UNIVERSITÀ DEGLI STUDI DI MILANO

SCUOLA DI DOTTORATO IN SCIENZE E TECNOLOGIE CHIMICHE

DIPARTIMENTO DI SCIENZE FARMACEUTICHE

CORSO DI DOTTORATO IN CHIMICA DEL FARMACO

CICLO XXVIII

High-resolution mass spectrometry strategies for the identification of small and large bioactive molecules.

SETTORE CHIM/08 CHIMICA FARMACEUTICA

Dott.ssa Altomare Alessandra Anna

Matricola: R10092

Tutor: Prof. Giancarlo Aldini

Coordinatore del dottorato: Prof. MARCO DE AMICI

ANNO ACCADEMICO
2014/2015

Contents

INTRODUCTION 7

1. MS in drug discovery 8

2. MS for the identification of bioactive molecules 8

3. Mass spectrometry-based proteomic 10

3.1 Sample preparation in proteomics 10

3.2 Sources and MS analyzers for proteomics 11

3.3 The different stages of a proteomic experiment 12

3.4 Computational proteomics and system biology 14

References 16

AIM OF THE WORK 17

PART I-A Health benefits of bovine colostrum revealed by an in depth proteomic analysis 20

Abstract 21

1. Introduction 22

2. Materials and methods 24

2.1 Chemicals 24

2.2 Bovine IgG purification 25

2.3 Separative methods of proteins on polyacrylamide gel 26

One-dimensional analysis (SDS-PAGE)

Two-dimensional gel electrophoresis

2.4 Sample preparation for mass spectrometry analysis 28

CPLL treatment (ProteoMiner)

In-gel protein digestion

2.5 Mass spectrometry analysis 29

2.6 Data processing using different bioinformatic tools 30

Protein identification

Generation of protein-protein interaction network

3. Results 32

3.1 SDS-PAGE analysis of bovine colostrum 32
3.2 Proteome treatment with CPLL 34
3.3 Immunoglobulin G (IgG) depletion by affinity chromatography 37
3.4 IgG depletion combined with CPLL Technology 38
3.5 Descriptive proteomic analysis 44
3.6 Functional proteomic analysis 46
4. Discussion **52**
4.1. Off-target based proteomics to elucidate the bases of its biological activities 52
4.2. Protein identification via mass spectrometry analysis 53
4.3. Functional study 55
5. Conclusions **57**
References 58
Bovine Colostrum Protein Database (Table S1) 60
PART I-B Wound healing in bovine colostrum **91**
References 106
PART II Set-up and application of an analytical approach for the quality control of purified colostrum as food supplement **109**
Abstract **110**
1. Introduction **111**
2. Materials and methods **113**
2.1 Chemicals 113
2.2 Separative methods of proteins on polyacrylamide gel 115
One-dimensional analysis (SDS-PAGE)
Two-dimensional gel electrophoresis
2.3 Bovine IgG purification 116
Affinity Chromatography
Tangential Flow Filtration
2.4 Profiling most abundant species 117

2.4.1 SEC-UV for IgG and IgM analysis *117*
2.4.2 Profiling intact proteins by HPLC-MS - TOF as MS analyser *119*
2.4.3 Profiling intact proteins by HPLC-MS - Ion trap as MS analyser *120*
2.5 Targeting bioactive proteins by ELISA assay 122
2.6 Protein Functionality 122
2.6.1 Lactoperoxidase Activity Assay - Spectrophotometric method *122*
2.6.2 Determination of Xanthine Dehydrogenase/Oxidase Activity - Fluorimetric Assay *123*
2.6.3 Determination of Xanthine Dehydrogenase/Oxidase Content - ELISA Assay *124*
3. Results **124**
3.1 Profiling intact proteins by electrophoresis and MS 124
3.1.1 1D and 2D gel electrophoresis of commercially available BC raw materials *124*
3.1.2 Profiling intact proteins by LC-ESI-MS *130*
3.2 Affinity chromatography – IgG and IgG-depleted BC fractions 133
3.3 IgG and IgM analysis by SEC-UV 135
3.3.1 Immunoglobulin G analysis *135*
3.3.2 Immunoglobulin M analysis *139*
3.4 β-lactoglobulin and α-lactalbumin analysis by LC-ESI/MS 141
3.5 Targeting bioactive proteins by ELISA assay 147
3.6. Protein Functionality 147
3.6.1 Lactoperoxidase Activity Assay - Spectrophotometric method *147*

3.6.2 Determination of Xanthine Dehydrogenase/Oxidase Activity - Fluorimetric Assay *148*

4. Discussion **151**

5. Conclusions **155**

References *157*

PART III The secrets of Oriental panacea: Panax ginseng **158**

Abstract **159**

1. Introduction **160**

2. Materials and methods **162**

2.1 Chemicals 162

2.2. Panax ginseng root treatment 163

2.3. SDS-PAGE analysis and trypsin digestion 165

2.4. Mass spectrometry analysis 166

2.5. Protein identification 166

2.6. Analysis of identified proteins 167

2.7. Analysis of proteolytic peptides obtained by simulated gastrointestinal digestion 168

3. Results **168**

4. Discussion **176**

4.1. Protein identification 176

4.2. Protein functional data 177

4.3. Analysis of proteolytic peptides 183

5. Conclusions **186**

References *188*

Panax Ginseng Protein Database (Table S1) *192*

Lisp of bioactive peptide in Panax Ginseng (Table S2) *203*

Conclusions **208**

Scientific Pubblication & Communication **209**

Introduction

1. MS in drug discovery

The first main application of mass spectrometry (MS) in pharmaceutical analysis and in particular in drug discovery started in the 60's, when it was used as an analytical tool for structure characterization of both synthetic and natural derived products and of their metabolites.

We had to wait several decades before seeing a broader application of MS, not only limited to structure characterization, but also extended to the different branches of the drug discovery program, and this was due to some instrumental achievements: (a) the advent in the 90's of ion sources (ESI and MALDI), able to ionize macromolecules, including peptides and proteins; (b) the commercial availability of high resolution mass analyzers, such as the Orbitrap in 2005; (c) the development of tandem and hybrid mass spectrometers as well as of bioinformatics tools for peptide sequence determination, even if contained in complex mixtures. [1]

Nowadays, high-resolution mass spectrometry (HR-MS), coupled to an electrospray ionization (ESI) source, is an emerging and powerful analytical tool in the drug discovery field, applied for target and hit compound identification, protein-ligand characterization, lead optimization and this through a high-throughput screening (HTS) approach. [2]

2. MS for the identification of bioactive molecules

An emerging field of MS application in drug discovery regards the identification of large bioactive macromolecules, namely proteins, naturally existing in biological systems. The rationale of such a discovery approach is to consider natural matrices such as plant extracts or animal fluids characterized by a well-established

pharmacological/clinical activity and then identifying the proteins constituting the proteome. Bioinformatics tools are then used not only to address the biological activity of the identified proteins and to establish protein-protein interaction but also to predict the bioactivity of the digested peptides arising from gastro-intestinal digestion. The aim of such an approach is to identify proteins or their clusters or peptides arising from their gastrointestinal digestion that are responsible of the pharmacological activity of the starting bioactive biological matrix. This knowledge permits not only the isolation of the bioactive fraction or its reproduction by recombinant technology, with the final goal of producing a standardized and high quality bioactive material, but also allows the set-up of suitable analytical protocols aimed at analyzing qualitatively and quantitatively the bioactive proteins.

The large-scale study of proteins, particularly their structures and functions, falls within the scope of proteomics. Proteomics has evolved substantially over the last 20 years; the term proteomics was first coined in 1997 [3] to make an analogy with genomics, the study of the genome.

The rapidly emerging field of proteomics is strictly dependent on the success of whole genome sequencing projects, since it provides an essential support for more comprehensive characterization of the genes encoded within the genomes. In this respect, MS-based proteomics has established itself as the method of choice to interpret the information encoded in genomes.

Traditional proteomics works aimed at characterizing individual proteins and understanding their functionality, and the amount of publicly available mass spectrometry-based proteomics data has grown substantially in recent years.

Proteomics could be considered as a field where researchers provide insights into cellular processes and functions by regrouping different pre-fractionation methods, mass spectrometry (MS), and bioinformatics.

3. Mass spectrometry-based proteomic

Although proteomics has enjoyed great success, it must still deal with significant technical challenges, including the dynamic nature of the proteome and its complexity. Recent development in mass spectrometry has improved the throughput in protein identification and quantification; most notably, enhancement of the sensitivity of mass spectrometers, as well as improved sample preparation and protein fractionation techniques have resulted over time in a more comprehensive analysis of proteomes.

3.1. Sample preparation in proteomics

Proteins present several striking analytical challenges, in particular the enormous dynamic range across which they are expressed. The analysis of complex biological matrices by MS is strictly dependent on off-line separation technologies that simplify such samples prior to mass analysis. Multiple separation methods can now, more easily, be combined for proteome analysis, and each approach available has its advantages and disadvantages. [4]

“No proteome can be considered *democratic*, but rather *oligarchic*”, declared P.G Righetti, to introduce the very helpful invention regarding Protein Equalizer Technology, an enrichment method used to reduce protein concentration differences in a sample. This technique comprises a diverse library of combinatorial peptide

ligands coupled to spherical porous beads. In more detail, when these beads are added to a complex proteome characterized by a widely differing protein composition and relative abundance, they are able to “equalize” the protein composition, by sharply reducing the concentration of the most abundant protein components, and at the same time by enhancing the concentration of less abundant species. This innovation introduces a useful tool in bringing the “unseen proteome” within the detection capabilities of current proteomics detection methods. [5]

Such an innovative technique has been largely applied with great success in this Ph.D work as described in the experimental part.

3.2 Sources and MS analyzers for proteomics

The production of ions of the molecules in the gas phase is an absolute requirement for successful mass spectrometric protein analysis. As reported above, the development of electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), the two soft ionization techniques capable of ionizing peptides or proteins, revolutionized protein analysis using MS. For any MS experiment, more than one consideration should be mad about the instrumentation and fragmentation method best suited to the analytical strategy chosen (top-down, bottom-up or shotgun).

The mass analyzer is central to MS experiments, and for proteomics research, four types of mass analyzer are commonly used: quadrupole (Q), ion trap (quadrupole ion trap - QIT, linear ion trap – LIT / LTQ), time-of-flight (TOF) mass analyzer, and Fourier-transform ion cyclotron resonance (FTICR) mass analyzer; they vary in their physical principles and analytical performance. In 1999, Makarov invented a new type

of mass analyzer, the Orbitrap. It is a “hybrid” instrument consisting of a LTQ coupled to a C-trap and the Orbitrap. It combines the robustness, sensitivity and MS/MS capability of the LTQ with the very high mass accuracy and the high-resolution capabilities of the Orbitrap, and, for all these characteristics, has become the instrument of choice for proteomic works. [6]

The use of mass spectrometry for proteomics should not be considered as the application of a single technique for all purposes, since no instrument or existing method allows the identification and quantification of all the components of complex protein samples in a simple, single-step operation.

The proteomic studies described in the present research work have been carried out by using an Orbitrap MS analyzer coupled to a Nano-LC system by a Nano-ESI source. The nano system together with a high-resolution MS analyzer represents a state of the art system for proteomics.

3.3 The different stages of a proteomic experiment

A typical proteomics experiment starts from the sample preparation stage, aimed at isolating proteins from the original sample (biological matrix, cell lysates or subcellular compartments). The second step consists of protein fractionation from complex mixtures using chromatographic or electrophoretic techniques.

Each fraction can be directly analyzed by mass spectrometry, and the protein analysis carried out in the form of whole-protein analysis (top-down approach). Large proteins (>~50KDa), however, could not be readily analyzed, due to the increasing complexity of the gas-phase protein ion’s tertiary structure. [6]

The second approach, based on peptide detection to infer a protein sequence, implies protein digestion (shotgun approach). The obtained peptide mixture is consequently subjected to a chromatographic separation in order to reduce its complexity before it reaches the mass analyzer. Usually the resulting peptide mixture is extensively fractionated by a nanoscale reversed-phase high pressure liquid chromatography (RP-HPLC), easily hyphenated to a mass spectrometer. [7] The mass spectrometer scans the population of eluting ions, measures their *m/z*, and in the case of tandem MS experiments proceeds to the fragmentation step. Because peptides tend to fragment into recognizable patterns, peptide characterization can be achieved from this spectrum. [8] In this respect, due to the wide dynamic range of protein species in a complex biological matrix, the shotgun approach leads to the loss of information upon the conversion of intact proteins. Since the limited dynamic range of MS analysis only allows for the peptides present at high relative abundance to be preferentially sampled, the information regarding low abundant species is not commonly obtained. Alternatively, sodium dodecyl sulfate (SDS)-PAGE or 2D-PAGE separation methods can be applied prior to MS analysis; once proteins are separated, selected spots are excised and digested in-gel with a specific protease like trypsin, followed by MS measurement of the peptide mixture obtained (bottom-up approach). In each of these methods, the identified peptides are listed on a protein “hit list”, which is the typical output of a proteomics experiment. Since protein identifications rely on matches with sequence databases, in-depth proteome characterization is actually restricted to those species for which comprehensive sequence databases are available. [9] The computational peptide identifications are known to return both correct assignment against the experimental spectra, as well as a number of false

positives, strictly dependent on the introduction of many sources of errors through all experimental steps. For instance, to solve the mass error introduced in the experimental observation of peaks, related to the mass spectrometer resolution, a mass error distribution is often used in the matching process. [8]

The above described bottom-up approach has been used for protein identification together with suitable databases obtained by the genomes.

3.4 Computational proteomics and system biology

Proteomics has recently come to the forefront offering a new perspective to biological systems and promising to transform biology and medicine; it has revolutionized biological systems studies leading to advancements in the identification of some pathologies as well as a more exhaustive explanation of a given biological system. Proteins do not usually function in isolation; most interact non – covalently with other molecules of the same protein or with other proteins. Groups of interacting molecules for a specific function make up bio-modules whose interconnections give rise to networks.

With the advent of system biology and more accurate computational modeling methodologies, it has been possible to realize biological networks aimed at elucidating both a system's structure and functional dynamics, and to predict the behavior of such a biological system as a whole. [4] System biology approaches, including genomics, transcriptomics, and proteomics have now been adapted to update natural product discovery platforms. With ever more sequenced genomes in hand, it is possible now to use bioinformatics analysis to predict the biosynthetic potential of an organism, or more generally of a biological complex matrix. [10]

In particular, MS-based proteomics, to be biologically useful, should be interfaced well with different bioinformatics tools capable of accurate data elaboration. Although, bioinformatics for proteomics has developed with great speed and complexity many open source software are available with statistical methods and filtering algorithms for proteomic data validation, their usage can be problematic for researchers without programming experience. The main application in this regard is the generation of protein-protein linkage maps. STRING and similar platforms (Cytoscape) are useful tools that can provide integration of external databases with user data sets and present these elaborations in rich graphical formats.

The current growth in the number of deep proteome studies means that the cell biology community needs consistent ways to process, analyze and share these large proteomic data sets.

References

1. Chowdhury, S.K. and B.T. Chait, *A mass spectrometric technique for detecting and identifying by-products in the synthesis of peptides.* Anal Biochem, 1989. **180**(2): p. 387-95.
2. Hofstadler, S.A. and K.A. Sannes-Lowery, *Applications of ESI-MS in drug discovery: interrogation of noncovalent complexes.* Nat Rev Drug Discov, 2006. **5**(7): p. 585-95.
3. James, P., *Protein identification in the post-genome era: the rapid rise of proteomics.* Q Rev Biophys, 1997. **30**(4): p. 279-331.
4. Moore, J.B. and M.E. Weeks, *Proteomics and systems biology: current and future applications in the nutritional sciences.* Adv Nutr, 2011. **2**(4): p. 355-64.
5. Righetti, P.G., et al., *Protein Equalizer Technology : the quest for a "democratic proteome".* Proteomics, 2006. **6**(14): p. 3980-92.
6. Han, X., A. Aslanian, and J.R. Yates, *Mass spectrometry for proteomics.* Curr Opin Chem Biol, 2008. **12**(5): p. 483-90.
7. Van Oudenhove, L. and B. Devreese, *A review on recent developments in mass spectrometry instrumentation and quantitative tools advancing bacterial proteomics.* Appl Microbiol Biotechnol, 2013. **97**(11): p. 4749-62.
8. CANNON, W., ***COMPUTATIONAL PROTEOMICS: HIGH-THROUGHPUT ANALYSIS FOR SYSTEMS BIOLOGY***, B.-J. WEBB-ROBERTSON, Editor. 2007: Pacific Symposium on Biocomputing. p. 403-408.
9. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.
10. Bumpus, S.B., et al., ***A Proteomics Approach to Discovery of Natural Products and Their Biosynthetic Pathways***. 2009: Nat Biotechnol. p. 951–956.

Aim of the work

The aim of the work is to set-up and apply an analytical approach based on proteomics for the identification of bioactive macromolecules contained in complex bioactive matrices, such as natural extracts and biological fluids. The approach starts with proteome characterization, followed by bioinformatics analysis aimed not only at addressing the biological activity of the identified proteins and at establishing protein-protein interaction, but also at predicting the bioactivity of the digested peptides arising from gastro-intestinal digestion. We believe that the knowledge generated by such an approach permits not only the production of a safe and high quality/standardized bioactive matrix, obtained by isolating the bioactive fraction or by recombinant technology, but also the set-up of suitable analytical protocols aimed at analyzing qualitatively and quantitatively the bioactive proteins.

The above mentioned strategy was applied for the identification of bioactive macromolecules contained in bovine colostrum and panax ginseng, two natural sources of nutraceutical and pharmaceutical interest. The goal was reached through an integrated analytical strategy based on a proteomic and bioinformatics approach and by applying innovative techniques for sample preparation, including affinity chromatography and the combinatorial peptide library technology (CPLL or PM, ProteoMiner), aimed at reducing the dynamic range of protein concentrations.

Bovine colostrum (BC) is a mixture of diverse components, including proteins, fat, lactose, vitamins and minerals, providing the first nutritional components to newborn calves. Certain beneficial effects of BC may be shared across species and it could also benefit humans. Based on the presence of several bioactive components, BC has been included in several health products that are claimed to improve the immune system and treat gastro-intestinal infection as well as to maintain skin and mucous

membrane integrity, favoring their healing. Most of the beneficial effects of BC have been ascribed to protein components, and so far explained on the basis of the known components or by searching, through an intrinsically limited on target approach, only putative components whose activity would explain the effect of BC. In the first part of the work, an innovative proteomic approach was used to generate an exhaustive protein data bank of bovine colostrum, including the minor protein components. By using a bioinformatics approach, a cluster of bioactive proteins effective for treating wound healing was identified. Such an approach also permitted the exploration of novel biological functions of colostrum through the identification of novel bioactive proteins not yet reported as colostrum components. An integrated analytical strategy was then set-up for the quality control of the bioactive protein components of bovine colostrum.

Asian ginseng (Panax ginseng C. A. Meyer), one of the most highly regarded of herbal medicines in the Orient, was selected as the second matrix to be studied; most of the health effects of such plant have been ascribed to ginsenosides. The aim of our research was to investigate the bioactive protein components and peptides arising from their gastrointestinal digestion, and to do this a comprehensive proteomic analysis was carried out together with a bioinformatics approach for protein/peptide bioactive constituents.

Part I-A

Health benefits of bovine colostrum revealed by an in depth proteomic analysis

Abstract

Bovine colostrum (BC), the initial milk secreted by the mammary gland immediately after parturition, is widely used for several health applications. We here propose an off-target method based on proteomic analysis to explain at molecular level the health benefits of BC. The method is based on the set-up of an exhaustive protein data bank of bovine colostrum, including the minor protein components, followed by a bioinformatics functional analysis. The proteomic approach based on ProteoMiner technology combined to a highly selective affinity chromatography approach for the immunoglobulins depletion, identified 1786 proteins (medium confidence; 634 setting high confidence), which were then clustered on the basis of their biological function. Protein networks were then created on the basis of the biological functions or health claims as input. A set of 93 proteins involved in the wound healing process was identified. Such an approach also permits the exploration of novel biological functions of BC by searching in the database the presence of proteins characterized by innovative functions. In conclusion an innovative approach based on an in depth proteomic analysis is reported which permits an explanation of the wound healing effect of bovine colostrum at molecular level and allows the search of novel potential beneficial effects.

1. Introduction

Bovine colostrum (BC), the initial milk secreted by the mammary gland immediately after parturition, is a mixture of diverse components, such as proteins, fat, lactose, vitamins and minerals, providing the first nutritional components to the new-born calves [1, 2].

Besides the nutritional effect, colostrum has a fundamental biological function in calves due to the presence of a complex mixture of proteins that actively participate in the protection of the neonate against pathogens and other postpartum environmental challenges (passive immune transfer) [3-6].

Certain effects of BC may be species specific, whereas other effects may be shared across species. Hence, the unique nutritional and biological activities of BC that benefit neonatal calves may also benefit humans and be effective for the treatment of some human pathologies [5, 7]. Most of the beneficial effects of BC have been ascribed to the protein components. For instance immunoglobulins, that represent the main protein class of colostrum, have direct antimicrobial and endotoxin-neutralizing effects throughout the alimentary tract, playing a direct role in the defence of the gastro-intestinal (GI) tract [8-11]. BC also has growth-promoting effects on human epithelial cells, which is attributable to the combination of several growth factors, including insulin-like growth factors (IGFs), basic fibroblast growth factors and platelet-derived growth factor [5, 12, 13].

Based on the beneficial effects, as reported above, BC has been included in several health products making various claims such as supporting the immune system, treating gastro-intestinal infection and for the maintenance of the skin and mucous

membranes integrity, favouring their healing [9, 13] and also for the prevention and treatment of vaginal dryness [14].

Although several studies on the health applications have so far been reported, an in depth knowledge of the biological effects of BC at molecular level is still missing. Most of the beneficial effects of BC are explained on the basis of the known components or by searching, through an on target approach, putative components whose activity would explain the effect of BC [2, 15, 16]. This on target method has several limits and in particular it permits the identification of only known or putative components proposed on the basis of their biological activity but not of unknown compounds.

We here propose an off-target method based on proteomic analysis. The method is based on the set-up of an exhaustive protein data bank of bovine colostrum, including the minor protein components. The proteins are then clustered on the basis of their biological function and protein networks created on the basis of the biological functions or health claims as input. Such an approach also permits the exploration of novel biological functions of colostrum by searching in the database for the presence of proteins characterized by innovative functions.

Hence, an in depth proteome analysis of BC is required to generate such a workflow, aimed to better investigate the health benefits of BC and to find novel applications.

From an analytical point of view, the complete BC proteome resolution represents a difficult challenge because, similarly to other body fluids, the proteins are contained in a large dynamic range, spanning up to 12 orders of magnitude. This hinders in-depth proteomic analyses due to the lack of sensitivity in protein detection, primarily because of the influence of the dominating proteins (immunoglobulins and caseins). [17] [18]

Up to now several approaches have been reported and, to our knowledge, that proposed by Nissen et al., based on an extensive fractionation of colostrum prior to 2D-LC-MS/MS analysis, which brought to the identification of 403 proteins, is, to date, the most extensive list of bovine colostrum proteins available in the literature [18].

We here report a strategy based on *the combinatorial peptide library technology* (CPLL or PM, ProteoMiner), aimed at reducing the dynamic range of protein concentrations, thus maintaining representatives of all proteins within the original sample under analysis [19]. The *CPLL technology* was also combined with a highly selective affinity chromatography approach, based on protein G stationary phase that binds specifically the Fc portion of the immunoglobulin and allows a removal of about 95% of IgG.

2. Materials and methods

2.1. Chemicals

ProteoMine (combinatorial hexapeptide ligand library beads, CPLL), Laemmli buffer, 40% acrylamide/Bis solution, N,N,N',N'-tetramethylethylenediamine (TEMED), molecular mass standards and electrophoresis apparatus for one-dimensional electrophoresis were supplied by Bio-Rad Laboratories, Inc., Hercules CA. β-mercaptoethanol, dithiothreitol (DTT), ammonium persulfate, 3-[3-cholamidopropyl dimethylammonio]-1-propanosulfonate (CHAPS), acetonitrile (ACN), trifluoroacetic acid (TFA), sodium dodecyl sulphate (SDS), iodoacetamide (IAA), formic acid (FA) and all other chemicals used throughout the experimental work were current pure analytical grade products and purchased from Sigma-Aldrich Corporation, St Louis,

MO. Water and acetonitrile (OPTIMA LC/MS grade) for LC/MS analyses were purchased from Fisher Scientific, UK. Complete protease inhibitor cocktail tablets and sequencing grade trypsin were supplied by Roche Diagnostics (Basel, CH).

Decaseinated and defatted bovine colostrum was supplied by the company Advances in Medicine (AIM, Bologna, IT). According to the supplier, bovine colostrum was collected from Holstein cows until the fifth hour after birth and immediately frozen at -20°C. After a suitable dilution with demineralized water, the suspension obtained was introduced into a reactor (controlled continuous stirring), where it was heated at 35-36°C for about 30 minutes. The suspension was then subjected to the skimming step, in a special cream centrifugal separator, and then caseins were removed by precipitation by adjusting the pH at their isoelectric point. The product was then ultrafiltered on a ceramic membrane with cut-off 0.05 µm and the permeate concentrated (about 10/20 times) and dialyzed with demineralized water. The following steps consisted of the clarification through 0.45 and 0.22 µm filters followed by lyophilisation.

2.2 Bovine IgG purification

Affinity Chromatography – IgG depleted BC was prepared by removing the IgG fraction by affinity chromatography. The affinity column was prepared by packing 400 mL of Protein G Sepharose 4 Fast Flow resin (GE Healthcare) in a column support HiScale 50 (GE Healthcare) which was connected to an FPLC system (ÄKTAprime plus, GE Healthcare).

