83(3):696-700.

Acknowledgments

First of all, I thank my tutor dott.ssa Anna Maria Timperio: her understanding, encouraging and personal guidance have provided a good basis for the present thesis. I owe my most sincere gratitude to Prof. Zolla Lello for giving me the opportunity to perform my research in his laboratory.

I wish to express my warm and sincere thanks to dott.ssa Rinalducci Sara .

I thank the members of my reading committee: Prof.Rossella Maione, dott.ssa Gabriella D'Orazi and Prof. Riccardo Negrini for careful reading and approving my manuscript.

During this work I have collaborated with many colleagues for whom I have great regard in particular for : Angelo, Barbara, Cristina, Cristiana, GianMaria,Valeria, and Valerio.

Last but not the least, I would like to thank my family: my parents and my boyfriend for their patience with me during this period.