The chromatographic purification started by equilibrating the column with 5 volumes (5 x chromatographic bed volume) of buffer A (Binding Buffer: 20 mM sodium

phosphate, pH 7), and then the sample was loaded at a flow rate of 20 ml min^{-1}. The absorbance of the flow-through was monitored at 280 nm and the different fractions were automatically collected when the absorbance was > 0.01 A.U. . The subsequent step consisted of recovering the IgG fraction by eluting the column with 100% of Elution Buffer (1 M glycine hydrochloride pH 2.5).

Tangential Flow Filtration – The collected fractions were mixed and subjected to concentration and desalting using hollow fibers cross flow filtration cartridges with 3000 NMWC (Nominal Molecular Weight Cutoff) and a surface area of 650 cm^2 (GE Healthcare) coupled to a tangential flow filtration system equipped with a peristaltic pump essential to keep the flow recirculation continuous (Kross Flo®- Tangential flow Filtration System Research III). The IgG depleted fraction was concentrated 20/30 times, dialysed with 5 volumes and then lyophilized. Glass vials filled with 3mL dialyzed solution and partially stoppered with lyo-stoppers were placed on a pre-cooled shelf at -20 °C inside the ice condenser chamber of an Alpha 1-4 LSC Christ freeze dryer (G). Once frozen (30 min), the samples were dried under the following condition. Main drying: the shelf temperature was ramped from -20 °C to 15 °C in 8 h; afterwards the shelf temperature was kept constant over a 10 h period. The chamber pressure was set at 1.030 mBar; final drying: the chamber pressure was set at 0.001 mBar and the shelf temperature ramped to 25 °C in 2 h and then kept constant for 2.5h.

2.3 Separative methods of proteins on polyacrylamide gel

One-dimensional analysis (SDS-PAGE) - Protein separation was performed under both reducing and non reducing conditions; aliquots of 10 µL of samples containing 20-25 µg of proteins were mixed with 10 µL of Laemmli sample buffer (containing 50

mM DTT in case of reducing conditions) and heated at 95°C for 5 minutes. Samples and the standard proteins mixture (Precision Plus Protein Standards) were loaded on precast gels (Any KD Mini Protean TGX), then placed in the electrophoresis cell (Mini-PROTEAN Tetra) and run at 200 V until the samples reached the front of the gel. Gels were stained using Coomassie blue (Biosafe G250, Bio-Rad). Images were acquired by the GS800 densitometer and analyzed by software Quantity One (all Bio-Rad).

Two-dimensional gel electrophoresis **-** 150 µg of each sample were solubilised in the denaturing buffer (7 M Urea, 2 M thiourea, 40 mM Tris, 3% CHAPS) and incubated for one hour at room temperature with 5 mM TCEP in order to reduce the protein disulphide bonds. After spiking with DESTREAK (150 mM), Ampholine (0.5%) and Blue Bromophenol, samples were loaded on IPG strips (Bio-Rad 7 cm, pH gradient 3-10); after incubation (4 hours) strips were laid in the strip housing for isoelectric focusing (IEF, I12 Protean IEF Cell-Bio-Rad) and covered with Mineral Oil (Bio-Rad) to prevent proteins oxidation and strip dehydration. The instrument was set to follow the voltage gradient reported in the Table 1.

Voltage (V)	Time (h)	Voltage Ramping Method (Slope)
100	1	Rapid
1000	1	Rapid
1000	1	Linear
5000	1	Rapid
5000	30	Linear

Table 1- Isoelectric focusing voltage gradient.

Once the IEF was completed, strips were incubated with 2.5 mL of equilibration buffer (6 M Urea, 2% SDS, 0.05 M Tris-HCl pH 8.8 and 20% glycerol) on gentle shacking for 10 minutes. Proteins separation based on their molecular mass was obtained by one-dimensional electrophoresis (SDS-PAGE) using precast gel (Any KD ™ Mini Protean® TGX ™). The strips were laid on the top of the gel, orienting them so that pI followed an ascending order (from 3 to 10) from left to right, and coated with a layer of agarose (0.5%). The electrophoretic device was assembled as described above and the electrophoretic run was carried out at 200 V for 30 min. After electrophoresis, gels were washed, stained and acquired as image as described in the previous paragraph.

2.4 Sample preparation for mass spectrometry analysis

CPLL treatment (ProteoMiner) - 1.5 g of lyophilized colostrum (IgG depleted and non depleted) was dissolved in 50 mL of solubilization buffer (phosphate buffer 10 mM and 25 mM KCl pH≈7) in order to obtain a concentration of 3% w/v. The pH of the solutions was adjusted at pH 2.2, 4.0, and 9.0. The solutions at the different pH values (2.2, 4.0, 7.0, 9.0) were then incubated overnight with 100 µL of beads (ProteoMiner, Protein Enrichment Kit) under continuous stirring provided by a rotary mixer. All the beads were then recovered by filtering the solutions with Micro Bio-Spin chromatographic columns (Bio-Rad) connected to an Erlenmeyer flask coupled with a vacuum pump. Proteins bound to the beads, now trapped in the mesh of the chromatography column, were eluted after two consecutive bead incubations with 4% sodium dodecyl sulphate (100 µL first elution, 80 µL second elution) at 99°C. After each incubation the samples were centrifuged at 4000 rpm for two minutes. The first

and the second elution, maintained in separate eppendorf tubes, contained the enriched proteins.

In-gel protein digestion – After gel separation by mono/bi-dimensional electrophoresis, bands/spots of interest were excised and washed with 200 µL of MilliQ H_2O. Gel pieces were destained using 100 µl of destaining solution (50% 25 mM NH_4HCO_3/50% ACN) for 10 min and then with 100 µl of acetonitrile for other 10 min.

100 µl of reducing solution (10 mM DTT in 50 mM NH_4HCO_3) was added to gel pieces and incubated for one hour at 56 °C; after the removal of the solution and a washing step in Digestion buffer (50 mM ammonium bicarbonate), 100 µl of alkylating solution (55mm IAA in 50 mM NH_4HCO_3) were added to gel pieces and incubated in the dark for 45 minutes. After the removal of the alkylating solution, gel pieces were washed in Digestion buffer. Trypsin (sequencing grade, Roche) was diluted in Digestion buffer and added to gel pieces (about 1 µg of trypsin to each sample). After an overnight incubation at 37°C, the tryptic mixtures were acidified with 1 µl of formic acid and eluted using 70 µl of 3% TFA / 30% ACN in H_2O MilliQ, which was added and incubated for 10 minutes at 37°C. This step was repeated once with the same solution and two times using acetonitrile; collected peptides were pooled for each sample. The peptide mixture was dried using a Speed Vac (Martin Christ) and stored at -20°C.

2.5. Mass spectrometry analysis

The peptide mixtures obtained by in-gel digestion were solubilized in 20 µl of buffer A (0.1% HCOOH). Five µl of each sample were injected on a C18 column (HALO

PicoFrit, 75 mM x 10 cm, 2.7 µ particles, 100 Å pore size, New Objective, USA) by a nano-chromatographic system (UltiMate 3000 RSLCnano System, Thermo Scientific) operating at a constant flow rate of 0.4 µl/min. The separating gradient ramped linearly from 1% acetonitrile to 35% acetonitrile in 90 min. The eluting peptides were on-line sprayed in a LTQ-Orbitrap XL mass spectrometer by a nano-ESI source (all Thermo Scientific) set as follows: positive ion mode, spray voltage 1.8 KV; capillary temperature 220°C, capillary voltage 35 V; tube lens offset 120 V. The mass spectrometer operated in data-dependent acquisition mode (DDA) to acquire both full MS and MS/MS spectra. Full MS spectra were acquired in a "profile" mode, by the Orbitrap (FT) analyzer, between 300-1500 *m/z*, AGC target = $5x10^5$ and resolving power = 60000 (FWHM at 400 m/z). Tandem mass spectra MS/MS were acquired by the Linear Ion Trap (LTQ) in CID mode, automatically set to fragment the nine most intense ions in each full MS spectrum (exceeding $1x10^4$ counts) under the following conditions: centroid mode, isolation width = 2.5 *m/z*, AGC target = $1x10^4$ and normalized CE = 35 eV. Dynamic exclusion was enabled for 45 seconds for ions already observed 3 times in 30 seconds. Charge state screening and monoisotopic precursor selection was enabled, singly and unassigned charged ions were not fragmented. Instrument control and spectra analysis were provided by the software Xcalibur 2.0.7 and Chromeleon Xpress 6.80.

2.6. Data processing using different bioinformatics tools

Protein identification - Proteins was identified using the software Proteome Discoverer 1.3 (Thermo Scientific) using the Sequest algorithm. The Bos Taurus database used for data analysis was downloaded from UniProt (released on October

16, 2013). Default settings were used for protein identification, with the following exceptions: mass range = 350 - 5000 Da, signal/noise threshold = 5, precursor mass tolerance = 5 ppm, fragment mass tolerance = 0.5 Da. Trypsin was set as the proteolytic enzyme, cysteine carbamidomethylation (+57.021 Da) and methionine oxidation (+15.995 Da) were set as variable modifications. Proteins were searched both in the Bos Taurus and in the Decoy Database. The false discovery rate of protein identification was set to FDR = 0.01 (Strict) or FDR = 0.05 (Relaxed). Two lists of proteins were obtained from Proteome Discoverer: the first list was generated by considering a medium-confidence identification (FDR = 0.05); the second one was generated by setting a high-confidence identification (FDR = 0.01).

These proteins were clustered on the basis of their cellular localization, molecular functions or biological processes by using the Annotation tool, that allows automatic information retrieval from the website Protein Center.

Generation of protein-protein interaction network - STRING database (http://string-db.org/) aims at providing a comprehensive understanding of biological processes, known or predicted, in which a group of proteins is involved, within a given organism. Functional analysis aimed at creating interaction and association protein networks was conducted subjecting the identified proteins (gene IDs), identified in all the colostrum samples analyzed, in STRING v.9.1, selecting Bos Taurus as organism reference.

In the final interactive network view proteins are represented by nodes and interactions by connecting lines, continuous for direct interactions (physical), discontinuous for indirect (functional). Each connecting line is supported by at least

one bibliographical reference or by canonical information present in the dataset of STRING. The confidence value (score) was set to 0.4 (mean level).

Based on functional annotations, including Gene Ontology, KEGG, Pfam and InterPro, proteins involved in selected biological processes were highlighted. On the basis of BC applications mentioned above, networks of protein interaction were enriched on the basis of the following biological processes: wound healing, bactericidal/bacteriostatic, neurogenesis, regulation of bone growth.

3. Results

3.1 SDS-PAGE analysis of bovine colostrum. Figure 1 displays the SDS-PAGE profile of BC in reducing and non-reducing conditions. As expected, the protein profile of native colostrum (non reducing conditions) shows a single, broad band of immunoglobulins centred at about 150 kDa, which disappears under reducing conditions in favour of two bands at about 50 and 25 kDa, representing the heavy and light IgG chains, respectively. Not many more bands can be detected, except for a few rather intense zones, which, upon cutting the gel into ten segments along the electrophoretic track, could be identified via MS as: α-lactalbumin (14.1 kDa), ß-lactoglobulin (19.9 kDa), serum transferrin (77.7 kDa) and α_2-macroglobulin (167.5 kDa), i.e. the classical set of high-abundance species normally found in just about any type of milk of animal origin (although in colostrum IgGs alone represent about 80% of the total protein mass).

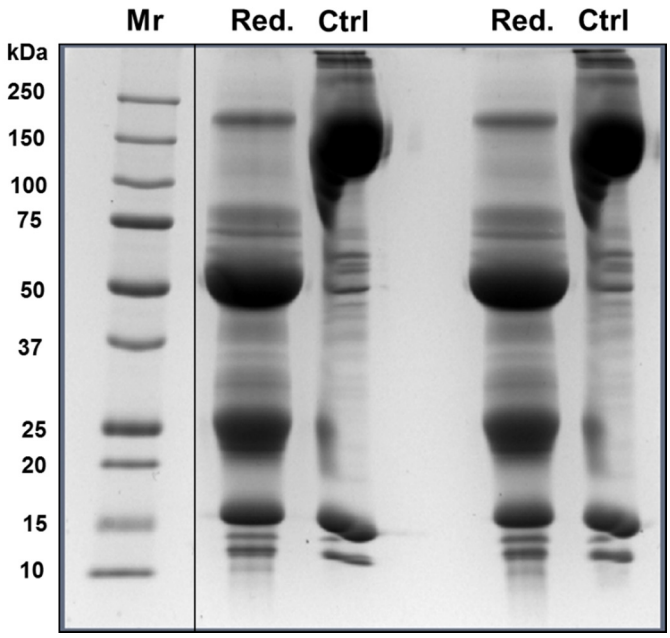


Figure 1 - SDS-PAGE profiling of colostrum proteins. Mr: molecular mass ladder. Samples: Ctrl: control; red: reduced with 50 mM DTT. The two samples to the left have been dialyzed so as to remove sugars; the two to the right are untreated. Staining with colloidal Coomassie blue.

The two lines are duplicated since they represent dialyzed vs. non-dialyzed colostrum. Although the pattern is superimposable, dialysis removes different sugars present in the fresh sample that, if not eliminated, upon storage form extensive adducts, producing, after Amadori rearrangement, stable glycated forms that would greatly increase the sample heterogeneity and produce a yellowish tinge, in turn diminishing the biological activity of originally active proteins [20] (all data presented hereafter were obtained using dialyzed colostrum).

3.2 Proteome treatment with CPLL. Given the impossibility of exploring to a deeper extent the colostrum proteome and so visualizing the low-abundance species whose signal is masked by the high-abundance proteins (especially IgGs), we resorted to sample treatment with combinatorial peptide ligand libraries (CPLL or PM, ProteoMiner), a technique well known to amplify the signal of less abundant species. The "democratic" extraction which occurred is shown in Figure 2, which shows the SDS-PAGE profiling of the control material (Starting – untreated sample) versus the sample captured with CPLLs at four different pH values (pH 2.2 with synthesis beads and pH 4.0, 7.0 and 9.0, with beads from ProteoMiner Kit).

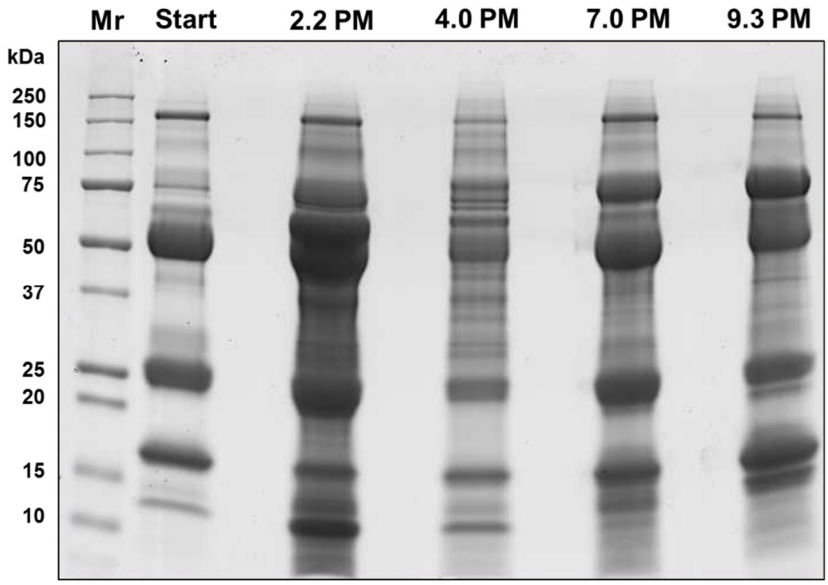


Figure 2 - SDS-PAGE profiling of colostrum proteins. Mr: molecular mass ladder. Samples: Start (control, untreated colostrum), 2.2 PM, 4.0 PM , 7.0 PM and 9.3 PM (colostrum captured with CPLL beads at pH 2.2, 4.0, 7.0 and 9.3, respectively). In these last 4 lanes notice the marked reduction of the 50 and 25 kDa chains of IgGs and the sudden appearance of a multitude of fine bands in the entire Mr region. Staining with colloidal Coomassie blue.

Two phenomena are immediately apparent: first of all, the massive amounts of heavy and light IgG chains present in the control are substantially reduced in the captured samples; concomitantly, a very large number of fine protein zones appears and carpets essentially the entire electrophoretic track between 10 to 250 kDa (and even higher).

Just to give a visual impact of the large increment of novel protein species detected, the two-dimensional (2D) maps of Figure 3 show the pI/Mr distribution of the visible species in the control (first panel) versus the eluate from CPLL beads after pH 4.0 capture (second panel). Many more spots are visible, especially in the acidic map region and in the Mr 10-40 kDa range.

Although the CPLL Technology was found very effective in normalizing the protein abundance, the isoforms of the heavy and light IgG chains respectively were still predominant and for this reason, we combined the CPLL to IgG depletion by affinity chromatography.

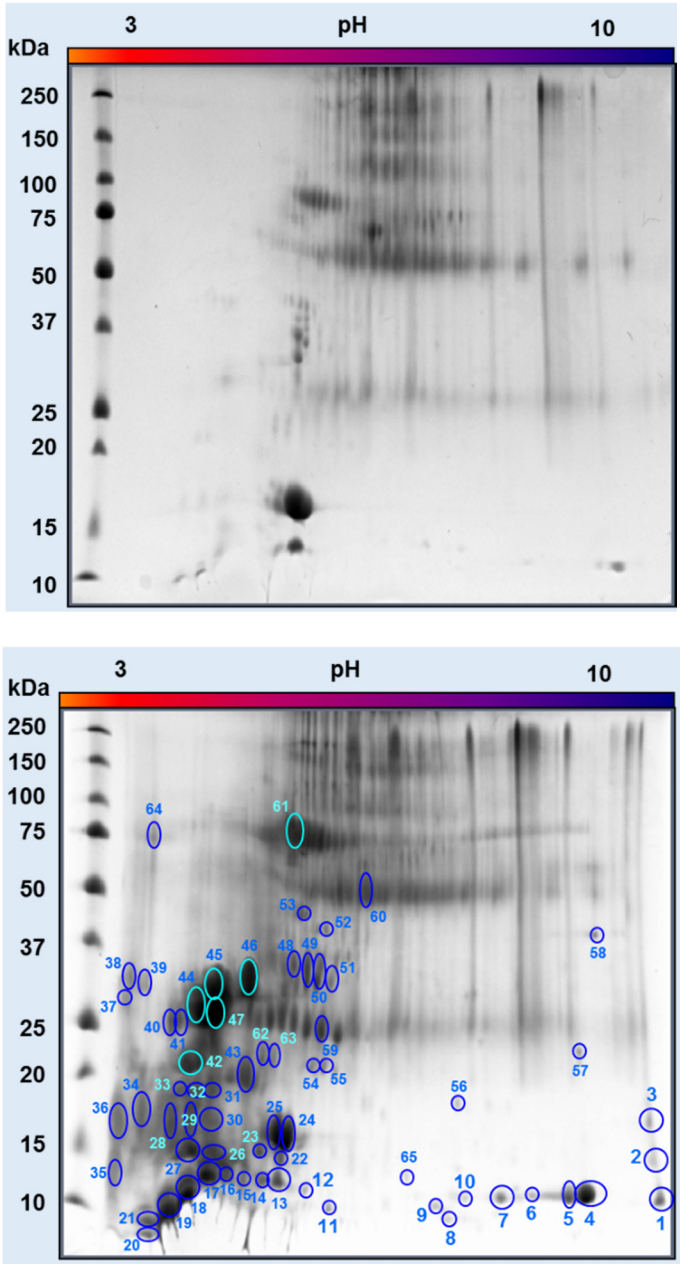


Figure 3 - Two-dimensional maps of control (first panel) and CPLL captured colostrum at pH 4.0 (second panel). Staining with colloidal Coomassie blue.

3.3 Immunoglobulin G (IgG) depletion by affinity chromatography. The affinity chromatogram displayed two peaks, the first, not retained, corresponding to the non-immunoglobulin protein fraction and the second to IgG (data not shown). According to the relative abundance of IgG in colostrum (80%), the second peak was the most abundant, followed by a peak attributed to the immunoglobulin aggregates at high Mr. The electrophoretic patterns obtained for each collected fraction (data shown in the second part of the work), indicate a good depletion of immunoglobulins, whose characteristic bands at 160-50-25 KDa, are very weak in the aliquots eluted within the first peak, and clearly much more intense in the fractions of the second peak corresponding to the immunoglobulins.

The efficiency of the depletion process was then confirmed by measuring the IgG content by enzyme immunoassay (ELISA assay) of the fractions collected and liophylized; the IgG content in the depleted fraction was less than 5% of the total amount of IgG (100%) as determined before the depletion. Data were confirmed by Western blot analyses (Figure 4), showing an intense dark spot referred to the large abundance of IgG in the untreated sample that almost disappears in the depleted colostrum (central lane) to be well evident in the IgG fraction.

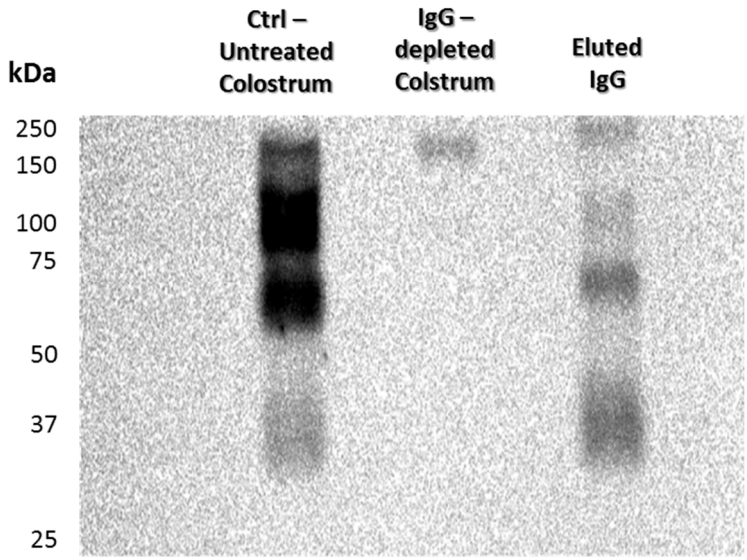


Figure S1- Western blotting analysis using monoclonal antibody anti-bovine IgG. The marked reduced IgG content in the purified Bovine colostrum by Affinity Chromatography can be observed. Samples: Ctrl (untreated colostrum), IgG depleted colostrum and eluted IgG. Position of the molecular marker (in kDa) is indicated on the left.

3.4 IgG depletion combined with CPLL Technology. IgG Depleted colostrum was treated with combinatorial peptide ligand libraries (CPLL or PM, ProteoMiner) under different pH conditions (pH 4.0 ,7.0, 9.3, 2.2) as described in Material and Methods section. Figure 5 shows the SDS-PAGE analysis, in which it is possible to appreciate the enrichment effect and in particular that the intensity of the characteristic bands of the IgG subunits (50 and 25 kDa) is extremely decreased while a very large number of fine protein zones appears.

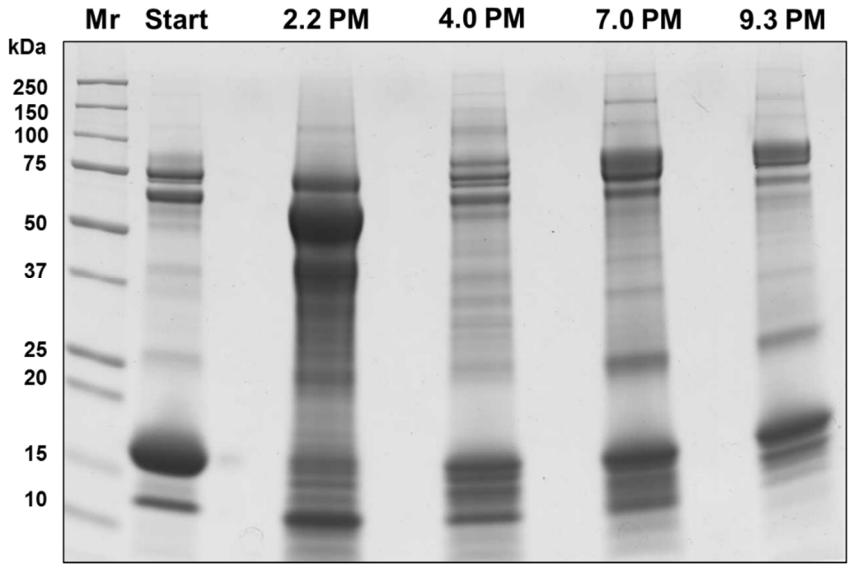


Figure 4 - SDS-PAGE profiling of colostrum proteins. Mr: molecular mass ladder. Samples: Start (IgG depleted colostrum), 2.2 PM, 4.0 PM , 7.0 PM and 9.3 PM (IgG-depleted colostrum captured with CPLL beads at pH 2.2, 4.0, 7.0 and 9.3, respectively). In the first lane chains of IgGs are clearly reduced; while in the last 4 lanes notice the sudden appearance of a multitude of fine bands in the entire Mr region. Staining with colloidal Coomassie blue.

The 1D SDS-PAGE was used for separation of the proteins prior to the protein identification by MS. Two distinct experiments, in which peptides extracted from a PAGE slices, were performed.

The Venn diagrams on Figure 6 show the number of proteins identified by the CPLL treaments at different pH. It can be seen that only 156 proteins are common to all three different pH treatments and that each capture at each individual pH value gives a unique contribution in terms of detected proteins.

Figure 6 - Venn diagrams showing the separate contributions to protein discovery of the three different CPLL treatments at pH 4.0, 7.0 and 9.3. Yellow: Colostrum captured with CPLL beads at pH 4.0 (421 proteins); blue: Colostrum captured with CPLL beads at pH 9.3 (435 proteins); Orange: Colostrum captured with CPLL beads at pH 7.0 (672 proteins).

The largest contribution, though, appears to be the one produced by the pH 7.0 treatment, which allows the identification of 326 unique proteins, followed by the pH 9.3 (169 ID) and finally by the pH 4.0 capture (147 unique ID). It is of interest, here, to compare what can be found in untreated colostrum vs. the different CPLL treatments.

This is shown in the Venn diagrams of Figure 7, in which the contribution of yet another capture at pH 2.2 (Figure 7 - panel A) is shown (this last capture is different from the other three, because here the mechanism of harvesting is largely induced by hydrophobic interaction, whereas at pH 4.0, 7.0 and 9.3 the main operative mechanism is based on ionically-driven interactions).

In general control colostrum allows identification of an average of 224 unique proteins, whereas, as shown in the four panels, the four CPLL treatments allow the detection of many more species. The newly found proteins obtained in the four captures can thus be summarized: 420 at pH 2.2 (panel A), 299 at pH 4.0 (panel B), 539 at pH 7.0 (panel C) and 330 at pH 9.3 (panel D). Interestingly, in all cases, an average of only 20% detected proteins are in common between the control colostrum and each of the four treatments.

Figure 7 - panel E summarizes the overall exploration of enriched colostrum proteome, the two Venn diagrams showing again, on the left side (small circle) the total IDs in control colostrum, while the large circle on the right represents the overall findings in all CPLL treatments, by subtracting the various redundant proteins present in more than one CPLL capture. It can be appreciated that, when comparing what can be seen in the control (224 proteins) vs. what has been found upon all four treatments (1167 proteins) a 5.5 fold increment of (550%) in protein visibility is obtained, this huge increase likely being due to low- to very-low abundance proteins present in colostrum, whose signal had been largely masked by that handful of proteins (5-6 major components), which alone constituted >95% of the protein mass in colostrum.

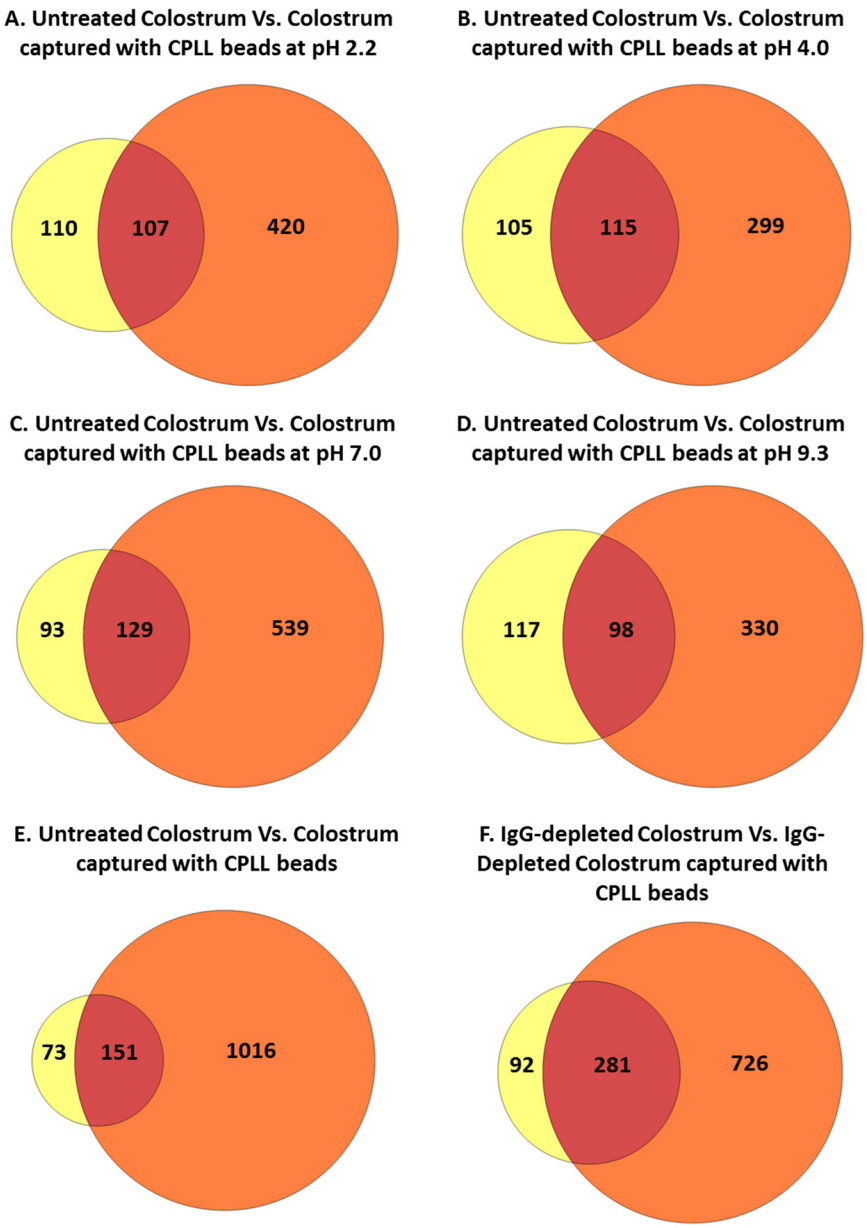


Figure 7 - Venn diagrams showing the separate contributions to protein discovery of four different CPLL treatments at pH 2.2 (panel A), 4.0 (panel B), 7.0 (panel C) and 9.3 (panel D) as compared to the proteome discovery of untreated colostrum samples; panel E summarizes the overall findings in all CPLL treatments against the total IDs in control colostrum; panel F shows the effect of IgGs depletion by affinity chromatography; against an average of about 371 IDs in the depleted colostrum, an additional 726 species are detected (representing the sum of all four CPLL treatments).

In order to investigate more deeply the colostrum proteome, we combined the depletion of the IgGs by affinity chromatography with the CPLL capture. The overall IDs are given in the Venn diagrams of Figure 7 – panel F: here too, against an average of about 371 IDs in the depleted colostrum, additional 726 proteins are detected (representing the sum of all four CPLL treatments). Interestingly, here too the overlap between the two samples is 20-25%.

The *merging tool*, available among the Proteome Discoverer options, automatically deletes the redundancies from different tabular reports; it was used to generate the final list of proteins identified in the different samples of bovine colostrum, which includes 1786 ID. This number was reduced to 634 proteins upon setting high peptide confidence as filter (Peptide RANK 1) [Table 2]. This list of proteins was used as input for functional analyses, as described above.

SAMPLE	**N° EXCLUSIVE PROTEINS**	**N° SHARED PROTEINS**
DIALYZED COLOSTRUM	**73**	**151**
DIALYZED COLOSTRUM + CPLLs	**1016**	
DIALYZED COLOSTRUM	**117**	**101**
IgG-DEPLETED - DIALYZED COLOSTRUM	**257**	
IgG-DEPLETED -DIALYZED COLOSTRUM	**92**	**281**
IgG-DEPLETED -DIALYZED COLOSTRUM + CPLLs	**726**	

Table 2 – Comparison of the number of proteins identified in all the fractions analyzed: dialyzed colostrum, colostrum treated with CPLLs, depleted colostrum, colostrum depleted and further treated with CPLLs.

3.5 Descriptive proteomic analysis. The proteins identified by mass spectrometry were subdivided into specific clusters, grouped on the basis of their cellular localization, molecular function and the biological process in which they are involved. Regarding the biological process (Figure 8 – panel A) the largest clusters include: proteins involved in metabolic processes (19.5%) and in biological process regulation (14.4%) while fewer proteins are involved in cellular processes including cellular division, proliferation and differentiation. A group of proteins accounting for 11.3% resulted with no-annotation.

Regarding the molecular function (Figure 8 – panel B), 27.03% of proteins are involved in protein binding, while fewer proteins are involved in metallic ions, nucleotides, DNA, RNA binding. It is also possible to find a great percentage of protein with catalytic activity. Whereas, with regard to intra/extra cellular localization (Figure 8 – panel C), we surprisingly found a high percentage of proteins localized in the cellular membrane (18.64 %), followed by those localized in the cytoplasm (18.34%) and extracellular proteins (17.19%).

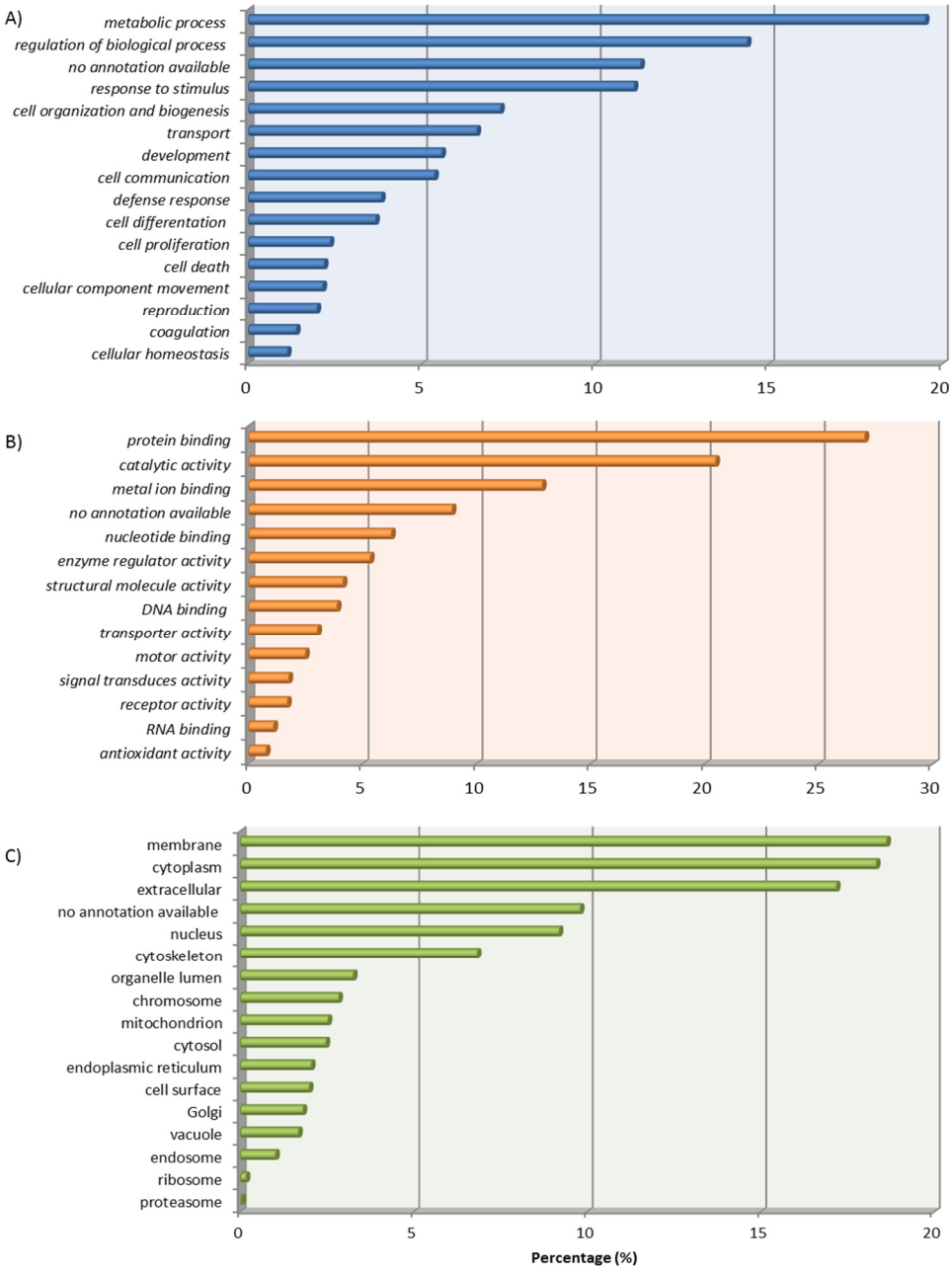


Figure 8 – BC protein distribution based on the biological process (panel A), molecular function (panel B), and cellular localization (panel C), obtained by using enrichment tool by Proteome Discoverer.

3.6 Functional proteomic analysis. The first goal of the functional analysis was the generation of a protein-interaction network. The interaction network, consisting of a complex grid based on bibliographic evidence or canonical information within STRING database, was integrated with functional information obtained from others databases, namely Gene Ontology, KEGG, Pfam e InterPro, in order to create different biological process clusters. Moreover, the enrichment tool of STRING allowed us to highlight protein groups based on targeted biological process, molecular function, cellular component localization such as cytoskeletal binding, calcium ion binding, growth factor activity and antioxidant activity (Figure 8). Besides the above-mentioned clusters, it should be stressed that most of the identified protein (spheres) are not involved in protein-protein interactions and this is in agreement with the heterogeneous composition of colostrum, which is not associated with a well-defined cellular system. The functional analysis was then focused on the wound healing process, which was selected since there is some evidence reporting the colostrum capacity to induce, enhance and promote wound healing in different animal models. To do this, the bovine colostrum database was filtered by considering the main physiological steps involved in the wound healing process such as: haemostasis, inflammation, proliferation and tissue re-modelling.

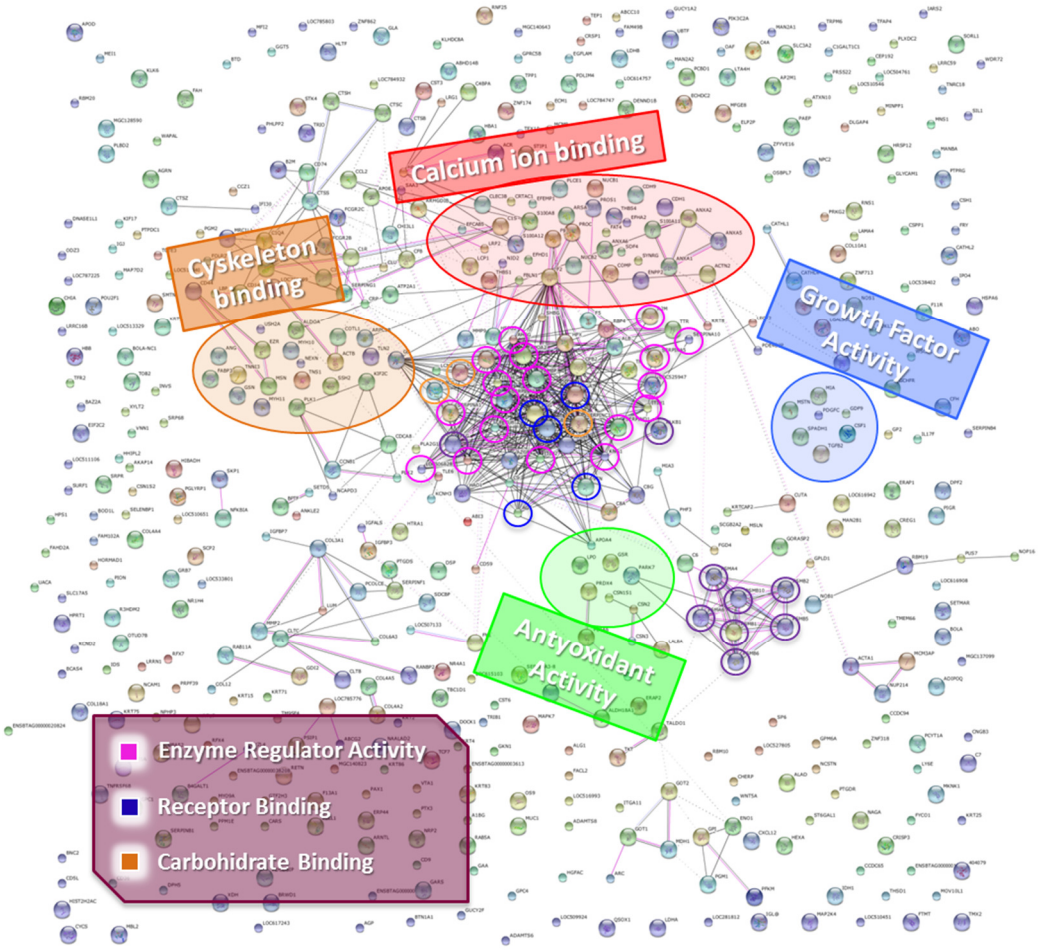


Figure 8- Protein interactome network for the bovine colostrum proteome using the STRING software. Thicker lines represent stronger associations.

A set of 93 proteins involved in the wound healing process were identified (Table 3), which can clustered on the basis of their biological functions as follow: 17 involved in the haemostasis process, 9 in the inflammatory process; 9, 15, and 13 in angiogenesis, extracellular matrix organization and regulation of epithelial cell proliferation, respectively; 5 proteins involved in collagen synthesis and 25 acting in epithelial remodelling.

HAEMOSTASIS	Vitamin K-dependent protein S Coagulation factor XIIIa Vitamin K-dependent protein C Coagulation factor V Plasma kallikrein Coagulation factor IX Prothrombin Kininogen-1 Carboxypeptidase B2 Fibrinogen alpha chain Fibrinogen beta chain Fibrinogen gamma-B chain Antithrombin-III Plasminogen Factor XIIa inhibitor Annexin A5 Hermansky-Pudlak syndrome 1 protein
INFLAMMATION	CD74 antigen Thrombospondin-4 C-C motif chemokine 2 Colony stimulating factor 1 Prothrombin Inter-alpha-trypsin inhibitor heavy chain H4 Alpha-2-antiplasmin Alpha-1-acid glycoprotein Fibronectin
ANGIOGENESIS	Angiogenin-1 Fibronectin Thrombospondin-1 Serine/threonine-protein kinase 4 Lactadherin (Milk fat globule-EGF factor 8) alpha 1 type XVIII collagen Nuclear receptor subfamily 4 group A member 1 Glucose-6-phosphate isomerase Beta-1,4-galactosyltransferase 1 Annexin A2 Ephrin-A1 72 kDa type IV collagenase (Matrix metalloproteinase-2)
ECM ORGANIZATION	Fibulin 1 92 kDa type IV collagenase (Matrix metalloproteinase-9) Vitronectin Angiotensinogen Plasminogen Alpha-2-antiplasmin Collagen alpha-4(IV) chain Alpha 1 type XVIII collagen

	Transforming growth factor beta-2
	Beta-1,4-galactosyltransferase 1
	Annexin A2
	Galectin 3
	Pikachurin
	Collagen alpha-1(III) chain
	Melanoma-derived growth regulatory protein
REGULATION OF EPITHELIAL CELL PROLIFERATION	Angiogenin-1
	Thrombospondin-1
	Thrombospondin-4
	Nuclear receptor subfamily 4 group A member 1
	Beta-1,4-galactosyltransferase 1
	NDP kinase A ;
	Pigment epithelium-derived factor
	Wingless-type MMTV integration site family, member 5A
	Annexin A1
	Desmoplakin
	Colony stimulating factor 1
	Cathepsin Z
	Cathepsin H
COLLAGEN SYNTHESIS	92 kDa type IV collagenase (Matrix metalloproteinase-9)
	72 kDa type IV collagenase (Matrix metalloproteinase-2)
	Collagen alpha-1(III) chain
	Prothrombin
	Fibrinogen gamma-B chain
REMODELLING	92 kDa type IV collagenase (Matrix metalloproteinase-9)
	72 kDa type IV collagenase (Matrix metalloproteinase-2)
	Actin, cytoplasmic 2 (Gamma-actin)
	Myosin-10
	Inversin
	Serine/threonine-protein kinase 4
	Retinol-binding protein 4
	Alpha-2-antiplasmin
	Fibronectin
	Beta-1,4-galactosyltransferase 1
	Collagen alpha-1(III) chain
	Transforming growth factor beta-2
	Neuropilin 2
	Alpha 1 type XVIII collagen
	Lysosomal alpha-glucosidase
	Glucose-6-phosphate isomerase
	Lactadherin (Milk fat globule-EGF factor 8)
	Troponin I, cardiac muscle
	Lymphocyte antigen 6 complex, locus E
	Annexin A2

	Ephrin-A1 LAMA4 protein Angiogenin-1 Thrombospondin-1 Plasminogen

Table 3 - BC proteins involved in wound healing clustered on the basis of the physiological step involved in the process.

The involvement of the identified proteins in each step of wound healing can be further explored on the basis of the detailed functional analysis of the proteins. As an example, Figure 9 shows the involvement of the identified proteins in the molecular events leading to haemostasis, which happens just before the inflammatory phase during wound healing. BC contains the main protein factors of the coagulation cascade and in particular the main coagulation factors (VIIIa, XII, V, IX), the Thrombin precursor and its corresponding inhibitor, and two serine-proteases (Protein C, Protein S), whose active form plays an important role in the regulation of blood coagulation.

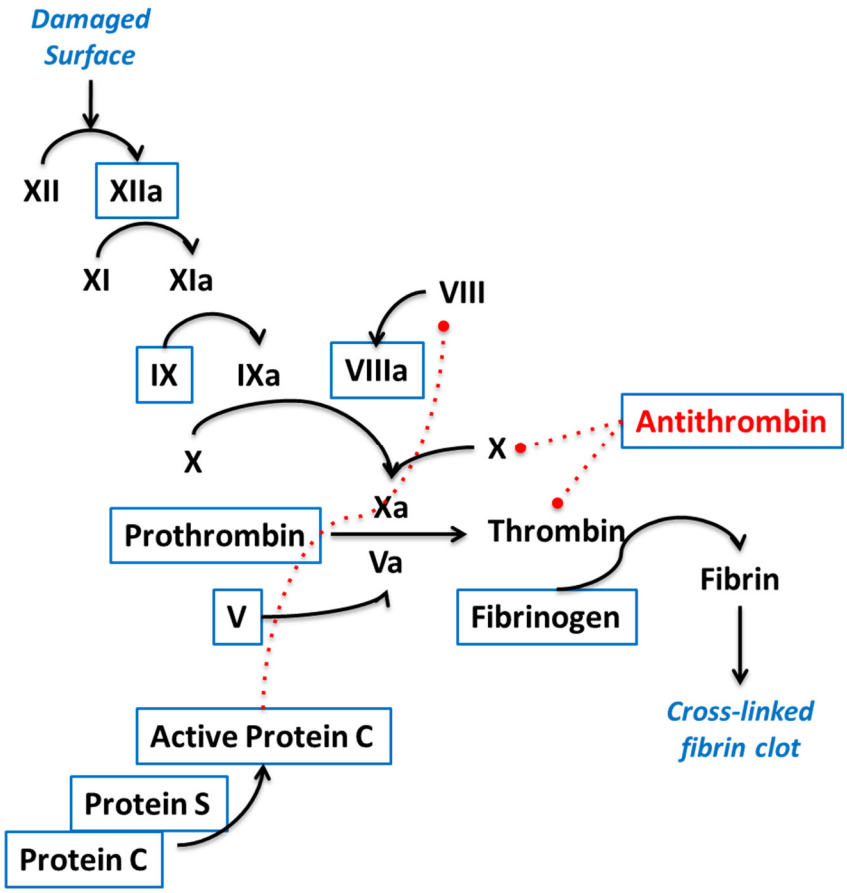


Figure 9 - Haemostasis pathway; the gene products highlighted by a blue colour have been detected in bovine colostrum.

The BC proteome was then used to search novel potential health applications of BC. To do this we created a list of proteins that are potentially useful for health and with an external use. Such proteins were sought with either recombinant or natural origin and using PubMed or patent search engines. The set of selected proteins was then sought in the database of BC. By using such an approach we found some proteins in

BC that are currently used or are potentially useful for treating some human diseases and mostly produced by recombinant technology. As an example we found Lumican, a 40 kDa keratan sulfate proteoglycan which is receiving growing interest since it is involved in corneal epithelial wound healing [21]. The content of Lumican in BC (as determined by ELISA) ranged between 10 and 30 pg mg^{-1}.

4. Discussion

4.1. Off-target based proteomics to elucidate the bases of its biological activities

The health benefits of bovine colostrum have been largely reported and partially demonstrated in some pharmacological models and clinical studies, but the molecular mechanisms are poorly understood [5, 22]. The health claims of BC are usually associated with the presence of a certain cluster of known proteins (i.e. growth factors for wound healing) whose well-defined biological function is taken into account to furnish the explanation for the BC activity. Another approach consists of searching for BC putative proteins whose presence would explain the BC activity. These approaches, based on selecting bioactive components *a priori* (on-target approach), are limited, since they do not permit the association of the biological function to unknown compounds, a very important issue to extend our knowledge of the molecular mechanisms of BC. To overcome such limit, we here propose an off-target method based on proteomic analysis that consists of identifying proteins involved in a certain biological effect and not selected *a priori*. Clearly, to make such an approach reliable, an exhaustive protein data bank of BC, including the minor protein components is required and this represented the first step of the work.

4.2. Protein identification via mass spectrometry analysis

Since bovine colostrum, in addition to being an extremely complex biological matrix, is characterized by the presence of abundant proteins (caseins and immunoglobulins), which limit the separation and identification of the minor protein components by conventional techniques, an innovative analytical approach was required.

Traditional systematic proteomic approaches are essentially based on two consecutive analytical steps: protein separation followed by identification. The conventional approach was initially applied for our purpose, but it turned out to be severely limiting due to its failure to separate/identify the minor protein components, including growth factors present in trace amounts, presumably responsible for some of the biological activities.

The need to solve a complex proteome, characterized by a wide dynamic range, has focused our attention on proteomic techniques able to reduce protein concentration differences and enrich the low-abundance species, among which the well-ingrained “CPLL technology” appears to offer the best solution, since it is the only one able to “normalize” a protein population, by sharply reducing the concentration of the most abundant components, while simultaneously enhancing the concentration of the most rare species [19].

An important goal was to set-up the best working conditions of the "CPLL technology" in order to extract the largest number of proteins. Since in the capture process ionic interactions are the most prominent ones, the extraction was conducted at four

different pH values (2.2, 4.0, 7.0, and 9.3) to modulate the surface charge of both the proteins and the hexa-peptide baits.

The one and two-dimensional electrophoretic analysis of extracts confirmed that the sample treatment with the CPLL technology greatly improved the resolution of the proteome. Moreover, both the one- and two-dimensional electrophoretic maps of the extracts corresponding to the four different pH are visibly different and characterized by an extremely heterogeneous protein composition. In this way the importance and advantage of the parallel incubations at different pH values was established.

Although this technology enabled a significant improvement in the resolution of the colostrum proteome, the immunoglobulin component still represented the main component. The only partial reduction of this protein component can be explained considering the heterogeneity of the variable portion of the immunoglobulin, so that the class of immunoglobulins is not recognized as a single protein, and therefore normalized, but as a set of more heterogeneous proteins.

Therefore, in order to achieve a more comprehensive “picture” of the bovine colostrum proteome, the next step was to analyze the IgG-depleted colostrum obtained by affinity chromatography. Immunoassays and electrophoretic analysis confirmed the efficiency and robustness of the depletion method.

As we expected, our previous results showed that the IgG-depletion on the one hand allowed the enrichment of a unique set of proteins, and on the other led to the loss of an equivalent number of proteins. This concept is in agreement with what reported in the literature: the selective removal of very abundant proteins, at the same time, cause the co-depletion of many species associated to the proteins removed [23].

Therefore, an in-depth analysis of the proteome cannot be achieved by considering only one sample preparation strategy, and/or a single extraction method, but through the integration of all the data obtained by using different approaches, since these are complementary.

By applying specific tools provided by Proteome Discoverer software, we compared the number of proteins identified in the following conditions: dialyzed colostrum, colostrum treated with CPLLs, depleted colostrum, colostrum depleted and further treated with CPLLs (Table 2).

An accurate analysis of the values in the Table 2 leads to the conclusion that every single treatment, in a different way, increases knowledge of the proteome, therefore, in order to achieve a complete and comprehensive database of proteins belonging to bovine colostrum, it was necessary to merge all the results obtained.

By using such an approach a list of 634 proteins was obtained by setting a high peptide confidence and 1786 proteins by using a medium peptide confidence. The number of the identified proteins represents to our knowledge the most extensive list of bovine colostrum proteins available so far reported.

4.3. Functional study

The final database shown in Table S1 was then used for the functional investigation. It was firstly applied for the off-target analysis aiming at better understanding the health benefit activity of colostrum at molecular level. We firstly clustered the proteins based on their biological function and in turn on the health claims. We started by focusing on the effect of BC to modulate and furnish a support effect on the immune system, since it is definitely one of the most studied and well known health effects

exerted by BC [5]; besides immunoglobulins, we found a cluster of 4.02% proteins involved in the defensive response and which can contribute to such effect. Another important biological effect already demonstrated for BC is its ability to promote epithelial cell proliferation, wound healing, accelerated proliferation of keratinocytes, and its positive effect on the re-epithelialization. A set of 2.43% of BC proteins is involved in the proliferation processes and 12.57% and 12.51% of the total protein content, are involved in cell division and cell growth, respectively, which together with the proteins having a proliferative activity, are supposed to contribute actively to the widely demonstrated wound healing and regeneration process of BC.

A more detailed function analysis was then carried out for the wound healing process. The off-target approach identified 93 proteins that were then clustered on the basis of the 7 biological events leading to wound healing. Most of the identified proteins in BC were not reported as acting in the wound healing process, which is usually related to the presence of growth factors. Besides an important insight into the molecular mechanism, such an approach can also be important in setting up suitable analytical procedures to guarantee the quality control. For instance if BC is used as a wound healing product then a panel of proteins involved in this biological process can be selected for the qualitative and quantitative analysis thus assuring the quality and safety of the final product.

Finally, by cross searching the BC database with a list of proteins, natural or recombinant, claimed to have health benefits for external use, novel applications of BC can be explored. As an example, we found that BC contains lumican (10-30 pg/mg) which is a keratan sulfate proteoglycan that may be beneficial for treating epithelial defects in the cornea [21, 24].

5. Conclusions

In conclusion this paper on the one hand highlights the importance of ProteoMiner in greatly expanding the visibility of low- to very-low abundance species and, on the other hand, underlines the importance of proteomics in the pharmaceutical field and in particular for clarifying the biological effects of a complex protein matrix.

References

1. Kehoe, S.I., B.M. Jayarao, and A.J. Heinrichs, *A survey of bovine colostrum composition and colostrum management practices on Pennsylvania dairy farms.* J Dairy Sci, 2007. **90**(9): p. 4108-16.
2. Sacerdote, P., et al., *Biological components in a standardized derivative of bovine colostrum.* J Dairy Sci, 2013. **96**(3): p. 1745-54.
3. Bechara, E.J.H., et al., *The dual face of endogenous alpha-aminoketones: Pro-oxidizing metabolic weapons.* Comparative Biochemistry and Physiology C-Toxicology & Pharmacology, 2007. **146**(1-2): p. 88-110.
4. Conneely, M., et al., *Effect of feeding colostrum at different volumes and subsequent number of transition milk feeds on the serum immunoglobulin G concentration and health status of dairy calves.* J Dairy Sci, 2014. **97**(11): p. 6991-7000.
5. Rathe, M., et al., *Clinical applications of bovine colostrum therapy: a systematic review.* Nutr Rev, 2014. **72**(4): p. 237-54.
6. Weaver, D.M., et al., *Passive transfer of colostral immunoglobulins in calves.* J Vet Intern Med, 2000. **14**(6): p. 569-77.
7. Kelly, G.S., *Bovine colostrums: a review of clinical uses.* Altern Med Rev, 2003. **8**(4): p. 378-94.
8. Bühler, C., et al., *Small intestinal morphology in eight-day-old calves fed colostrum for different durations or only milk replacer and treated with long-R3-insulin-like growth factor I and growth hormone.* J Anim Sci, 1998. **76**(3): p. 758-65.
9. Cairangzhuoma, et al., *A preparation of cow's late colostrum fraction containing αs1-casein promoted the proliferation of cultured rat intestinal IEC-6 epithelial cells.* Biosci Biotechnol Biochem, 2013. **77**(5): p. 992-7.
10. Cairangzhuoma, et al., *Skimmed, sterilized, and concentrated bovine late colostrum promotes both prevention and recovery from intestinal tissue damage in mice.* J Dairy Sci, 2013. **96**(3): p. 1347-55.
11. Støy, A.C., et al., *Bovine colostrum improves intestinal function following formula-induced gut inflammation in preterm pigs.* Clin Nutr, 2014. **33**(2): p. 322-9.
12. Doillon, C.J., et al., *Modulatory effect of a complex fraction derived from colostrum on fibroblast contractibility and consequences on repair tissue.* Int Wound J, 2011. **8**(3): p. 280-90.
13. Kovacs, D., et al., *Bovine colostrum promotes growth and migration of the human keratinocyte HaCaT cell line.* Growth Factors, 2009. **27**(6): p. 448-55.
14. Tucci, S., et al., *Colostrum in menopause effects on vaginal cytology/symptoms.* Clin Exp Obstet Gynecol, 2013. **40**(2): p. 219-21.
15. Galitsopoulou, A., et al., *Polyamine profile in ovine and caprine colostrum and milk.* Food Chem, 2015. **173**: p. 80-5.
16. Purup, S., et al., *Biological activity of bovine milk on proliferation of human intestinal cells.* J Dairy Res, 2007. **74**(1): p. 58-65.

17. Kramski, M., et al., *Hyperimmune bovine colostrum as a low-cost, large-scale source of antibodies with broad neutralizing activity for HIV-1 envelope with potential use in microbicides.* Antimicrob Agents Chemother, 2012. **56**(8): p. 4310-9.
18. Nissen, A., et al., *In-depth analysis of low abundant proteins in bovine colostrum using different fractionation techniques.* Proteomics, 2012. **12**(18): p. 2866-78.
19. Righetti, P.G., et al., *Protein Equalizer Technology : the quest for a "democratic proteome".* Proteomics, 2006. **6**(14): p. 3980-92.
20. Renzone, G., S. Arena, and A. Scaloni, *Proteomic characterization of intermediate and advanced glycation end-products in commercial milk samples.* J Proteomics, 2015. **117**: p. 12-23.
21. Liu, C.Y. and W.W. Kao, *Lumican promotes corneal epithelial wound healing.* Methods Mol Biol, 2012. **836**: p. 285-90.
22. Bagwe, S., et al., *Bovine colostrum: an emerging nutraceutical.* J Complement Integr Med, 2015.
23. Restuccia, U., et al., *pI-based fractionation of serum proteomes versus anion exchange after enhancement of low-abundance proteins by means of peptide libraries.* J Proteomics, 2009. **72**(6): p. 1061-70.
24. Amjadi, S., et al., *The role of lumican in ocular disease.* ISRN Ophthalmol, 2013. **2013**: p. 632302.

Table S1 – List of the 634 unique gene products identified in bovine colostrum.

Accession Number	Protein Name	Coverage	N° Peptides
F1MTE8	**1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase epsilon-1**	1.17	1
Q2HJD7	**3-hydroxyisobutyrate dehydrogenase, mitochondrial**	5.36	1
F1MKI5	**45 kDa calcium-binding protein (Fragment)**	25.59	4
F1N2B5	**4F2 cell-surface antigen heavy chain**	22.04	7
E1BMZ8	**5'-3' exoribonuclease 1**	1.80	1
Q0IIG5	**6-phosphofructokinase, muscle type**	4.49	1
F1MKH8	**72 kDa type IV collagenase**	4.69	2
Q0VCX2	**78 kDa glucose-regulated protein**	9.92	5
E1BK03	**A disintegrin and metalloproteinase with thrombospondin motifs 6**	1.25	1
F1MP11	**A disintegrin and metalloproteinase with thrombospondin motifs 8**	1.33	1
A4FUD5	**ABI3 protein**	2.79	1
Q95M17	**Acidic mammalian chitinase**	3.18	1
P60712	**Actin, cytoplasmic 1**	47.73	13
Q58CQ2	**Actin-related protein 2/3 complex subunit 1B**	3.76	1
G5E5R8	**Activity-regulated cytoskeleton-associated protein**	3.03	1

Q0VCZ8	**Acyl-CoA synthetase long-chain family member 1**	1.43	1
Q3Y5Z3	**Adiponectin**	5.00	1
G3MYZ3	**Afamin**	20.70	9
F1MSI2	**Agrin**	0.69	1
G5E5D8	**A-kinase anchor protein 14**	11.86	1
Q2KJH7	**Aldehyde dehydrogenase 18 family, member A1**	1.38	1
A7YY28	**Alpha/beta hydrolase domain-containing protein 14B**	4.76	1
Q3SZR3	**Alpha-1-acid glycoprotein**	52.97	13
P34955	**Alpha-1-antiproteinase**	46.15	23
Q2KJF1	**Alpha-1B-glycoprotein**	55.07	19
P28800	**Alpha-2-antiplasmin**	29.27	10
P12763	**Alpha-2-HS-glycoprotein**	28.97	12
Q7SIH1	**Alpha-2-macroglobulin**	59.34	67
F1MRD4	**Alpha-actinin-2 (Fragment)**	1.88	1
F1MJQ3	**Alpha-amylase 2**	24.07	10
F1MB08	**Alpha-enolase**	17.28	5
E1B725	**Alpha-galactosidase A**	5.69	2
P00711	**Alpha-lactalbumin**	83.10	15
F1N7T2	**Alpha-mannosidase 2x - 1**	12.31	11
E1BGJ4	**Alpha-mannosidase 2x -2**	11.74	11
Q1RMM9	**Alpha-N-acetylgalactosaminidase**	6.81	2

P02662	**Alpha-S1-casein**	37.85	8
P02663	**Alpha-S2-casein**	19.37	3
E1B958	**Alstrom syndrome protein 1**	0.49	1
E1BJN3	**Amine oxidase**	12.10	7
P10152	**Angiogenin-1**	57.43	13
P80929	**Angiogenin-2**	45.53	6
Q3SZH5	**Angiotensinogen**	54.95	15
E1BFD1	**Ankyrin repeat and LEM domain-containing protein 2**	2.90	1
F1N650	**Annexin**	57.51	20
F6QVC9	**Annexin**	10.90	3
P04272	**Annexin A2**	31.86	7
P81287	**Annexin A5**	10.90	3
P79134	**Annexin A6**	2.08	1
F1MSZ6	**Antithrombin-III**	59.14	24
Q3ZC13	**AP-2 complex subunit mu**	4.83	1
G3N2W6	**APOL3 protein**	8.04	1
P15497	**Apolipoprotein A-I**	61.89	20
P81644	**Apolipoprotein A-II**	16.00	1
F1N3Q7	**Apolipoprotein A-IV**	15.53	4
E1BNR0	**Apolipoprotein B (Fragment)**	0.74	2
Q32KY0	**Apolipoprotein D**	13.23	2

Q03247	**Apolipoprotein E**	37.34	9
E1BLF5	**Aryl hydrocarbon receptor nuclear translocator-like protein 1**	2.88	1
Q08DD1	**Arylsulfatase A**	11.83	3
P33097	**Aspartate aminotransferase, cytoplasmic**	15.98	4
P12344	**Aspartate aminotransferase, mitochondrial**	2.09	1
F1MH20	**Ataxin-10**	3.37	1
E1B7A9	**ATP-binding cassette sub-family C member 11**	1.99	1
Q4GZT4	**ATP-binding cassette sub-family G member 2**	1.53	1
G3N0Q8	**Azurocidin**	42.68	7
F1MER7	**Basement membrane-specific heparan sulfate proteoglycan core protein**	0.30	1
P08037	**Beta-1,4-galactosyltransferase 1**	46.52	13
P17690	**Beta-2-glycoprotein 1**	55.65	14
P01888	**Beta-2-microglobulin**	66.10	11
P02666	**Beta-casein**	8.48	1
Q0V8R6	**Beta-hexosaminidase subunit alpha**	3.21	1
P02754	**Beta-lactoglobulin**	86.52	27
A6QLB0	**Beta-mannosidase**	7.96	4
F1MTT1	**Bile acid receptor**	2.28	1
F1N2G1	**Biorientation of chromosomes in cell division protein 1-like 1**	1.61	2
A6QQ07	**Biotinidase**	9.52	3
P13753	**BOLA class I histocompatibility antigen, alpha chain BL3-7**	31.59	10

A7YWH4	**BoLA protein**	22.04	6
F1N4Y5	**BOLA-NC1 protein (Bos taurus)**	38.60	11
E1BNJ9	**Brain-specific serine protease 4**	43.53	7
Q0IIA9	**Breast carcinoma amplified sequence 4**	8.54	1
F1N6I8	**Bromodomain adjacent to zinc finger domain protein 2A**	0.79	1
F1MTQ0	**Bromodomain and WD repeat-containing protein 1**	0.99	1
P18892	**Butyrophilin subfamily 1 member A1**	23.00	8
Q3SX46	**C1GALT1-specific chaperone 1**	3.14	1
Q28065	**C4b-binding protein alpha chain**	2.46	1
A8YXZ2	**C8G protein**	37.97	6
Q6R8F2	**Cadherin-1**	8.73	6
E1BHE0	**Cadherin-9**	3.39	1
B5UBG1	**Calcitonin**	15.38	1
F1MHU1	**Calcium uptake protein 3, mitochondrial**	3.00	1
Q2KIG3	**Carboxypeptidase B2**	34.52	11
Q17QK3	**Carboxypeptidase Q**	2.33	1
A6QNS7	**Cardiomyopathy-associated protein 2**	2.24	1
A4FUC8	**CARS protein**	3.97	1
F1MFI4	**Cartilage acidic protein 1**	48.46	27
P35445	**Cartilage oligomeric matrix protein**	3.57	2
P22226	**Cathelicidin-1**	50.32	9

P19660	**Cathelicidin-2**	9.66	1
P33046	**Cathelicidin-4**	20.83	2
P07688	**Cathepsin B**	15.52	3
E1BEQ4	**Cathepsin G**	3.61	1
P25326	**Cathepsin S**	19.34	4
P05689	**Cathepsin Z**	32.89	8
P00760	**Cationic trypsin**	70.73	18
P28291	**C-C motif chemokine 2**	12.12	2
F1MHC3	**CD44 antigen**	9.29	3
Q32PA1	**CD59 molecule, complement regulatory protein**	20.66	2
F1N514	**CD5L protein**	36.64	18
Q29630	**CD74 molecule, major histocompatibility complex, class II invariant chain**	15.20	1
G3MYH4	**CD81 antigen (Fragment)**	16.82	2
G8JKX6	**CD9 antigen (Fragment)**	31.37	5
A4IFM5	**CDCA8 protein**	6.07	1
G3X7L5	**cDNA FLJ54643, highly similar to Homo sapiens SET domain and mariner transposase fusion gene (SETMAR), mRNA**	5.97	1
F1N3A0	**cDNA FLJ60976**	3.84	1
F1MAX3	**Centrosomal protein of 192 kDa**	0.84	1
F1N192	**Centrosome and spindle pole-associated protein 1**	2.26	1
F1MDQ6	**cGMP-dependent protein kinase 2**	2.30	1

A8E648	**CHERP protein**	2.73	1
G3X7D2	**Chitinase-3-like protein 1**	60.61	20
F1N2J7	**Chitobiosyldiphosphodolichol beta-mannosyltransferase**	6.68	1
P09611	**Chorionic somatomammotropin hormone 1**	11.44	1
F1MPU0	**Clathrin heavy chain (Fragment)**	1.20	1
P17697	**Clusterin**	32.80	13
Q2HJ57	**Coactosin-like protein**	21.83	4
F1MBC5	**Coagulation factor IX**	8.95	3
Q28107	**Coagulation factor V**	1.04	2
F1MW44	**Coagulation factor XIII A chain**	22.68	13
Q17QL3	**Coiled-coil domain containing 94**	9.06	1
Q2TA16	**Coiled-coil domain-containing protein 65**	2.21	1
F1N2Y4	**Collagen alpha-1(III) chain (Fragment)**	1.72	1
E1B7H2	**Collagen alpha-1(VII) chain**	1.74	3
F1MZ30	**Collagen alpha-1(X) chain**	4.45	1
G3MZI7	**Collagen alpha-1(XI) chain**	12.22	16
F1N401	**Collagen alpha-1(XII) chain**	0.35	1
F1N6W9	**Collagen alpha-1(XVIII) chain**	1.46	2
F1N7Q7	**Collagen alpha-2(IV) chain (Fragment)**	2.00	1
G3N3E4	**Collagen alpha-3(VI) chain**	2.47	2
F1N474	**Collagen alpha-4(IV) chain**	1.30	1

Q29442	**Collagen alpha-4(IV) chain (Fragment)**	7.95	1
P42916	**Collectin-43**	38.63	11
Q5E9E3	**Complement C1q subcomponent subunit A**	13.93	2
Q2KIV9	**Complement C1q subcomponent subunit B**	5.67	1
Q0VCX1	**Complement C1s subcomponent**	5.81	4
Q3SYW2	**Complement C2**	38.53	25
Q2UVX4	**Complement C3**	79.11	152
P01030	**Complement C4 (Fragments)**	34.13	22
E1BH06	**Complement C4-B**	37.56	47
F1MVK1	**Complement C4-B**	21.71	27
A5D9E9	**Complement component 1, r subcomponent**	6.67	3
Q29RU4	**Complement component C6**	47.10	31
F1N045	**Complement component C7**	3.32	2
F1MX87	**Complement component C8 alpha chain**	23.60	10
Q3MHN2	**Complement component C9**	60.40	29
P81187	**Complement factor B**	62.29	54
Q3T0A3	**Complement factor D**	43.63	9
F1MW79	**Complement factor H**	4.46	1
Q28085	**Complement factor H**	69.66	71
F1MC45	**Complement factor H (Fragment)**	65.80	48
Q17QC8	**Complement factor properdin**	2.80	1

F1MPS0	**Condensin-2 complex subunit D3**	1.21	1
C4T8B4	**C-reactive protein**	41.96	5
Q0P5I9	**CXCL12 protein**	35.96	4
F1MZC9	**Cyclic nucleotide-gated cation channel beta-3**	3.31	1
E1BM19	**Cyclin-dependent kinase-like 3**	3.38	1
F6QEL0	**Cystatin-B**	18.18	1
P01035	**Cystatin-C**	64.19	9
G3N3P6	**Cystatin-M**	32.86	1
F6R3I5	**Cysteine-rich secretory protein 1**	21.49	7
P62894	**Cytochrome c**	36.19	3
Q0V896	**Cytochrome P450, family 4, subfamily F, polypeptide 2**	3.63	1
A6QQS0	**D4, zinc and double PHD fingers family 2**	4.86	1
F1MRG2	**Death domain-containing protein 1**	1.79	1
F1MIX6	**Dedicator of cytokinesis protein 1**	1.07	1
Q58DK5	**Delta-aminolevulinic acid dehydratase**	3.04	1
F1N0M9	**DENN domain-containing protein 1B**	1.86	1
F1MW13	**Deoxyribonuclease (Fragment)**	12.06	2
E1BKT9	**Desmoplakin**	0.45	1
Q3ZCJ8	**Dipeptidyl peptidase 1**	7.13	4
Q5E982	**Diphthine synthase**	3.86	1
F1N5I2	**Disks large-associated protein 4**	2.12	1

F1N2W9	**DNA helicase MCM9**	1.23	1
Q3MHE4	**DNA mismatch repair protein Msh2**	1.82	1
E1BFW7	**DNA-binding protein RFX7**	3.22	2
F1MXS1	**E3 SUMO-protein ligase RanBP2**	0.56	1
Q5E9N3	**E3 ubiquitin-protein ligase RNF25**	5.02	1
F1MN61	**Early endosome antigen 1**	3.11	2
F1N0Z9	**Early lactation protein**	19.42	2
A5PJT7	**ECM1 protein**	5.42	2
A1A4K5	**Ectonucleotide pyrophosphatase/phosphodiesterase family member 2**	33.90	23
E1B7P6	**EF-hand calcium-binding domain-containing protein 5**	1.66	1
Q17QM6	**EF-hand domain-containing protein D1**	4.66	1
A2VE41	**EGF-containing fibulin-like extracellular matrix protein 1**	17.67	5
E1BED8	**Elongation factor 1-alpha (Fragment)**	3.16	1
B8Y9S9	**Embryo-specific fibronectin 1 transcript variant**	16.38	23
F1MU34	**Endoplasmic reticulum aminopeptidase 2**	14.33	11
A6QPT7	**Endoplasmic reticulum aminopeptidase 2**	1.36	1
Q2TBT3	**Enoyl-CoA hydratase domain-containing protein 2, mitochondrial**	11.82	1
Q3ZC64	**Ephrin-A1**	19.51	3
P79345	**Epididymal secretory protein E1**	47.65	5
F1MW08	**ERO1-like protein alpha**	5.10	1
B0JYP5	**Eukaryotic translation initiation factor 2C, 2**	1.98	1

P31976	**Ezrin**	6.54	4
E1BMJ0	**Factor XIIa inhibitor**	30.13	14
P50448	**Factor XIIa inhibitor**	30.13	14
P10790	**Fatty acid-binding protein, heart**	20.30	2
A8DC37	**Fc-gamma-RII-D**	17.99	2
F1MMZ2	**Ferritin**	5.79	1
Q58D62	**Fetuin-B**	21.45	5
Q3SZZ9	**FGG protein**	61.15	22
A5PJE3	**Fibrinogen alpha chain**	30.57	20
F1MAV0	**Fibrinogen beta chain**	53.94	26
Q9MZ06	**Fibroblast growth factor-binding protein 1**	30.77	7
F1MYN5	**Fibulin-1**	15.86	7
Q0VCN9	**Folate receptor 2 (Fetal)**	37.94	6
P02702	**Folate receptor alpha**	58.51	11
E1BJL8	**Folate receptor alpha**	38.13	8
Q6SJV6	**Foveolin**	8.11	1
A6QLL8	**Fructose-bisphosphate aldolase**	9.62	4
A5PKH3	**Fumarylacetoacetase**	5.97	2
F1MLX0	**Fumarylacetoacetate hydrolase domain-containing protein 2A**	9.24	1
G3N2N5	**FYVE and coiled-coil domain-containing protein 1**	1.25	1
E1B9P1	**FYVE, RhoGEF and PH domain-containing protein 4**	2.03	1

Q1JPD9	**G protein-coupled receptor, family C, group 5, member B**	6.72	3
Q1LZG6	**G2/mitotic-specific cyclin-B1**	2.34	1
A6QLZ0	**Galectin**	23.02	5
A7E3W2	**Galectin-3-binding protein**	7.39	3
G3N2D8	**Gamma-glutamyltranspeptidase 1**	17.08	9
A6QPN6	**Gamma-interferon-inducible lysosomal thiol reductase**	21.31	3
F1MBQ5	**Gamma-secretase-activating protein**	1.83	1
A5D7A2	**GARS protein**	3.11	1
Q3SX14	**Gelsolin**	49.11	28
G3X6N2	**General transcription factor IIH subunit 3 (Fragment)**	6.56	1
E1BJE7	**Germinal-center associated nuclear protein**	0.76	1
Q3ZBD7	**Glucose-6-phosphate isomerase**	16.16	6
F1MGS8	**Glutamate carboxypeptidase 2**	1.22	1
E1BKZ1	**Glutathione reductase, mitochondrial**	2.86	1
P80195	**Glycosylation-dependent cell adhesion molecule 1**	30.07	6
G3X745	**Glypican-1**	5.01	2
F1MCX1	**Glypican-4**	28.32	10
G5E6P1	**Gm628 protein**	6.13	1
F1MSB5	**Golgi reassembly-stacking protein 2**	5.71	1
F1MJV6	**G-protein coupled receptor 126**	3.53	3
Q8WMP9	**Group XV phospholipase A2**	3.44	1

Q1RMW5	**Growth factor receptor-bound protein 7**	4.51	1
O18836	**Growth/differentiation factor 8**	34.93	8
Q9GK68	**Growth/differentiation factor 9**	4.86	1
Q32L41	**GTP cyclohydrolase 1 feedback regulatory protein**	20.24	1
E1BP75	**Guanylate cyclase soluble subunit alpha-2**	1.92	1
E1BD63	**H-2 class I histocompatibility antigen, D-P alpha chain**	9.72	4
G3X6K8	**Haptoglobin**	21.45	6
F1MWU9	**Heat shock 70 kDa protein 6**	2.80	1
E1B812	**Heat shock factor protein 5**	3.82	1
G5E507	**Heat shock protein HSP 90-beta**	0.97	1
F1MLM2	**Helicase-like transcription factor**	1.39	1
P02081	**Hemoglobin fetal subunit beta**	33.79	6
P01966	**Hemoglobin subunit alpha**	15.49	1
P02070	**Hemoglobin subunit beta**	57.93	9
Q3SZV7	**Hemopexin**	73.64	33
Q2KII3	**Hepatitis A virus cellular receptor 1 N-terminal domain containing protein**	45.04	6
E1BCW0	**Hepatocyte growth factor activator**	10.77	2
E1B7M9	**Hermansky-Pudlak syndrome 1 protein**	2.42	1
E1BGX8	**HHIP-like protein 2**	1.38	1
E1B8L8	**High affinity immunoglobulin alpha and immunoglobulin mu Fc receptor**	4.45	2
F1MKS5	**Histidine-rich glycoprotein**	9.33	3

P33433	**Histidine-rich glycoprotein (Fragments)**	7.58	2
F1N6G4	**Histo-blood group ABO system transferase**	34.35	8
A1A4R1	**Histone H2A type 2-C**	21.71	2
E1BGW2	**Histone H2B**	11.90	1
E1BLK2	**Histone H2B.2**	9.77	1
G3N081	**Histone H4**	10.64	1
E1BKN0	**Histone-lysine N-methyltransferase 2B**	0.55	1
G3MYV9	**Histone-lysine N-methyltransferase SETMAR**	7.01	1
Q0VD24	**Histone-lysine N-methyltransferase SETMAR**	7.52	1
F1MXQ1	**Homeobox protein engrailed (Fragment)**	7.53	1
F1MCY3	**HORMA domain-containing protein 1**	4.65	1
G3MZ19	**HRPE773-like**	14.65	2
E1BC79	**Hydroxyacid oxidase 1**	5.66	1
G3N0T0	**Hypoxanthine-guanine phosphoribosyltransferase**	4.59	1
F1N2D5	**Iduronate 2-sulfatase**	15.02	4
G3MXB5	**Ig alpha-1 chain C region**	25.51	4
G3N342	**Ig epsilon chain C region**	19.72	5
G3N0V0	**Ig gamma-4 chain C region**	71.47	17
G3N3Q3	**Ig heavy chain V region MC101**	44.52	4
G3MZH0	**Ig heavy chain V region MC102**	25.00	3
G3MXG6	**Ig heavy chain V-II region ARH-77**	46.76	5

G3N1R1	**Ig heavy chain V-II region SESS**	24.09	4
F1MZ96	**Ig kappa chain C region**	48.75	15
G5E5H2	**Ig lambda chain V-IV region Bau**	36.69	3
F1MLW7	**Ig lambda-2 chain C regions**	64.96	16
F1MCF8	**Ig lambda-2 chain C regions**	49.15	15
F1MLW8	**Ig lambda-2 chain C regions**	42.06	11
G5E513	**Ig mu chain C region**	61.05	24
Q3T101	**IGL@ protein**	15.00	2
G5E604	**IGL@ protein**	63.55	5
G3X7I5	**IGL@ protein**	15.60	1
Q3SYR8	**Immunoglobulin J chain**	54.14	10
E1BLF1	**InaD-like protein**	0.72	1
F1MJZ4	**Insulin-like growth factor binding protein, acid labile subunit**	9.51	4
P20959	**Insulin-like growth factor-binding protein 3**	2.75	1
F1MPP2	**Insulin-like growth factor-binding protein 7**	52.13	11
F1N157	**Integrin alpha-11**	1.20	1
Q0VCM5	**Inter-alpha-trypsin inhibitor heavy chain H1**	51.10	40
F1MNW4	**Inter-alpha-trypsin inhibitor heavy chain H2**	34.25	32
F1MMD7	**Inter-alpha-trypsin inhibitor heavy chain H4**	40.61	27
F1MLN4	**Interleukin-17F**	12.94	1
A0JN59	**Inversin**	2.94	1

A4IFB8	**IPO4 protein**	1.39	1
Q9XSG3	**Isocitrate dehydrogenase [NADP] cytoplasmic**	38.41	14
G3N0D2	**Isoform 2 of Maestro heat-like repeat-containing protein family member 9**	3.52	1
G3MWG4	**Isoleucine--tRNA ligase, mitochondrial**	1.95	1
Q9XT56	**Junctional adhesion molecule A**	4.03	1
P02668	**Kappa-casein**	42.63	5
E1BHV7	**Kelch domain-containing protein 8A**	2.29	1
A6QNZ7	**Keratin 10 (Epidermolytic hyperkeratosis; keratosis palmaris et plantaris)**	11.79	7
A1A4M9	**Keratin associated protein 13-1**	7.32	1
A1L595	**Keratin, type I cytoskeletal 17**	6.58	3
Q0P5J4	**Keratin, type I cytoskeletal 25**	4.00	2
A4FUZ0	**Keratin, type II cuticular Hb3**	13.79	6
E1B898	**Keratin, type II cuticular Hb6**	13.99	6
G3MZ71	**Keratin, type II cytoskeletal 2**	7.93	4
M0QVZ6	**Keratin, type II cytoskeletal 5**	12.91	8
M0QVY0	**Keratin, type II cytoskeletal 6A**	14.54	10
Q148H5	**Keratin, type II cytoskeletal 71**	4.00	2
Q08D91	**Keratin, type II cytoskeletal 75**	12.34	7
P05786	**Keratin, type II cytoskeletal 8**	9.62	5
G3N171	**Kielin/chordin-like protein**	0.92	1
F6RB08	**KIF17 protein**	4.49	1

A6QPE8	**KIF2C protein**	1.66	1
P01044	**Kininogen-1**	35.27	17
P01045	**Kininogen-2**	26.01	13
A6QQ95	**KLK6 protein**	52.44	7
Q17QL7	**KRT15 protein**	9.49	4
A4IFP2	**KRT4 protein**	9.11	5
Q95114	**Lactadherin**	58.78	25
G3MXZ0	**Lactoperoxidase (Fragment)**	58.76	25
P24627	**Lactotransferrin**	80.93	97
F1MYX9	**Laminin subunit alpha-4**	2.47	3
Q95M12	**Legumain**	12.24	6
Q2KIF2	**Leucine-rich alpha-2-glycoprotein 1**	41.62	12
F1MV58	**Leucine-rich repeat and calponin homology domain-containing protein 4**	2.58	1
E1BMH2	**Leucine-rich repeat-containing protein 16B**	2.48	1
Q5E9X4	**Leucine-rich repeat-containing protein 59**	5.56	1
O62644	**Leukocyte cell-derived chemotaxin-2**	20.53	2
Q1JPB0	**Leukocyte elastase inhibitor**	30.77	10
Q3SZH7	**Leukotriene A-4 hydrolase**	9.17	4
G3MWK8	**LGP2 protein**	3.75	1
F1MNN7	**Lipopolysaccharide-binding protein**	18.50	6
E1BM42	**LisH domain and HEAT repeat-containing protein KIAA1468**	2.15	1

P19858	**L-lactate dehydrogenase A chain**	9.34	3
Q5E9B1	**L-lactate dehydrogenase B chain**	31.14	10
A6QPC7	**LOC510340 protein (Fragment)**	5.94	1
A7Z077	**LOC510651 protein**	1.38	1
A5PJJ4	**LOC787225 protein**	7.94	1
A5PJH7	**LOC788112 protein**	17.44	2
E1B9L0	**Low affinity immunoglobulin gamma Fc region receptor II**	22.73	2
F1N6H1	**Low-density lipoprotein receptor-related protein 2**	14.59	51
Q05443	**Lumican**	14.91	4
Q1RMM2	**Lymphocyte antigen 6 complex, locus E**	8.46	1
Q9MYM4	**Lysosomal alpha-glucosidase**	4.27	3
Q29451	**Lysosomal alpha-mannosidase**	1.00	1
Q6B411	**Lysozyme C, milk isozyme**	15.54	2
F1MEC3	**Lysozyme C, tracheal isozyme**	34.69	5
Q27996	**Lysozyme C, tracheal isozyme**	15.65	2
F1MGS9	**Macrophage colony-stimulating factor 1**	1.99	1
F1MWM8	**Macrophage mannose receptor 1**	0.82	1
E1BJP1	**Major urinary protein 3**	8.79	1
Q3T145	**Malate dehydrogenase, cytoplasmic**	13.77	3
G3MY87	**Maltase-glucoamylase**	2.84	3
Q08DW8	**Mammary tumor virus receptor 2 homolog**	14.06	1

G3N120	**Mannose-binding protein C**	7.00	1
O02659	**Mannose-binding protein C**	13.25	2
F1N1A7	**MAP kinase-interacting serine/threonine-protein kinase 1**	2.86	1
A5PJP8	**MAP2K4 protein**	3.88	1
E1BF55	**MAP7 domain-containing protein 2**	2.38	1
P52176	**Matrix metalloproteinase-9**	6.74	3
E1BFW3	**Meiosis inhibitor protein 1**	1.56	1
F1MH18	**Meiosis-specific nuclear structural protein 1**	2.43	1
Q0VC16	**Melanoma inhibitory activity protein 3**	1.05	1
Q28038	**Melanoma-derived growth regulatory protein**	19.23	2
E1BG25	**Melanotransferrin**	4.87	3
A5PKJ4	**Mitogen-activated protein kinase 7**	1.92	1
Q2HJ49	**Moesin**	5.55	3
Q95122	**Monocyte differentiation antigen CD14**	46.92	11
A6QP39	**MSLN protein**	2.98	1
Q8WML4	**Mucin-1**	3.45	3
E1BNL1	**Multidrug resistance-associated protein 7**	1.61	1
Q2TBT6	**Multiple inositol polyphosphate histidine phosphatase, 1**	7.08	2
Q27991	**Myosin-10**	0.61	1
F1MYM9	**Myosin-14**	1.06	1
E1BH94	**N-acetylmuramoyl-L-alanine amidase**	33.99	9

A6QM01	**NAGLU protein**	21.74	9
E1BP93	**Nephrocystin-3**	1.05	1
P31836	**Neural cell adhesion molecule 1**	3.28	2
F1MCK2	**Neuroblast differentiation-associated protein AHNAK**	0.44	1
Q0VD07	**Neuronal membrane glycoprotein M6-a**	5.76	1
E1BEL6	**Neuropilin-2**	17.71	10
E1B6Z6	**Neutrophil gelatinase-associated lipocalin**	60.50	14
E1BDF0	**Nexilin**	2.07	1
Q8WNW7	**NF kappa B inhibitor alpha (Fragment)**	8.28	1
F1MT12	**Nicastrin**	1.53	1
A7E306	**NID2 protein**	2.53	1
E1BKI3	**Nitric oxide synthase**	2.19	1
G3X6Z0	**NMDA receptor synaptonuclear signaling and neuronal migration factor**	2.00	1
F1MGC2	**Non-secretory ribonuclease**	23.42	3
F1MV74	**Non-specific lipid-transfer protein**	3.50	1
F1MM55	**Nuclear pore complex protein Nup133**	2.16	1
F1MH31	**Nuclear pore complex protein Nup214**	1.21	1
F1MNL6	**Nuclear pore membrane glycoprotein 210-like**	1.08	1
Q0V8F0	**Nuclear receptor subfamily 4 group A member 1**	2.68	1
Q0IIH5	**Nucleobindin 2**	13.49	7
Q0P569	**Nucleobindin-1**	52.53	30

Q5E996	**Nucleolar protein 16**	6.18	1
F1MCZ7	**Nucleolar transcription factor 1**	1.59	1
Q3T0Q4	**Nucleoside diphosphate kinase B**	44.74	5
F1N3U7	**Nucleosome-remodeling factor subunit BPTF**	0.49	1
Q32KV6	**Nucleotide exchange factor SIL1**	15.58	7
Q0IIA2	**Odorant-binding protein-like**	48.26	7
F1MQH5	**Olfactory receptor**	4.21	1
E1BKS7	**Olfactory receptor**	4.81	1
G3MZJ6	**Olfactory receptor 51S1**	11.18	1
E1B9S7	**OTU domain-containing protein 7B**	2.85	1
E1BLL9	**Out at first protein homolog**	6.03	1
E1BBD8	**Oxysterol-binding protein**	1.78	1
F1MH87	**Paired box protein Pax-1**	5.53	1
F1N726	**Pancreatic secretory granule membrane major glycoprotein GP2**	12.17	5
Q58CQ9	**Pantetheinase**	6.47	2
Q8MJG1	**PC4 and SFRS1-interacting protein**	2.45	1
Q2KIR5	**PCYT1A protein**	3.69	1
A6QNL5	**PDIA6 protein (Fragment)**	3.09	1
Q3T005	**PDZ and LIM domain protein 4**	6.34	1
F1MXC4	**Pecanex-like protein 1**	0.56	1
G3X6Q8	**Pentraxin-related protein PTX3**	7.59	1

Q8SPP7	**Peptidoglycan recognition protein 1**	73.68	12
Q9BGI2	**Peroxiredoxin-4**	54.74	10
A6QQ11	**PGM2 protein (Fragment)**	2.87	1
E1BMD4	**PH domain leucine-rich repeat-containing protein phosphatase 2**	0.83	1
G5E5M1	**PHD finger protein 3**	0.84	1
P13696	**Phosphatidylethanolamine-binding protein 1**	67.91	10
E1BE68	**Phosphatidylinositol 4-phosphate 3-kinase C2 domain-containing subunit alpha**	2.30	1
F1MNM2	**Phosphatidylinositol-glycan-specific phospholipase D**	4.29	3
Q08DP0	**Phosphoglucomutase-1**	17.08	5
Q95121	**Pigment epithelium-derived factor**	63.22	24
A3KN33	**Pikachurin**	1.47	1
Q2KJ63	**Plasma kallikrein**	4.09	1
Q9N2I2	**Plasma serine protease inhibitor**	21.53	5
E1B726	**Plasminogen**	33.99	21
P06868	**Plasminogen**	34.24	21
F1MYX5	**Plastin-2**	31.42	12
P26201	**Platelet glycoprotein 4**	16.95	8
E1BJY4	**Platelet-derived growth factor C**	11.73	3
A0JN47	**Plexin domain containing 2**	5.46	2
P81265	**Polymeric immunoglobulin receptor**	58.26	43

P0CG53	**Polyubiquitin-B**	39.34	3
F1MVI5	**Potassium voltage-gated channel subfamily D member 2**	1.90	1
E1B760	**Potassium voltage-gated channel subfamily H member 3**	1.40	1
F1MMF9	**POU domain, class 2, transcription factor 1**	1.85	1
Q29437	**Primary amine oxidase, liver isozyme**	51.31	28
Q3T0I2	**Pro-cathepsin H**	23.58	4
Q2HJB6	**Procollagen C-endopeptidase enhancer**	12.42	3
E1BHJ0	**Profilin**	31.43	3
E1BKM4	**Programmed cell death 6-interacting protein**	7.77	6
A5D7K8	**Prostaglandin D2 receptor**	4.16	1
O02853	**Prostaglandin-H2 D-isomerase**	28.80	8
Q2TBX6	**Proteasome subunit beta type-1**	7.47	1
F1MMK9	**Protein AMBP**	45.45	23
F1MX50	**Protein CREG1**	67.42	10
F1MTI7	**Protein CutA**	9.09	1
Q5E946	**Protein DJ-1**	7.94	1
F1MIT7	**Protein FAM102A**	6.27	1
Q2KJI3	**Protein FAM49B**	8.02	2
F1N4N5	**Protein FAM81A**	4.38	1
F1MYQ7	**Protein furry homolog**	0.40	1
Q2KIT0	**Protein HP-20 homolog**	52.88	11

Q2KIX7	**Protein HP-25 homolog 1**	12.74	2
Q2KIU3	**Protein HP-25 homolog 2**	20.93	3
F1MX66	**Protein NYNRIN**	0.88	1
F1MX65	**Protein OS-9**	1.65	1
F1N580	**Protein phosphatase 1E**	2.33	1
E1BPN5	**Protein phosphatase Slingshot homolog 2**	2.18	1
F1MNT2	**Protein RTF2 homolog**	10.03	1
F1MX83	**Protein S100-A11**	10.78	1
P79105	**Protein S100-A12**	10.87	1
P28782	**Protein S100-A8**	28.09	2
F1N1D4	**Protein tweety homolog**	2.67	1
A7E379	**Protein tyrosine phosphatase domain-containing protein 1**	2.01	1
E1BDG5	**Protein Wnt**	6.05	1
P00735	**Prothrombin**	60.32	48
E1B949	**Protocadherin Fat 4**	0.42	1
A8E4M9	**PRPF39 protein**	3.47	1
F1ML88	**Pseudouridylate synthase 7 homolog**	2.28	1
Q3ZBD3	**Pterin-4-alpha-carbinolamine dehydratase**	23.08	2
G3X7G5	**Putative helicase Mov10l1**	1.89	1
Q2KIY5	**Putative phospholipase B-like 2**	5.09	2
A0JNC2	**R3H domain-containing protein 2**	1.62	1

P50397	**Rab GDP dissociation inhibitor beta**	11.69	3
E1BKQ2	**Rab11 family-interacting protein 1**	1.23	1
E1BDG6	**Rab11 family-interacting protein 4**	2.43	1
F2Z4D5	**Ras-related protein Rab-11A**	6.02	1
Q0IIG7	**Ras-related protein Rab-5A**	5.12	1
F1MFJ3	**Receptor-type tyrosine-protein phosphatase gamma**	2.99	3
P82943	**Regakine-1**	16.30	1
Q762I5	**Resistin**	12.84	1
O02740	**Retinal guanylyl cyclase 2**	1.99	1
P18902	**Retinol-binding protein 4**	63.93	10
Q9TU03	**Rho GDP-dissociation inhibitor 2**	12.50	1
P61823	**Ribonuclease pancreatic**	42.00	5
Q3T114	**Ribonuclease UK114**	32.85	3
Q58DP6	**Ribonuclease, RNase A family, 4**	64.63	10
F1N0L5	**RNA binding motif protein 19**	1.96	1
F1MFZ9	**RNA-binding protein 10**	1.51	1
F1MF93	**RNA-binding protein 20**	1.74	1
Q3T042	**RNA-binding protein NOB1**	7.26	1
A5PKC6	**RNF12 protein**	3.62	1
F1MGE7	**Sarcoplasmic/endoplasmic reticulum calcium ATPase 1**	1.61	1
A6QPK0	**SCGB2A2 protein**	28.26	2

A0JNP2	**Secretoglobin family 1D member**	8.82	1
Q2KJ32	**Selenium-binding protein 1**	41.31	14
F1N152	**Serine protease HTRA1**	28.54	12
E1BDW7	**Serine/threonine-protein kinase 4**	18.82	8
F1MH45	**Serine/threonine-protein kinase MRCK alpha**	2.21	2
Q2TA25	**Serine/threonine-protein kinase PLK1**	3.16	1
E1BP17	**Serine/threonine-protein kinase PLK2**	2.48	1
F1N3E3	**Serine-rich coiled-coil domain-containing protein 1**	2.56	1
Q29443	**Serotransferrin**	92.05	89
Q2HJF0	**Serotransferrin-like (Bos taurus)**	13.34	7
Q9TTE1	**Serpin A3-1**	66.91	23
G8JKW7	**Serpin A3-6**	56.31	21
A2I7N2	**Serpin A3-6**	45.65	16
A2I7N3	**Serpin A3-7**	28.30	14
A6QPQ2	**Serpin A3-8**	36.12	14
Q3SYR0	**Serpin peptidase inhibitor, clade A (Alpha-1 antiproteinase, antitrypsin), member 7**	18.49	5
A5PJ69	**SERPINA10 protein**	12.61	4
A6QPZ4	**SERPINB4 protein**	79.65	20
A6QPP2	**SERPIND1 protein**	28.43	9
P02769	**Serum albumin**	92.09	99

Q8SQ28	**Serum amyloid A protein**	32.06	4
Q3T004	**Serum amyloid P-component**	42.86	8
G3N2I7	**SET domain-containing protein 5**	1.19	1
A5PKC2	**SHBG protein**	8.48	2
E1B7R3	**Sialin**	4.04	1
F1MNJ2	**Signal recognition particle receptor subunit alpha**	4.18	1
A6QQW3	**Signal recognition particle subunit SRP68**	3.37	1
G3MYW5	**Single-pass membrane and coiled-coil domain-containing protein 2**	5.78	1
A6QP52	**SMTN protein**	2.53	1
G3MYM1	**Sodium channel protein type 10 subunit alpha**	1.22	1
E1BPZ1	**Sortilin-related receptor**	8.19	13
E1BD26	**Sorting nexin-14**	1.27	1
F1MUD3	**Sperm flagellar protein 2**	0.82	1
P29392	**Spermadhesin-1**	16.42	1
A5D796	**Spermatogenesis-defective protein 39 homolog**	3.53	1
Q3ZCF3	**S-phase kinase-associated protein 1**	9.20	1
Q2KIW7	**ST6GAL1 protein**	4.33	1
Q08E24	**Store-operated calcium entry-associated regulatory factor**	5.07	1
Q3ZBZ8	**Stress-induced-phosphoprotein 1**	3.50	1
F1MM32	**Sulfhydryl oxidase**	58.20	26
F1N6P2	**Surfeit locus protein 1**	2.32	1

Q3SZA6	**Syndecan binding protein (Syntenin)**	39.93	8
F1MI41	**Synergin gamma**	1.45	1
F1MQI1	**Talin-2**	0.43	1
O97790	**TBC1 domain family member 1**	1.29	1
A6QPP5	**TCF7 protein**	4.35	1
F1MTN5	**Tctex1 domain-containing protein 3**	7.73	1
E1BB53	**Telomerase protein component 1**	1.26	1
E1BMG7	**Teneurin-4**	0.82	1
Q9GLM4	**Tensin-1**	1.40	1
F1MND3	**Testis-expressed sequence 10 protein**	1.61	1
Q2KIS7	**Tetranectin**	50.00	11
A6QLV4	**TFAP4 protein**	7.46	1
G1K152	**Thioredoxin-related transmembrane protein 2**	6.67	1
F1N3A1	**Thrombospondin-1**	36.84	39
Q3SWW8	**Thrombospondin-4**	3.95	3
A5D7K2	**THSD1 protein**	4.12	1
A6QPW7	**TNFRSF6B protein**	37.16	9
A4FUY5	**TOB2 protein**	4.06	1
Q2TBL6	**Transaldolase**	13.35	3
F1MFD3	**Transcription factor RFX4**	2.81	1
E1B8P2	**Transcription factor Sp6**	7.20	1

F1MF62	**Transcription initiation factor TFIID subunit 1**	1.06	1
F1MIL2	**Transducin-like enhancer protein 6**	5.05	1
F1MF47	**Transferrin receptor protein 2**	2.74	1
P21214	**Transforming growth factor beta-2**	4.59	2
F1MBS3	**Transforming growth factor-beta-induced protein ig-h3 (Fragment)**	13.01	4
F1MWK1	**Transient receptor potential cation channel subfamily M member 6**	0.94	1
A7E3W4	**Transketolase**	25.17	9
A5D7E2	**Transmembrane 9 superfamily member 4**	2.65	1
O46375	**Transthyretin**	71.43	10
A6QLF4	**TRIB1 protein**	5.65	1
E1BLK1	**Trinucleotide repeat-containing gene 18 protein**	0.56	1
A6QQZ8	**TRIO protein**	1.92	1
G5E5Y3	**Tripartite motif-containing protein 5**	4.26	1
Q0V8B6	**Tripeptidyl-peptidase 1**	22.91	6
P08057	**Troponin I, cardiac muscle**	7.55	1
F1MMS7	**Uncharacterized protein**	13.11	6
F1MI18	**Uncharacterized protein**	5.62	6
E1BJV2	**Uncharacterized protein**	1.65	1
Q5E9S8	**Uncharacterized protein C15orf41 homolog**	7.83	1
Q2NL11	**Uncharacterized protein C8orf74 homolog**	8.19	1
E1BJV3	**Uncharacterized protein KIAA0895-like**	6.16	1

E1BKW7	**Unconventional myosin-IXa**	0.99	1
F1MLR9	**Unconventional myosin-VIIb**	0.65	1
F1MIA0	**UPF0544 protein C5orf45 homolog**	5.15	1
P62248	**UPF0556 protein C19orf10 homolog**	28.16	6
F1N3A8	**UPF0602 protein C4orf47**	4.21	1
E1BP05	**Usherin**	0.31	1
F1MKQ9	**Uveal autoantigen with coiled-coil domains and ankyrin repeats protein**	1.50	1
F1MX48	**UV-stimulated scaffold protein A**	4.82	1
Q0VD30	**Vacuolar fusion protein CCZ1 homolog**	4.58	1
F1N318	**Vacuolar protein sorting-associated protein VTA1 homolog**	3.58	1
Q3MHN5	**Vitamin D-binding protein**	74.05	36
P00745	**Vitamin K-dependent protein C (Fragment)**	9.43	3
P07224	**Vitamin K-dependent protein S**	27.41	19
Q3ZBS7	**Vitronectin**	11.13	4
E1BC63	**WD repeat-containing protein 72**	1.33	1
E1BGC3	**Wings apart-like protein homolog**	1.42	1
F1MJU6	**WSC domain-containing protein 1**	3.65	1
F6Q234	**Xaa-Pro dipeptidase**	7.71	3
F1MUT3	**Xanthine dehydrogenase/oxidase**	58.71	65
P80457	**Xanthine dehydrogenase/oxidase**	58.71	65
F1N283	**Xyloside xylosyltransferase 1**	4.83	1

Q5QQ49	**Xylosyltransferase 2**	2.54	1
E1BEG4	**Zinc finger FYVE domain-containing protein 16**	1.62	1
F1MB55	**Zinc finger protein 157**	3.40	1
Q3SZL2	**Zinc finger protein 174**	7.45	1
G3X6V5	**Zinc finger protein 208**	1.64	1
F1MJN7	**Zinc finger protein 318**	0.39	1
A8SM77	**Zinc finger protein 713**	4.26	1
F1N4Q9	**Zinc finger protein 839**	3.86	1
F1MFU7	**Zinc finger protein 862**	1.71	1
F1N447	**Zinc finger protein basonuclin-2**	2.15	1
E1BBZ5	**Zinc finger protein ZFPM2**	1.79	1
F1MWD7	**Zinc finger protein ZXDC isoform 1**	2.14	1
Q3ZCH5	**Zinc-alpha-2-glycoprotein**	64.88	32

Part I-B

Healing...is not a science but the intuitive art of wooing nature.
W.H. Auden, "The Art of Healing"

WOUND HEALING IN BOVINE COLOSTRUM

The initial idea was to highlight proteins group based on targeted biological process of pharmaceutical interest. Our functional analysis focused on **wound healing** process, since there are several evidences of colostrum capability to induce, enhance and promote the wound healing in different animal models. Wound healing process could be explained through four subsequently physiological step: *hemostasis, inflammation, proliferation and tissue re-modeling* (Figure 1).

For each biological process concerning in the healing phase we defined the proteins involved in; each protein "positioning" (Table 3 – Part I-A) has been combined with a mechanistic explanation of its involvement through the manual annotation of known or predicted functions, literature research, and sequence similarity confirmation with other organisms (in particular with the Human).

Figure 1 – Biological processes concerning in the healing phase.

During wound healing, just before the inflammatory phase, there is the cascade of coagulation regulation to obtain the arrest of blood loss, the so-called hemostasis. It represents the local response to hemorrhage, caused by the blood vessels disruption, based on thrombocytes action and the local activation of coagulation factors in tissue. This phase is characterized by the formation of a clot, typical structure consisting of a fibrin network in which remain trapped the corpuscular elements of the blood, which occupies the wound. Figure 9 (part I) shows the pathway of blood coagulation, as described in the literature; all the proteins identified, with a good degree of reliability, are highlighted in red.

Reasonably, the demonstrated presence of the main protagonists of the coagulation cascade could be a good starting point to explain the mechanism of action of all the processes concerning the wound healing.

Once the blood loss has stopped, the body immediately sends fluids containing plasma proteins, blood cells (primarily neutrophils, lymphocytes, and macrophages) and antibodies to the wound site to ensure that the proper equipment is in place to promote the *inflammation phase*.

Although the main “actors” of this phase are all the inflammatory cells, our proteomic work, as it has been set, does not allow to define the cell component, hence, we investigated on those pro-inflammatory factors (proteins) that can promote the chemotaxis, proliferation and activity enhancement of all the inflammatory cells involved (Figure 2).

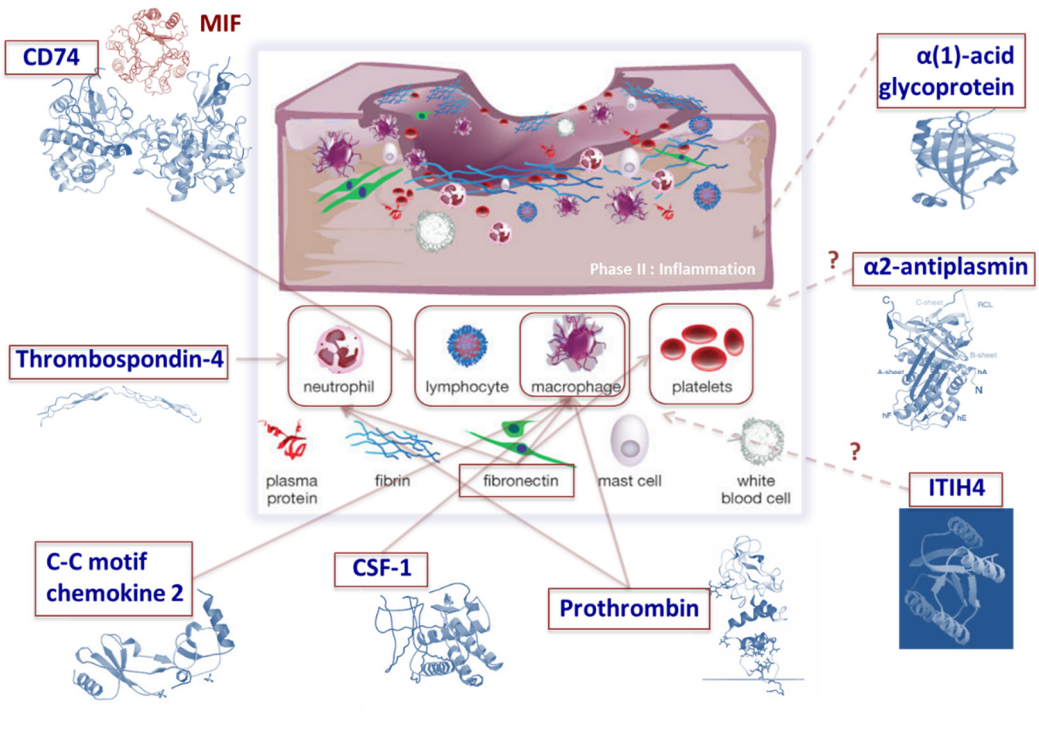


Figure 2 – Panel of pro-inflammatory factors (proteins) that can promote the chemotaxis, proliferation and activity enhancement of all the inflammatory cells involved.

The major pro-inflammatory protein is **Fibronectin**, since it promotes adhesion and migration of neutrophils, monocytes, into the wound region, enhances the phagocytosis of immune-opsonized particles by monocytes, and promote bacterial attachment and infection. [1]

Beside fibronectin, other pro-inflammatory factors showing a positive regulation of *neutrophil monocytes/macrophages*, B cells chemotaxis or exerts pro-proliferative effects on the same inflammatory cell lines are **CD74 antigen, Thrombospondin-4 (THBS4)**, **C-C motif chemokine 2, colony stimulating factor 1).** [2] [3]

Thrombin, whose role has been widely proved to go far beyond hemostasis, appears to play a key role in the early inflammatory phase, stimulating chemotaxis and production of inflammatory cytokines and chemokines.

A different class of inflammatory proteins are the *acute-phase proteins* (APP): proteins whose plasma concentration changes by at least 25% following inflammatory stimulus since they are able to initiate, sustain or resolve the inflammatory response. **Inter-alpha-trypsin inhibitor heavy chain H4** [4] and **alpha-1-acid glycoprotein** [5] appear to function in modulating the activity of the immune system during the acute-phase reaction (APR).

The fibrinolytic system is closely linked to the system of inflammation control. **Plasmin** in addition to lysing fibrin clots, also cleaves the complement system component C3. The system of **plasminogen** activation is controlled at several levels: plasminogen activator inhibitors (PAI-1 – stabilized by **vitronectin,** and PAI-2) counteract the activity of plasminogen activators and **alpha 2-antiplasmin** inhibits directly the activity of plasmin. [6]

Fibronectins is also involved in the *angiogenesis* pathway, by promoting the spreading of platelets at the site of injury, the adhesion and migration of fibroblasts, and endothelial cells into the wound region. [1]

Angiogenesis, *alias* neovascularization, is part of the longer *proliferation phase (Phase II in wound healing),* also characterized by *collagen deposition, granulation tissue formation, epithelialization, and wound contraction*, which overall restore skin integrity by filling in the wound with new tissue. During angiogenesis, new blood vessels form within the wound by extending from the wound's edges. Most notably, the process of *angiogenesis* occurs in overlapping phases in response to inflammation when endothelial cells migrate to the area of the wound.

As the wound *macrophages* switches from inflammatory to healing mode, it begins to secrete **endothelial chemotactic and growth factors** (**Angiogenin-1** and **Fibronectin).**

To migrate, endothelial cells need proteases able to break down existing vascular basal lamina: plasminogen activator and **collagenases** (**92 kDa type IV collagenase, also known as matrix metalloproteinase-9 (MMP-9), 72 kDa type IV collagenase - matrix metalloproteinase-2 (MMP-2))** are essential to degrade the clot and part of the ECM (extra-cellular matrix). [7]

On the other hand, it has been demonstrated that MMP2 also interacts with **Thrombospondin 1 (TSP1)** that is a known ligand for CD36 mediating anti-angiogenic properties. [8] *Thrombospondins* (TSPs), adhesive glycoprotein family, during wound healing act as angiogenic switches and control the extent of revascularization (Figure 3).

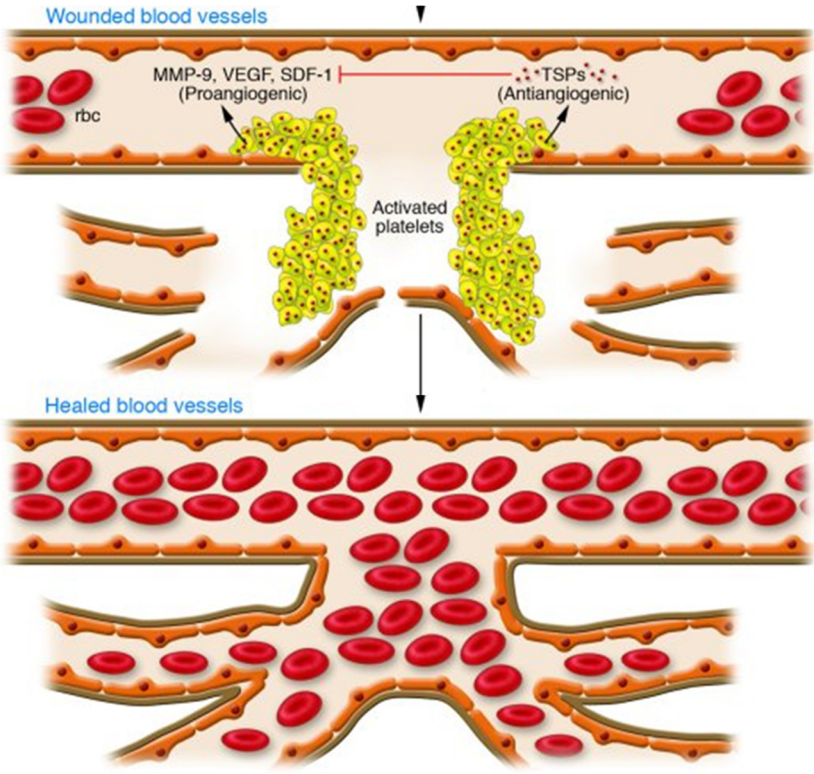


Figure 3 – Thrompospondins during wound healing act as angiogenic switches and regulate the revascularization process. [9]

Detached endothelial cells from pre-existing capillaries and post-capillary venues can divide and migrate chemotactically towards the wound, laying down new vessels in the process.

Neuropilin 2 (NRP2), commonly over-expressed in regions of physiological wound-healing, is a receptor for the vascular endothelial growth factor (VEGF) ligand family involved in angiogenesis by promoting endothelial cell survival.[10] [11]

Lactadherin (Milk fat globule-EGF factor 8) seems to play a crucial role in vascular endothelial growth factor (VEGF) – dependent neovascularization, contributing in this phase to phagocytic removal of apoptotic cells.[12]

Angiogenesis is tightly regulated by diverse growth factors with angiogenic activity: beside the *vascular endothelial growth factor (VEGF)* there are the *fibroblast growth factor (FGF)*, the **nerve Growth factor IB** (NGFIB), which in turn is an early response gene stimulated by FGF-2. [13] On the contrary **pigment epithelium-derived factor (PEDF)** is a multifunctional secreted protein with anti-angiogenic function interfering with VEGF signaling. Thrombospondins, antiangiogenic proteins, are upregulated by PEDF. [14]

The endothelium of vessels mature by laying down new endothelial extracellular matrix, followed by basal lamina formation. When macrophages and other growth factor-producing cells are no longer in a hypoxic they stop producing angiogenic factors. Thus, when tissue is adequately perfused, migration and proliferation of endothelial cells is reduced.

Simultaneously with angiogenesis, fibroblasts begin accumulating in the wound site becoming the main cells in the wound *(fibroplasia)*. [15] Previous studies have shown that the constituent chains of **fibrinogen** and its **thrombin-derived cleavage products** have the potential to promote fibroblast proliferation. At this stage phagocytes release proteases (**plasmin**) that break down the ECM of neighboring tissue, permitting activated fibroblasts from the adjacent uninjured cutaneous tissue to proliferate and migrate towards the wound.

Initially fibroblasts utilize the fibrin cross-linking fibers (well-formed by the end of the inflammatory phase) to migrate across the wound, subsequently adhering to **fibronectin**. Recently, it has been suggested that Angiotensin II (by-product from **Angiotensinogen)** and **Inversin** have a function in skin wound healing, in coordinating fibroblasts migration. [16] [17]

Fibroblasts then deposit ground substance (*granulation tissue*) into the wound bed, which consists of a provisional extracellular matrix (ECM). The ECM is a complex, highly organized three-dimensional structure that serves as a scaffold for tissue re-growth, but also as a signaling, binding, and activating substrate that interacts with the wound healing process. Components of the ECM play key roles in stimulating cell proliferation and differentiation, guiding cell migration, and modulating cellular responses.

Such components include ***structural proteins,*** primarily collagen, the most abundant proteins in the body. **Collagen** in the skin is primarily type I and **type III** (**Collagen alpha-1(III) chain)** and provide structure, strength, and integrity. In early granulation tissue **collagen III** is expressed at high levels, gradually replaced by collagen I in later phases of wound healing. [18] [38]

Type IV (**Collagen alpha-4(IV))** collagen is a component of epidermal basement membranes, where it forms complex protein networks which make up thin sheet-like structures, that separate and support cells in many tissues. While, an excess of **endostatin (alpha 1 type XVIII collagen),** and its precursor, in wound healing, leads to an increased fragility of the epidermal-dermal junction.[19]

Cell-adhesive glycoprotein. bind to cells and multiple components of the ECM and serve as modulators for growth factor activity (**fibronectin**, **laminin (laminin subunit alpha, LAMA4), fibulin (Fibulin -1) and vitronectin).**

Glycosaminoglycans draw large amounts of water into their structures, allowing them to resist compression forces, and interact with proteins (such as the adhesive protein **fibronectin**) in the ECM. **Pikachurin (EGF-like, fibronectin type-III and laminin G-like domain-containing protein**), for instance, binds to α-dystroglycan in the extracellular space and promotes matrix assembly and cell adhesiveness.

Proteoglycans serve as co-receptors for growth factors, participate in cell signaling, and help regulate the activity of many other molecules.

Matricellular protein class include **thrombospondins (THBS1 / THBS4)**, osteopontin, tenascins, and secreted protein acidic and rich in cysteine (**SPARC**), which modulate cell-matrix interactions and help regulate the response to certain growth factors.

Growth factors (PDGF, **TGF-β**) together with **fibronectin** encourage proliferation, migration to the wound bed, and production of ECM molecules by fibroblasts.

It has been largely demonstrated the presence of the isoform **TGF-beta 2** in bovine colostrum; exogenous application of the growth factor to enhance healing has been already investigated: TGF-beta2 appears to be of value in shortening the time for gain of incisional wound strength in healthy rats. [20]

Among all the duties of fibroblasts, the most important is the *production of collagen (Figure 5)*, because collagen itself is essential to increase the wound strength. .

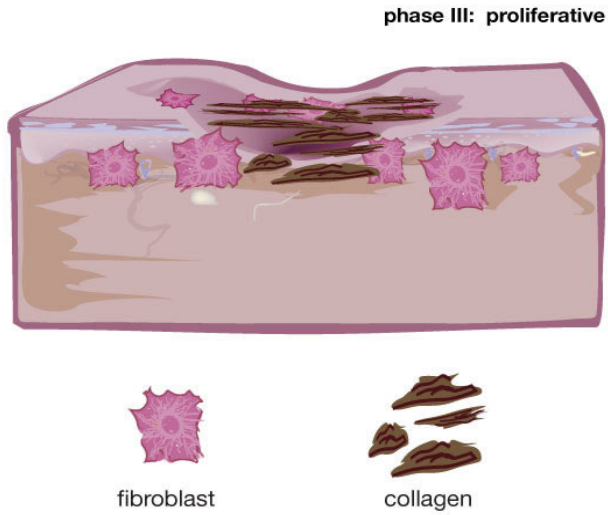


Figure 5 – Proliferative phase picture.

Even as fibroblasts are producing new collagen, **collagenases** (**Matrix metalloproteinase-9** / **Matrix metalloproteinase-2)** other factors degrade it. Shortly after wounding, synthesis exceeds degradation so collagen levels in the wound rise, but later production and degradation become equal so there is no net collagen gain signaling the onset of the later maturation phase.[21] At the end of the granulation phase, fibroblasts begin to commit apoptosis, converting granulation tissue in an environment rich in collagen; this is essential for the *re-epithelialization phase,* involving basal *keratinocytes* (highly specialized epithelial cells) migration across the new tissue to form a barrier between the wound and the environment. *Epidermal cell-derived factors (EDF)*, including the **nerve growth factor**, were found to stimulate the proliferation of keratinocytes, but to inhibit fibroblasts.

Keratinocytes must dissolve their desmosomes and hemi-desmosomes, which normally anchor the cells by intermediate filaments in their cytoskeleton to other cells and to the ECM to detach from the basement membrane and enter the wound bed. [22] Like fibroblasts, migrating keratinocytes use the **fibronectin cross-linked with fibrin**.[23] Factors responsible of epithelial cells migration are the cellular proteins belonging to the protein family Annexins (**Annexin A1**, **Annexin A2)**, up-regulated in migrating epithelial cells (IECs). [42] [24] [25] On the other hand, **ephrin-A1**, it has been found that attenuates cell migration when bound EphA2 receptor. [26]

Epithelial cells climb over one another in order to migrate. This growing sheet of epithelial cells is often called the epithelial tongue (Figure 6). [27]

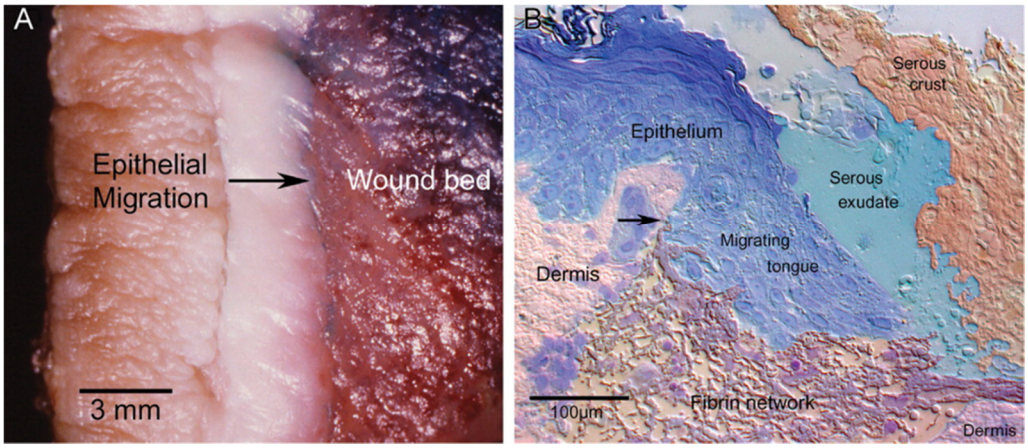


Figure 6 – Topography of human cutaneous wounds. (A) Macroscopic view of epithelial migration over the wound bed of a burn patient. (B) A cross-sectional view of human skin30 hr after incisional wounding. [28]

Cells migrate in two different manners: the first one is chemotaxis; the second type relies on *trans-differentiation or cellular transition*, applied by epithelial cell migration. The tight junctions between cells have to be partially dissociated; then, epithelial cells build up migration machinery with mesenchymal characteristics, which includes ECM components (**fibronectin**, **MMPs**, integrins, intermediate **actin** filaments and **myosin** chains, etc.). Since, keratinocytes at this stage extensively expressed mesenchymal-specific markers, they may be activated and mobilized by *epithelial to mesenchymal phenotype transition* (EMT), which is orchestrated by injury signals like inflammatory cytokines. [29] Several studies established the importance of some members of the **cathepsin (cathepsin Z, cathepsin H)** family in mediating this process. [30]

Keratinocytes themselves also produce and secrete factors, including growth factors (e.g. **Colony stimulating factor 1, Pigment epithelium-derived factor - PEDF)** and basement membrane proteins, which regulate both epithelialization. [31] Most notably, it has been demonstrated that **PEDF** influences keratinocytes at sites of

injury enhancing their adhesion, therefore impairing migration, while having no effect on cell proliferation. [32]

When keratinocytes have finished migrating, they reestablish desmosomes and hemi-desmosomes and become anchored once again to the basement membrane; [22] **desmoplakin** is a protein associated with desmosomes.[33] Basal cells begin to divide and differentiate in the same manner as they do in normal skin to reestablish the strata found in re-epithelialized skin.

An interesting finding in the bovine colostrum insight regards **Galectin-3,** a widely studied protein for its broad biological functionality (cell adhesion, cell activation and chemo-attraction, cell growth and differentiation, cell cycle, and apoptosis). Several studies confirmed to be involved in a variety of processes, including tissue repair, playing a crucial role in modulating *re-epithelialization* of corneal, dermal, intestinal, kidney and skin wounds. Overall, these findings have broad implications for developing novel therapeutic strategies for the treatment of non-healing wounds.

Contraction phase commences when fibroblasts have differentiated into myofibroblasts; such epithelial plasticity is accompanied by dramatic reorganizations of the actin cytoskeleton: **γ-cytoplasmic actin (γ-CYA)** regulates epithelial phenotype and suppression of EMyT. [34]

Myofibroblasts are attracted at wound edges by **fibronectin** and growth factors and they form connections to the ECM at the wound edges by desmosomes, in order to pull the ECM when they contract, reducing the wound size. [23] [35]

Myofibroblasts can contract by using smooth muscle type **actin-myosin complex**. **Myosin X**, found in bovine colostrum, is an unconventional myosin motor, well known to play an important role during cell spreading in cytoskeleton reorganization, focal

contacts formation, and lamellipodial extension. The contraction stage in proliferation ends as myofibroblasts stop contracting and commit apoptosis. [36]

As the final phase of wound healing, the *remodeling phase* is responsible for final scar tissue formation, tightly controlled by regulatory mechanism maintaining a delicate balance between degradation of fibronectin and synthesis of collagen (**type III collagen** is replaced by type I collagen). Synthesis and breakdown of collagen, exerts by **MMPs,** take place continuously. [37]

MMPs activity is regulated and synchronized by **inhibitory factors**. For instance, Vitamin A reduces matrix metalloproteinase expression and stimulates collagen synthesis; so that, *vitamin A deficiency* can impede wound healing. Human skin can specifically and efficiently take up vitamin A from holo-**Retinol-binding protein (**RBP). [38]

As the wound heals, the density of fibroblasts and macrophages is further reduced by apoptosis. The result is a fully matured scar with a decreased number of cells and blood vessels and a high tensile strength. [37]

References

1. Grinnell, F., *Fibronectin and wound healing.* J Cell Biochem, 1984. **26**(2): p. 107-16.
2. Marsh, L.M., et al., *Surface expression of CD74 by type II alveolar epithelial cells: a potential mechanism for macrophage migration inhibitory factor-induced epithelial repair.* Am J Physiol Lung Cell Mol Physiol, 2009. **296**(3): p. L442-52.
3. Kyriakides, T.R. and S. Maclauchlan, *The role of thrombospondins in wound healing, ischemia, and the foreign body reaction.* J Cell Commun Signal, 2009. **3**(3-4): p. 215-25.
4. Piñeiro, M., et al., *ITIH4 (inter-alpha-trypsin inhibitor heavy chain 4) is a new acute-phase protein isolated from cattle during experimental infection.* Infect Immun, 2004. **72**(7): p. 3777-82.
5. French, D., et al., *A preliminary evaluation of the functional significance of alpha-1-acid glycoprotein glycosylation on wound healing.* Biomed Chromatogr, 2002. **16**(6): p. 412-9.
6. Schaefer, B.M., et al., *alpha 2-Antiplasmin and plasminogen activator inhibitors in healing human skin wounds.* Arch Dermatol Res, 1996. **288**(3): p. 122-8.
7. Kyriakides, T.R., et al., *Mice that lack matrix metalloproteinase-9 display delayed wound healing associated with delayed reepithelization and disordered collagen fibrillogenesis.* Matrix Biol, 2009. **28**(2): p. 65-73.
8. Oikarinen, A., et al., *Demonstration of 72-kDa and 92-kDa forms of type IV collagenase in human skin: variable expression in various blistering diseases, induction during re-epithelialization, and decrease by topical glucocorticoids.* J Invest Dermatol, 1993. **101**(2): p. 205-10.
9. Varner, J.A., *The sticky truth about angiogenesis and thrombospondins.* J Clin Invest, 2006. **116**(12): p. 3111-3.
10. Staton, C.A., et al., *Neuropilins in physiological and pathological angiogenesis.* J Pathol, 2007. **212**(3): p. 237-48.
11. Favier, B., et al., *Neuropilin-2 interacts with VEGFR-2 and VEGFR-3 and promotes human endothelial cell survival and migration.* Blood, 2006. **108**(4): p. 1243-50.
12. Fens, M.H., et al., *Angiogenic endothelium shows lactadherin-dependent phagocytosis of aged erythrocytes and apoptotic cells.* Blood, 2008. **111**(9): p. 4542-50.
13. Jiang, Z.L., et al., *Fibroblast growth factor-2 regulation of Sprouty and NR4A genes in bovine ovarian granulosa cells.* J Cell Physiol, 2011. **226**(7): p. 1820-7.

14. Liu, J.T., et al., *Role of pigment epithelium-derived factor in stem/progenitor cell-associated neovascularization.* J Biomed Biotechnol, 2012. **2012**: p. 871272.
15. Stadelmann, W.K., A.G. Digenis, and G.R. Tobin, *Physiology and healing dynamics of chronic cutaneous wounds.* Am J Surg, 1998. **176**(2A Suppl): p. 26S-38S.
16. Yahata, Y., et al., *A novel function of angiotensin II in skin wound healing. Induction of fibroblast and keratinocyte migration by angiotensin II via heparin-binding epidermal growth factor (EGF)-like growth factor-mediated EGF receptor transactivation.* J Biol Chem, 2006. **281**(19): p. 13209-16.
17. Veland, I.R., et al., *Inversin/Nephrocystin-2 is required for fibroblast polarity and directional cell migration.* PLoS One, 2013. **8**(4): p. e60193.
18. Eckes, B., et al., *Fibroblast-matrix interactions in wound healing and fibrosis.* Matrix Biol, 2000. **19**(4): p. 325-32.
19. Seppinen, L., et al., *Lack of collagen XVIII accelerates cutaneous wound healing, while overexpression of its endostatin domain leads to delayed healing.* Matrix Biol, 2008. **27**(6): p. 535-46.
20. Wright, T.E., et al., *The effect of TGF-beta2 in various vehicles on incisional wound healing.* Int J Surg Investig, 2000. **2**(2): p. 133-43.
21. Muller, M.J., et al., *Retardation of wound healing by silver sulfadiazine is reversed by Aloe vera and nystatin.* Burns, 2003. **29**(8): p. 834-6.
22. Santoro, M.M. and G. Gaudino, *Cellular and molecular facets of keratinocyte reepithelization during wound healing.* Exp Cell Res, 2005. **304**(1): p. 274-86.
23. Deodhar, A.K. and R.E. Rana, *Surgical physiology of wound healing: a review.* J Postgrad Med, 1997. **43**(2): p. 52-6.
24. Rankin, C.R., et al., *Annexin A2 regulates β1 integrin internalization and intestinal epithelial cell migration.* J Biol Chem, 2013. **288**(21): p. 15229-39.
25. Babbin, B.A., et al., *Annexin 2 regulates intestinal epithelial cell spreading and wound closure through Rho-related signaling.* Am J Pathol, 2007. **170**(3): p. 951-66.
26. Kaplan, N., et al., *EphA2/Ephrin-A1 signaling complexes restrict corneal epithelial cell migration.* Invest Ophthalmol Vis Sci, 2012. **53**(2): p. 936-45.
27. Bartkova, J., et al., *Cell-cycle regulatory proteins in human wound healing.* Arch Oral Biol, 2003. **48**(2): p. 125-32.
28. Underwood, R.A., et al., *Ultrastructural localization of integrin subunits beta4 and alpha3 within the migrating epithelial tongue of in vivo human wounds.* J Histochem Cytochem, 2009. **57**(2): p. 123-42.
29. Yan, C., et al., *Epithelial to mesenchymal transition in human skin wound healing is induced by tumor necrosis factor-alpha through bone morphogenic protein-2.* Am J Pathol, 2010. **176**(5): p. 2247-58.
30. Tan, G.J., et al., *Cathepsins mediate tumor metastasis.* World J Biol Chem, 2013. **4**(4): p. 91-101.

31. Bayram, Y., et al., *The cell based dressing with living allogenic keratinocytes in the treatment of foot ulcers: a case study.* Br J Plast Surg, 2005. **58**(7): p. 988-96.
32. Chen, L. and L.A. DiPietro, *Production and function of pigment epithelium-derived factor in isolated skin keratinocytes.* Exp Dermatol, 2014. **23**(6): p. 436-8.
33. Garrod, D.R., et al., *Hyper-adhesion in desmosomes: its regulation in wound healing and possible relationship to cadherin crystal structure.* J Cell Sci, 2005. **118**(Pt 24): p. 5743-54.
34. Lechuga, S., et al., *Loss of γ-cytoplasmic actin triggers myofibroblast transition of human epithelial cells.* Mol Biol Cell, 2014. **25**(20): p. 3133-46.
35. Mirastschijski, U., et al., *Matrix metalloproteinase inhibitor GM 6001 attenuates keratinocyte migration, contraction and myofibroblast formation in skin wounds.* Exp Cell Res, 2004. **299**(2): p. 465-75.
36. Birbrair, A., et al., *Type-2 pericytes participate in normal and tumoral angiogenesis.* Am J Physiol Cell Physiol, 2014. **307**(1): p. C25-38.
37. Velnar, T., T. Bailey, and V. Smrkolj, *The wound healing process: an overview of the cellular and molecular mechanisms.* J Int Med Res, 2009. **37**(5): p. 1528-42.
38. Sun, H. and R. Kawaguchi, *The membrane receptor for plasma retinol-binding protein, a new type of cell-surface receptor.* Int Rev Cell Mol Biol, 2011. **288**: p. 1-41.

Part II

Set-up and application of an analytical approach for the quality control of purified colostrum as food supplement

Abstract – A validated analytical procedure is here described for the quality control of the protein fraction of purified bovine colostrum used in food supplements. The proposed procedure starts with 1D and 2D-gel electrophoresis. The sample is then separated into two fractions by protein G affinity chromatography: the IgG enriched and the IgG depleted fraction (IgG-d). A size exclusion chromatography coupled to UV is then applied to the IgG and IgG-d fractions for the quantitative analysis of IgG and IgM, respectively. The IgG-d fraction is then analyzed by HPLC-MS analysis for the quantitative analysis of β-lactoglobulins and α-lactoalbumin. The following step consists of quantitatively measuring a set of bioactive proteins selected from the bovine colostrum data bank on the basis of the claimed health benefits. The enzymatic activities of lactoperoxidase and xanthine dehydrogenase/Oxidase (XDH/XO) is then tested as an index of protein functionality.

1. Introduction

Bovine colostrum (BC) is the initial milk secreted by the mammary gland immediately after parturition and provides the first nutritional components to the new-born calves due to the presence of proteins, fat, lactose, vitamins and minerals [1].

Besides the nutritional effect, colostrum also has a key role in the protection of the newborn calves against pathogens and other postpartum environmental challenges, and this is due to the presence of a complex mixture of antimicrobial peptides and bioactive proteins (i.e. immunoglobulins, lactoferrin, lactoperoxidase and lysozyme) that can stimulate innate antiviral pathways [2, 3].

Some beneficial effects of BC may be shared across species making it effective for the treatment of some human pathologies [4]. Most of the beneficial effects of BC have been ascribed to the protein components consisting of more than 1200 proteins, as recently found by our team [5]. For instance immunoglobulins, which represent the main protein class of colostrum, have direct antimicrobial and endotoxin-neutralizing effects throughout the alimentary tract, playing the role of a direct GI tract defence [6]. BC also has growth-promoting effects on human epithelial cells, which is attributable to the combination of several growth factors, including insulin-like growth factors, basic fibroblast growth factors and platelet-derived growth factor [7].

Based on the presence of several bioactive components, BC has been included in several health products that are claimed to improve the immune system, treat gastro-intestinal infection as well as to maintain the skin and mucous membranes integrity, favouring their healing [8-11]. Moreover, clinical research has demonstrated that α-Lactalbumin in addition to being a highly nutritive protein, plays a central biochemical role as the regulatory subunit for lactose synthesis in the growing neonate [12]; having 72% sequence homology with the human isoform, bovine α-lactoalbumin is

widely added to infant formula-feed [13, 14]. On a commercial level, different bovine colostrum based products are available and they can be differentiated either on the basis of the degree of purification or on the process technologies. For instance, some products derive from a minimally processed raw material in order to maintain the original ingredients, while others are subjected to several steps aimed at removing the major ingredients such as fats, caseins, and lactose. Commercially available bovine colostrum also differs because of the technological processes needed for water removal (spray drying or freeze drying) or inhibiting the growth of bacteria (pasteurization or filtration). Due to the quite heterogeneous matrix and to the different methods of production, the quality control of bovine colostrum used in health products is a quite complex matter. Some validated analytical methods have so far been reported for the qualitative and quantitative analysis of some bioactive ingredients. In general the proposed analytical procedures target specific analytes such as IgG, lactoferrin and growth factors [7] but no integrated assays are yet available and which would be required to give a comprehensive qualitative and quantitative picture of the BC proteome. Such an approach is clearly required not only to guarantee the efficacy but also the safety of the product.

The aim of the present paper is the set-up of an integrated analytical strategy for the quality control of the protein components of bovine colostrum. The procedure we propose starts with 1D and 2D-gel electrophoresis in reducing and non-reducing conditions that permits the mapping of the main protein components and the presence of aggregates as well as of caseins. The sample is then fractionated by affinity chromatography to isolate the IgG purified and depleted fractions and IgG/IgM content is then determined in the two fractions by SEC-UV. The IgG-depleted fraction is then subjected to HPLC-MS analysis for the simultaneous quantitative analysis of

the main protein components, including α-lactoalbumin and β-lactoglobulins (isoforms A and B). The analysis then focuses on a set of bioactive proteins (Lumican, IL-17 and IGF-1) which are selected from the BC database based on the health claim made. Since such proteins are contained in life amounts, ELISA is proposed. Finally, the conformational protein functional integrity is assayed by evaluating the enzymatic activity of two representative enzymes: Lactoperoxidase (LPO) and Xanthine Dehydrogenase/Oxidase (XDH/XO).

2. Materials and methods

2.1 Chemicals

Laemmli buffer, 40% acrylamide/Bis solution, N,N,N',N'-tetramethylethylenediamine (TEMED), molecular mass standards and electrophoresis apparatus for one-dimensional electrophoresis were supplied by Bio-Rad Laboratories, Inc., Hercules CA. β-mercaptoethanol, dithiothreitol (DTT), ammonium persulfate, 3-[3-cholamidopropyl dimethylammonio]-1-propanosulfonate (CHAPS), acetonitrile (ACN), trifluoroacetic acid (TFA), sodium dodecyl sulphate (SDS), iodoacetamide (IAA), formic acid (FA), ABTS (2,2'-azino-di-3-ethylbenzthiazoline-6-sulphonic acid) solution, hydrogen peroxide, all the protein standards used for quantitative analysis (Bovine α-Lactalbumin, β-Lactoglobulin, Immunoglobulin M, Lactoperoxidase, Xanthine Dehydrogenase/Oxidase) and all other chemicals used throughout the experimental work were current pure analytical grade products and purchased from Sigma-Aldrich Corporation, St Louis, MO. Water and acetonitrile (OPTIMA® LC/MS grade) for LC/MS analyses were purchased from Fisher Scientific, UK. Purified IgG were prepared in our laboratory from bovine colostrum collected from Holstein cows until the fifth hour after birth and immediately frozen at -20°C. After a suitable dilution

with demineralized water, the suspension obtained was introduced into a reactor (controlled continuous stirring), where it was heated at 35-36°C for about 30 minutes. The suspension was then subjected to the skimming step, and then caseins were removed by precipitation by adjusting the pH at their isoelectric point (pH 4.6). The product was then ultra-filtered and IgG separated by affinity chromatography and using Protein G Sepharose 4 as stationary phase (see below). Purity of IgG was determined by measuring the water content (about 8.8%) by thermogravimetric analyses using a TGA 2050 thermogravimetrical analyser (TA Instruments, US). A sample of approximately 10 mg was heated in a platinum crucible at 5 K/min from 25 °C to 250 °C under nitrogen atmosphere and the loss of weight was recorded. The densitometric analysis of a 1D-gel electrophoresis was used to determine the presence of protein impurities that accounted for 13.5%: by considering the presence of 1.2% of inorganic material as determined by total ashes measurement, the purity of IgG prepared according our protocol accounted for 76.5%.

Analyses were carried out on commercially available BC raw materials sold as ingredient for the formulation of healthy products/food supplement; according to the vendors' data sheet the different BC raw materials were produced as follows: BC1 produced by a low temperature pasteurisation and freeze-drying process without the removal of any components; BC2 was defatted and spray-dried; BC3 no information on processing furnished; BC4 and BC5 were skimmed, decaseinated, micro-iltered, concentrated and freeze-dried. We also considered not processed BC which was collected from Holstein cows until the fifth hour after birth and immediately frozen at -20°C (BC0).

2.2 Separative methods of proteins on polyacrylamide gel

One-dimensional analysis (SDS-PAGE) - Protein separation was performed under reducing conditions; aliquots of 10µL of samples containing 20-25 µg of proteins were mixed with 10µL of Laemmli sample buffer containing 50 mM DTT and heated at 95°C for 5 minutes. Samples and the standard proteins mixture (Precision Plus Protein Standards) were loaded on precast gels (Any KD ™ Mini Protean® TGX ™) and then placed in the electrophoresis cell (Mini-PROTEAN Tetra) and run at 200 V (constant) for a variable time of about 30-40 min. Staining was carried out using Coomassie blue stain (Biosafe G250 Stain, Bio-Rad) and the images acquired by using the Bio-Rad GS800 densitometer and analyzed by using the software quantity One 1-D.

Two-dimensional gel electrophoresis - 150 µg of each sample were solubilised in the denaturing buffer (7 M Urea, 2 M thiourea, 40 mM Tris, 3% CHAPS) and incubated for one hour at room temperature with 5 mM TCEP in order to reduce the protein disulphide bonds. After spiking with DESTREAK (150 mM), ampholine (0.5%) and blue bromophenol, samples were loaded on IPG strips (Bio-Rad 7 cm, pH gradient 3-10); after incubation (4 hours) strips were laid in the strip housing for isoelectric focusing (IEF, I12 Protean IEF Cell-Bio-Rad) and covered with mineral oil (Bio-Rad) to prevent proteins oxidation. The instrument was set to follow the voltage gradient reported in Table 1 (Part I). Once the IEF was completed, strips were incubated with 2.5 mL of equilibration buffer (6 M urea, 2% SDS, 0.05 M Tris-HCl pH 8.8 and 20% glycerol) on gentle shacking for 10 minutes. Proteins separation based on their molecular mass was obtained by one-dimensional electrophoresis (SDS-PAGE) using precast gel (Any KD ™ Mini Protean® TGX ™). The strips were laid on the top

of the gel, orienting them so that pI followed an ascending order (from 3 to 10) from left to right, and coated with a layer of agarose (0.5%). The electrophoretic device was assembled as described above and the electrophoretic run was carried out at 200 V for 30 min. After electrophoresis, gels were washed and stained, de-staining and acquisitions carried out as described in the previous paragraph.

2.3 Bovine IgG purification

Affinity Chromatography – IgG depleted BC was prepared by removing the IgG fraction by affinity chromatography. The affinity column was prepared by packing 400 mL of Protein G Sepharose 4 Fast Flow resin (GE Healthcare) in a column support HiScale ™ 50 (GE Healthcare) which was connected to an FPLC system (ÄKTAprime plus, GE Healthcare line-up). The chromatographic purification started by eluting the column with five volumes (5 x chromatographic bed volume) of buffer A (binding buffer: 20 mM sodium phosphate, pH 7), followed by sample loading at a flow rate of 20 ml min^{-1}. The eluate was monitored at 280 nm and all the fractions characterized by a significant UV absorption were automatically collected. The subsequent step consisted of recovering the IgG fraction by eluting the column with 100% of elution buffer (1 M glycine hydrochloride pH 2.5).

Tangential Flow Filtration – The collected fractions were mixed and subjected to concentration and desalting using hollow fibers cross flow filtration cartridges with 3000 NMWC (Nominal Molecular Weight Cutoff) and a surface area of 650 cm^2 (GE Healthcare) coupled to a tangential flow filtration system equipped with a peristaltic pump essential to keep the flow recirculation continuous (Kross Flo®- Tangential flow Filtration System Research III). The IgG depleted fraction was concentrated 20/30 times, dialysed with 5 volumes and then lyophilized.

2.4 Profiling most abundant species

2.4.1 SEC-UV for IgG and IgM analysis

A size exclusion chromatography (SEC) method was set-up for the quantitative analysis of high-MW proteins (IgG and IgM); the technique is based on the separation of molecules depending on their size. SEC was performed on a Thermo Finnigan Surveyor HPLC system (ThermoFinnigan Italia, Milan, Italy) equipped with a variable wavelength detector and an auto-sampler, controlled by Xcalibur software (version 2.0.7 Thermo Fisher Scientific, Rodano, MI, Italy). The SEC separation was performed on a 4.6 × 300 mm Phenomenex Yarra™ 3u SEC-3000, with a 4 × 3 mm GFC4000 pre-column, by running an isocratic flow of mobile phase containing 0.1M sodium phosphate bibasic, 0.025% sodium azide pH 6.8, at a constant flow rate of 0.5 mL min^{-1}. The autosampler temperature was set at 8°C and UV detection was conducted at a wavelength of 280 nm, a typical wavelength for protein detection.

Preparation of standard - the dipeptide Tyrosine-Histidine (TH) was chosen as internal standard (IS). For calibration standards preparation, IgG standard was weighed and dissolved in the mobile phase (0.1M sodium phosphate bibasic, 0.025% sodium azide, pH 6.8) to obtain a solution of 1mg mL^{-1}. The stock solution was then diluted to obtain working solutions at concentrations of 0.05, 0.10, 0.25, 0.50, 1.00, 2.00, and 4.00 mg mL^{-1}. The IgM stock solution (nominal concentration of 1mg/mL) was diluted in mobile phase, to obtain working solutions at concentrations of 0.020, 0.050, 0.125, 0.250, 0.375, 0.500, 0.750, and 0.900 mg mL^{-1}. All the stock solutions were added by a fixed amount of the dipeptide Tyrosine-Histidine chosen as the internal (0.1 mg mL^{-1}).

Sample Preparation - The two fractions eluted from the purification step (depleted and IgG purified) were used for the IgG and IgM quantification in the bovine colostrum. Prior to SEC analysis the fractions were spiked with a fixed amount of the internal standard, 0.1 mg mL^{-1} TH.

Method validation - The method was validated for linearity, lower limit of quantification (LLOQ), precision, and accuracy, by using as standard proteins the bovine IgM purchased from Sigma Aldrich, and lyophilized IgG (purified by affinity chromatography from bovine colostrum as above described).

Linearity and carry over effect - To evaluate the linearity of the method eight calibration standards for each analyte were studied over a calibration range of 0.05-4.00 mg mL^{-1} (IgG) and 0.02-0.90 mg mL^{-1} (IgM). Each calibration curve was built considering a blank sample, a zero sample (TH 0.1 mg mL^{-1}) and eight calibration levels, among which blank and zero samples were necessary to confirm the absence of interferences. The calibration curves were realized by plotting peak area ratios of calibration standards to the IS versus the nominal molarity of analytes. Carry-over effects were assessed, with relevant criteria by injecting samples followed by the calibration standard at the highest concentration. The limit of detection (LOD) was calculated using a signal to noise (S/N) ≤ 3.

Precision and accuracy - The precision of the method was given by the coefficient of variation (CV). The acceptable limit of CV was <20% for the LLOQ (corresponding to the low quality control, LQC), and ≤15% for medium quality control (MQC) and high quality control (HQC) samples. The acceptable limit of BIAS % (accuracy) should not

exceed ±20% and ±15%for the LLOQ (LQC) and all the other concentrations, respectively. To investigate the intra- and inter-day precision and accuracy, for IgG quantitation seven calibration standards in the range of 0.05-4.00 mg mL^{-1} and three quality control samples LQC (0.1 mg mL^{-1}), MQC (1.0 mg mL^{-1}), and HCQ (4.0 mg mL^{-1}) were prepared separately in 6 replicates of each concentration, using chromatography for three consecutive days. The same procedure was used for IgM quantitation: in the range of 0.02-0.90 mg mL^{-1} and three quality control samples LQC (0.125 mg mL^{-1}), MQC (0.250 mg mL^{-1}), and HCQ (0.900 mg mL^{-1}) were prepared separately in 6 replicates. Both the accuracy and CV should be respect the imposed limits, as above reported. The data obtained for the intra-day assay accuracy and precision were re-analyzed for inter-day assay accuracy and precision using all three days' data for each concentration.

2.4.2 Profiling intact proteins by HPLC-MS - TOF as MS analyser

The BC sample (lyophilized) was solubilized in 50% of mobile phase A (0.1% formic acid in H_2O) and 50% of mobile phase B (0.1% HCOOH in CH_3CN) to obtain a solution with an approximate concentration of 2µg/µl. Aliquots of 10 µl of sample were injected on a ProSwift RP-H4 column, 1 x 250mm and using a Dionex Ultimate 3000 RSLC as HPLC system. The rate of sample infusion into the mass spectrometer was 300µL min^{-1}, and the separating gradient ramped linearly from 5% to 95% of mobile phase B in 30 min. The eluting proteins were on-line sprayed in a Bruker maXis plus – Time of flight - mass spectrometer by an-ESI source; the profile spectra were acquired in positive ion mode in the mass ranges 750-5000 *m*/*z* and 300-3000 *m*/*z*, with a resolving power of 50,000. External mass calibration was performed prior to

each run. Data were analyzed with the instrument software Bruker Daltonics Data Analysis.

2.4.3 Profiling intact proteins by HPLC-MS - Ion trap as MS analyzer

Reversed phase LC-UV/MS experiments were performed on a Thermo Finnigan Surveyor HPLC system (ThermoFinnigan Italia, Milan, Italy) equipped with an online degasser, a gradient pump, an autosampler, and a variable wavelength detector, coupled to an LCQ Advantage mass spectrometer through a Finnigan IonMax electrospray ionization (ESI) source. Separation was conducted on a Jupiter C4 300A column (2 × 150 mm, 5 µm) with a Widepore C4 pre-column (4 x 3.0mm), at a flow rate of 0.25 mL min^{-1} with the column temperature set at room temperature. Mobile phase A was 0.1% formic acid in water and B is 0.1% formic acid in acetonitrile. Total run time was 24 min for each sample injection; gradient elution was set as follows: mobile phase B from 30% to 45% in 3 min and from 45% to 60% mobile in 13 min. At the end of the analysis, the column was washed by using 95% B for 3 min, followed by re-equilibration using 30% B for 5 min. The UV detection wavelength was set to acquire in a range 200-600nm (channel A 280 nm). MS experimental conditions were: the capillary temperature was 275°C, source voltage 3.5kV, sheath gas flow was 35 arb and sweep gas flow was 10 arb. The profile spectra were acquired in positive ion mode in the mass range of 300-2000 m/z in profile mode. Mass calibration was performed prior to all the analysis. Full instrument control and data analysis were provided by Xcalibur software (version 2.0.7, Thermo Fisher Scientific, Rodano, MI, Italy). For each analysis 5µl of sample solution was injected by an autosampler.

Preparation of standard – For the preparation of the calibration standard solutions, α-lactoalbumin, β-lactoglobulins ,A/B and the internal standard, were separately weighed and dissolved in the mixture of 0.1% HCOOH in water and 0.1% HCOOH in CH_3CN (70/30 v/v) to obtain a stock solution of 60 µM. These stock solutions were then diluted and mixed to obtain working solutions at concentrations of 0.5, 1.0, 2.0, 5.0, 10.0, 15.0, and 20.0 µM. All the stock solutions were added by a fixed amount of ubiquitin (11.6 µM) chosen as internal standard.

Sample Preparation - All the BC fractions obtained by the IgG depletion conducted on 4 different batches were spiked with 11.6 µM of IS. Samples were desalted and filtered by using Amicon Ultra-0.5 mL Centrifugal Filters (MWCO 3000Da) prior the LC-MS analysis.

Linearity and carry over effect - Calibration curves were built considering a blank sample, a zero sample (ubiquitin 11.6µM) and seven calibration levels (5-50µM for β-lactoglobulins and 0.5-20.0 µM for α-lactoalbumin), among which blank and zero samples were necessary to confirm the absence of interferences. To simplify the method and to reduce the time of analysis we resorted to a combined calibration curve obtained by adding both the proteins in the same calibration standards over a range of 0.5-20.0 µM. The calibration curves were obtained by plotting peak area ratios of calibration standards to the IS versus the nominal molarity of analytes. Carry-over effects were assessed, with relevant criteria by injecting samples followed by the calibration standard at the highest concentration. The limit of detection (LOD) was calculated using a signal to noise ratio (S/N) ≤3.

Precision and accuracy - The precision of the method was given by the coefficient of variation (CV%). The acceptable limit of CV was <20% for the LLOQ (corresponding to the LQC), and ≤15% for medium quality control (MQC) and high quality control (HQC) samples. Accuracy was determined as the percentage of difference between the mean calculated concentration and the theoretical value (BIAS %). The acceptable limit of BIAS percentage (%) should not exceed ±20% and ±15% for the LLOQ (LQC) and all the other concentrations, respectively. To investigate the intra- and inter-day precision and accuracy, precision seven calibration standards in the range of 0.5–20.0 µM and three quality control samples LQC (0.5µM), MQC (2.0µM), and HCQ (20.0µM) were prepared separately in six replicates of each concentration. Both the accuracy and CV should be respect the imposed limits, as above reported. The data obtained for the intra-day assay accuracy and precision were re-analyzed for inter-day assay accuracy and precision using all three days' data for each concentration.

2.5 Targeting bioactive proteins by ELISA assay

Lumican, IL-17 and IGF-1 contents in BC were measured using a commercially available ELISA kits (Cell Biolabs Inc. Valter Occhiena S.R.L.) according to the manufacturer's instructions. Spectrofluorimetric readings were performed using a Wallac Victor 2 reader (Perkin Elmer).

2.6 Protein Functionality

2.6.1 Lactoperoxidase Activity Assay - Spectrophotometric method

Lactoperoxidase activity was determined according to the method previously described by Shindler et al [15].. Briefly, 10 µL of the sample (standards and unknown

samples) were added to 2.0 ml of substrate solution (1 mM ABTS in 0.1 M acetate buffer pH 4.4) and mixed thoroughly. A 570µL portion was transferred to each of two wells (replicates) in the multiplate, which was put in the spectrophotometer. The experiment was controlled that a stable and straight base line is obtained. Then, 5µl of 5mM H_2O_2 solution, prepared fresh, were added to the sample and mixed carefully in the multiwall reader Wallac Victor 2 reader (Perkin Elmer). The increase in absorbance was then measured for 3 to 5 min at 413 nm (405±15nm filter). The increase in absorbance/min (ΔA413 min^{-1}) was calculated drawing the tangent to the initial activity (1 minute). The activity of the enzyme is defined as: 1 unit (U) of lactoperoxidase activity is the amount of enzyme catalysing the oxidation of 1 µ-mole of substrate per min. The calibration curve was prepared by considering six standard concentrations of the enzyme (0.3, 0.4, 0.5, 0.6, 0.8 and 1.0U). The calibration curve was realized by plotting the ΔA413 min^{-1} (measured in the first minute of the reaction), *versus* the nominal concentration of the enzyme. The linearity was assessed considering the coefficient of linear regression obtained.

2.6.2 Determination of Xanthine Dehydrogenase/Oxidase Activity - Fluorimetric Assay

Xanthine Dehydrogenase/Oxidase (XDH/XO) efficiently metabolizes pterin (2-amino 4-hydroxypteridine) into a fluorescent product, isoxanthopterin (2-amino-4,7-pterinediol), which enables the highly sensitive quantitation of the enzyme's activity. Since the enzymatic activity of XDH/XO can greatly depends on the sample's matrix, a quantitative method based on standard addition was employed. For the analysis four different samples were prepared and analyzed simultaneously by using a four positions cuvettes holder. Samples were prepared by mixing BC with pterin as

substrate (40μM final concentration), in absence and presence of three increasing amounts of the enzyme standard solution to obtain the following concentrations: 0.005, 0.010 and 0.020 U ml^{-1}. Samples were then diluted to a final volume of 2.2ml of tris buffer (pH 7.8, 50 mM). The reaction mixtures were monitored for 30 min to measure the amount of isoxanthopterin produced, whose fluorescence intensity was measured by using the Luminescence Spectrometer LS 50B (Perkin Elmer) and setting the excitation and emission wavelengths at 345 nm at 390 nm, respectively.

2.6.3 Determination of Xanthine Dehydrogenase/Oxidase Content - ELISA Assay

Quantitative measurement of XDH/XO was carried by ELISA according to the suppliers' manuals (Cloud-Clone Corp., Houston, USA).

3. Results

3.1 Profiling intact proteins by electrophoresis and MS

3.1.1 1D and 2D gel electrophoresis of commercially available BC raw materials

Figure 1 displays the SDS-PAGE profile (non-reducing conditions) of a set of 5 commercially available BC raw materials used in healthy products, differing on the basis of the process technology and final composition, as reported by the suppliers.

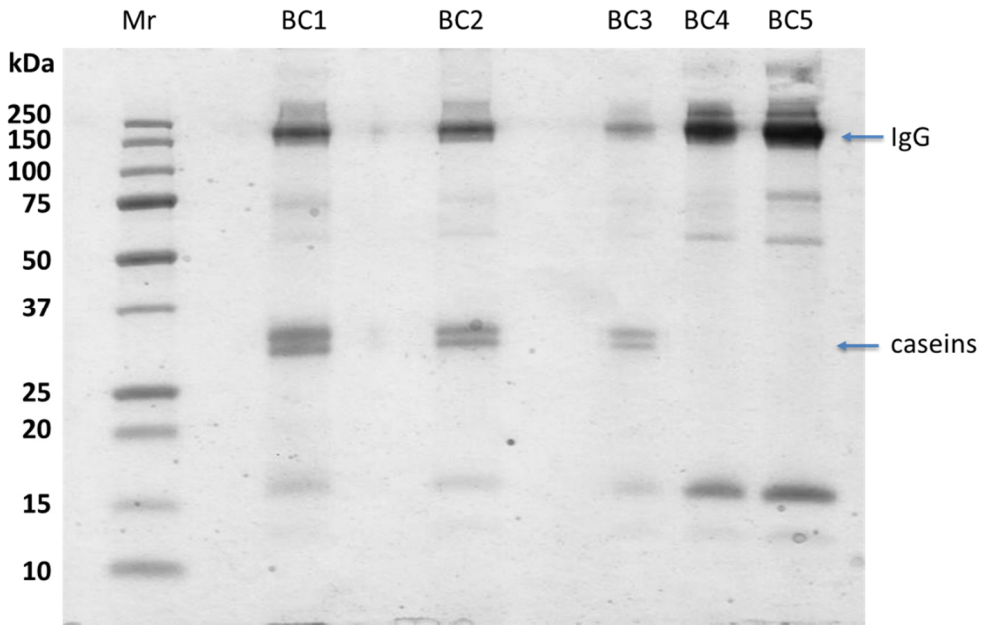


Figure 1 – SDS-PAGE profiling of colostrum proteins. Mr: molecular mass ladder. Samples: BC commercial products (BC1, BC2 and BC3 claimed to be minimally processed thus maintaining the original composition, BC4 and BC5 were defatted and decaseinated); gel was run in non-reducing conditions and stained with colloidal Coomassie blue.

In particular, samples BC1, BC2 and BC3 are claimed to be minimally processed thus maintaining the original composition while samples BC4 and BC5 were defatted and decaseinated. The mono-dimensional gel electrophoresis of all the samples shows a main single, heavy band of IgG centered at ca. 150 kDa, which in reducing conditions disappears in favor of two bands at ca. 50 and 25 kDa, representing the heavy and light IgG chains, respectively (data not shown). Low intensity bands at higher MW are also observed which can be attributed to IgA (mol. wt. about 410.000) and IgM (mol. wt. about 900.000). Samples BC1, BC2 and BC3 but not BC4 and BC5 show two intense bands between 25 and 30 kDa attributed to caseins. Sample BC4 depleted by caseins and defatted was then analyzed by 2D-gel electrophoresis.

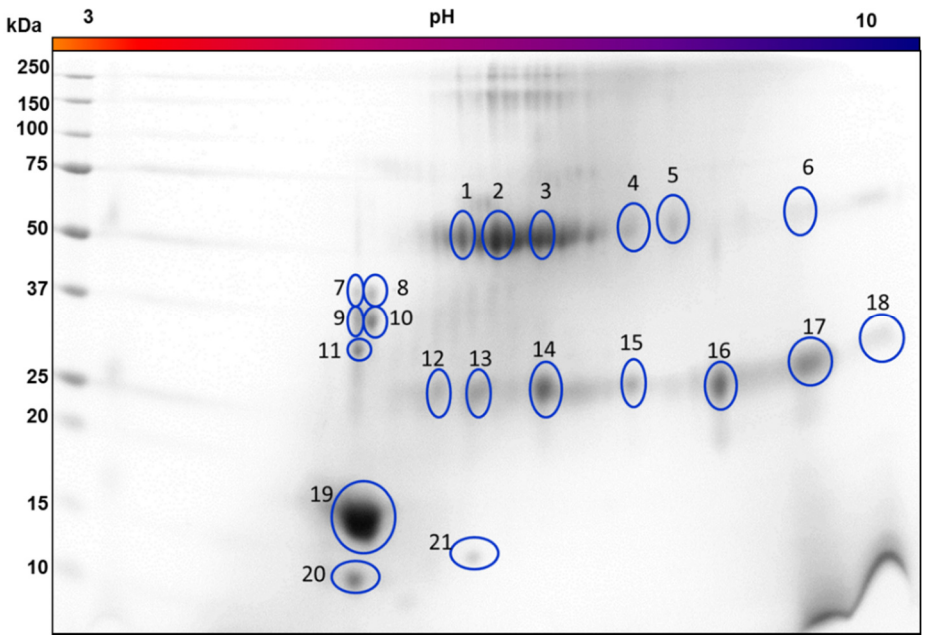


Figure 2 - Two-dimensional map of BC4 (casein-depleted and defatted bovine colostrum). The 21 numbered and circled spots have been cut out and subjected to MS analysis after trypsin digestion. Staining with colloidal Coomassie blue. Identified proteins are listed in table 2.

The sample is characterized by few rather intense zones in addition to all those spots corresponding to the isoforms of immunoglobulin subunits, which, upon cutting the gel (Figure 2) were identified via MS as the following well known milk proteins: α-lactalbumin (14.1 kDa), ß-lactoglobulin (19.9 kDa), serum transferrin (77.7 kDa) and α-S1-casein (24KDa) in traces. All the identified gene product are reported in table 2. Sample 4 was then used for further analysis as here below reported.

SPOT	PROTEIN NAME	ACC. NUM.	SCORE	COVERAGE	N° PEPT.	MW	P.I.	OMOLOGY
1	serum albumin	P02769	76.63	26.19	10	69KDa	6.18	
	Ig heavy chain V-II region ARH-77	G3MXG6	17.63	38.85	3	15KDa	9.45	50% Human
2	serum albumin	P02769	176	43.82	18	69KDa	6.18	
	Ig gamma-4 chain C region	G3N0V0	68.44	9.2	3	36KDa	7.84	50% Human
	Ig heavy chain V-II region ARH-77	G3MXG6	21.19	23.74	2	15KDa	9.45	50% Human
	Ig lambda-2 chain C regions	F1MCF8	18.47	21.37	3	24KDa	7.58	50% Human
3	Ig gamma-4 chain C region	G3N0V0		39.57	10	36KDa	7.84	50% Human
	Ig heavy chain V-II region ARH-77	G3MXG6		23.74	2	15KDa	9.45	50% Human
	Polymeric immunoglobulin receptor	P81265		8.72	4	82KDa	7.27	
4	Ig gamma-4 chain C region	G3N0V0	56.79	24.23	7	36KDa	7.84	50% Human
	Beta-lactoglobulin	P02754	18.76	32.58	4	20KDa	5.02	
	Ig heavy chain V-II region ARH-77	G3MXG6	14.72	25.9	2	15KDa	9.45	50% Human
5	Ig gamma-4 chain C region	G3N0V0	24.7	15.64	3	36KDa	7.84	50% Human
6	Ig gamma-4 chain C region	G3N0V0	27.88	20.86	5	36KDa	7.84	50% Human
	Beta-lactoglobulin	P02754	14.2	15.73	2	20KDa	5.02	
7	Beta-lactoglobulin	P02754	123.7	54.49	13	20KDa	5.02	
	Zinc-alpha-2-glycoprotein	Q3ZCH5	30.45	12.37	4	34KDa	5.24	
	Actin, cytoplasmic 1	P60712	22.97	20	4	42KDa	5.48	
8	Beta-lactoglobulin	P02754	182	62.92	17	20KDa	5.02	

	Zinc-alpha-2-glycoprotein	Q3ZCH5	84.71	31.77	13	34KDa	5.24	
9	Beta-lactoglobulin	P02754	275.7	77.53	18	20KDa	5.02	
10	Beta-lactoglobulin	P02754	163.1	54.49	13	20KDa	5.02	
11	Beta-lactoglobulin	P02754	165.1	60.67	15	20KDa	5.02	
12	Ig kappa chain C region	F1MH40	70.79	36.67	8	26KDa	5.94	50% Human
	Ig lambda-2 chain C regions	F1MLW7	62.75	31.2	5	24KDa	7.62	50% Human
	Beta-lactoglobulin	P02754	35.85	41.57	5	20KDa	5.02	
	Complement C3	G3X7A5	35.28	3.73	6	187KDa	7	
	Prostaglandin-H2 D-isomerase	O02853	11.53	13.61	2	21KDa	6.90	
13	Ig kappa chain C region	F1MZ96	134.2	42.5	11	26KDa	6.90	50% Human
	Complement C3	G3X7A5	76.36	6.2	12	187KDa	7	
	Ig lambda-2 chain C regions	F1MLW7	70.32	31.2	4	24KDa	7.62	50% Human
	Ig lambda-2 chain C regions	F1MLW8	56.87	22.32	2	24KDa	5.91	50% Human
	Beta-lactoglobulin	P02754	25.14	32.58	4	20KDa	5.02	
	Prostaglandin-H2 D-isomerase	O02853	13.91	13.61	4	21KDa	6.90	
14	Ig lambda-2 chain C regions	F1MLW7	409.8	58.97	10	24KDa	7.62	50% Human
	Ig lambda-2 chain C regions	F1MLW8	229.2	22.32	5	24KDa	5.91	50% Human
	Ig kappa chain C region	F1MZ96	193.1	42.5	12	26KDa	6.90	50% Human
	Complement C3	G3X7A5	34.61	4.7	6	187KDa	7	
	Beta-lactoglobulin	P02754	17.65	24.72	3	20KDa	5.02	
15	Ig kappa chain C region	F1MH40	106.1	38.33	9	26KDa	5.94	50% Human
	Ig lambda-2 chain C regions	F1MLW7	68.66	34.62	7	24KDa	7.62	50% Human
	Ig lambda-2 chain C regions	F1MLW8	40.63	17.6	5	24KDa	5.91	50% Human

	Beta-lactoglobulin	P02754	16.62	21.35	3	20KDa	5.02	
	Ig gamma-4 chain C region	G3N0V0	15.68	11.96	3	36KDa	7.84	50% Human
16	Ig lambda-2 chain C regions	F1MLW7	143.5	31.2	7	24KDa	7.62	50% Human
	Ig lambda-2 chain C regions	F1MCF8	137.6	27.78	7	24KDa	7.58	50% Human
	Ig lambda-2 chain C regions	F1MLW8	106.6	22.32	4	24KDa	5.91	50% Human
	Beta-lactoglobulin	P02754	16.47	32.58	4	20KDa	5.02	
	Ig kappa chain C region	F1MH40	15.11	14.58	2	26KDa	5.94	50% Human
17	Ig lambda-2 chain C regions	F1MLW7	96.71	29.91	10	24KDa	7.62	50% Human
	Ig lambda-2 chain C regions	F1MLW8	57.04	22.32	5	24KDa	5.91	50% Human
	Ig kappa chain C region	F1MH40	30.66	14.58	3	26KDa	5.94	50% Human
18	Ig lambda-2 chain C regions	F1MLW7	129	33.33	8	24KDa	7.62	50% Human
	Ig lambda-2 chain C regions	F1MLW8	114.7	22.32	5	24KDa	5.91	50% Human
	Beta-lactoglobulin	P02754	90.81	46.07	11	20KDa	5.02	
	Ig kappa chain C region	F1MH40	16.08	14.58	2	26KDa	5.94	50% Human
	Alpha-S1-casein	P02662	9.05	10.28	2	24KDa	5.02	
19	Beta-lactoglobulin	P02754	770.8	86.52	24	20KDa	5.02	
20	Alpha-lactalbumin	P00711	76.14	38.73	8	16KDa	5.14	
	Beta-lactoglobulin	P02754	71.36	46.07	11	20KDa	5.02	
21	Beta-lactoglobulin	P02754	8.48	17.42	2	20KDa	5.02	

Table 2 - BC proteins identified by trypsin digestion and MS analysis of the spots circled in the 2D gel electrophoresis shown in Fig. 2.

3.1.2 Profiling intact proteins by LC-ESI-MS

Profiling of intact proteins of BC4 was then carried out by LC-ESI-TOF. Figure 3a shows the average MS spectrum relative to the Total Ion Chromatogram (TIC) which is characterized by a not resolved broad peak from *m/z* 2740 to 4250 as magnified in Figure 3b, absent in the IgG depleted fraction, and attributed to the polyclonal immunoglobulins. From such a broad peak a series of more defined peaks arises and, as determined by the deconvoluted spectrum (Figure 3c), were assigned to proteins with an MW from 112 kDa to 163 kDa. Several abundant proteins that we identified in colostrum fall into this MW range such as two different isoforms of xanthine dehydrogenase/oxidase (146.6Kda), thrombospondin-1/2 (129Kda), complement factor H (140KDa), two different isoforms of alpha-mannosidase (130KDa), G-protein coupled receptor 126 (132KDa), aldehyde oxidase (147KDa) [5]. Figure 4a reports the TIC of BC4 depleted of IgG determined by LC-ESI-MS-TOF, which identifies at least 9 peaks.

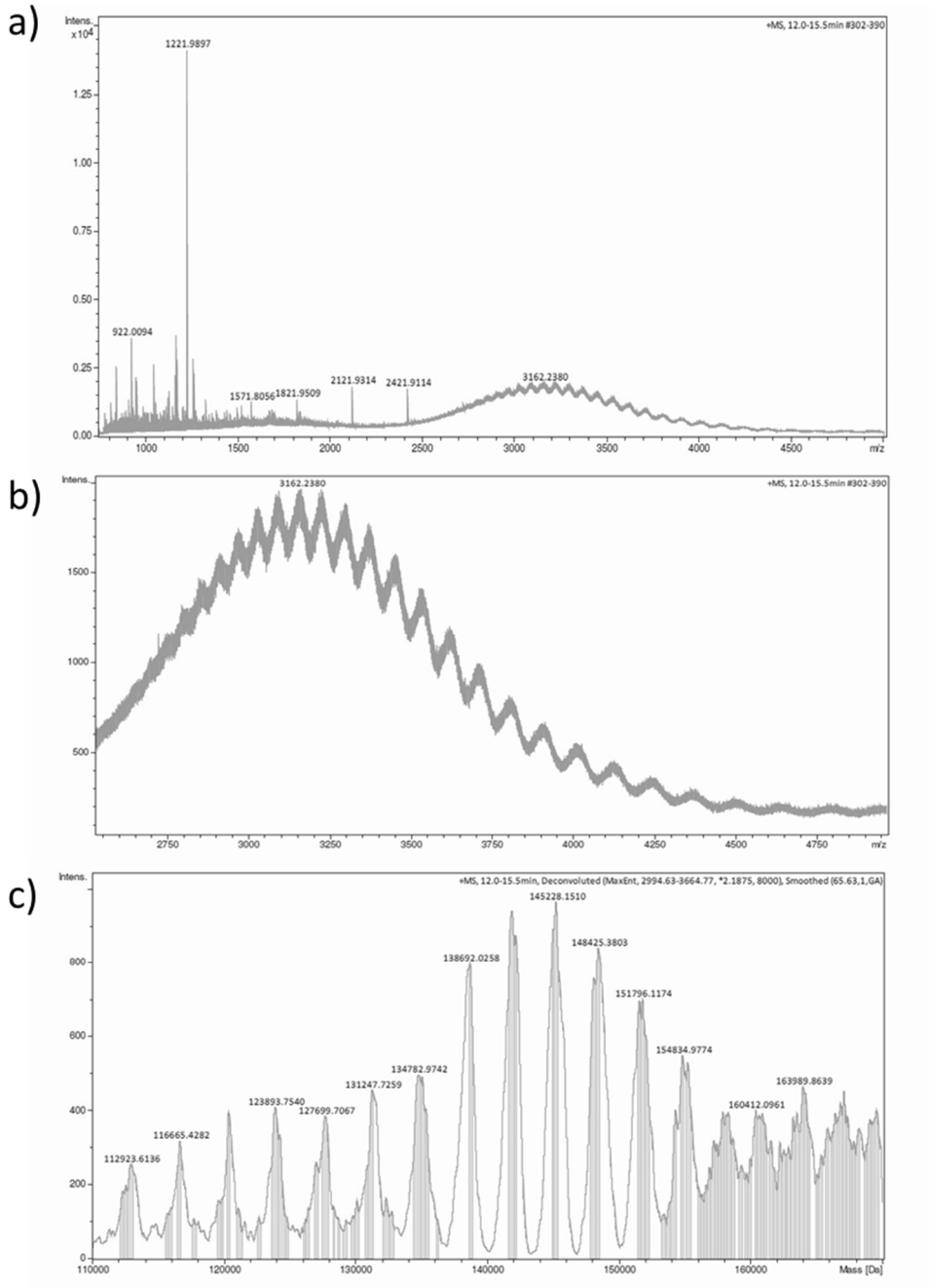


Figure 3 – LC-ESI-MS-TOF analysis of BC4 (casein-depleted and defatted bovine colostrum). Panel a) average MS spectrum relative to the Total Ion Chromatogram (TIC); panel b) magnified MS spectrum; panel c) deconvoluted MS spectrum.

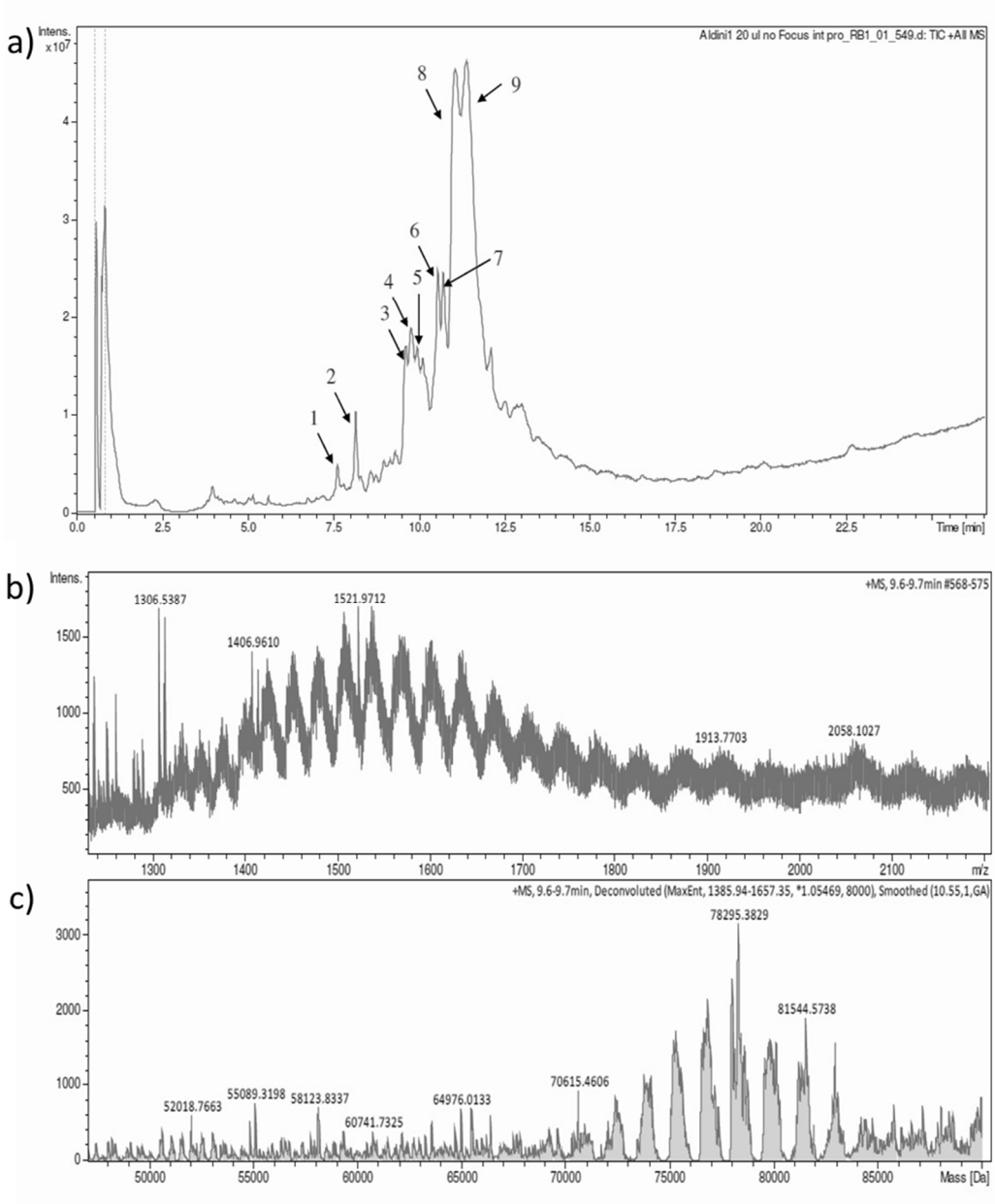


Figure 4 – LC-ESI-MS-TOF analysis of BC4 depleted of IgG. Panel a) Total ion current (TIC); panel b) MS spectrum relative to the peak eluting at 9.6 – 9.7 min (peak 3); panel c) deconvolution analysis of the MS spectrum showed in panel b) and referred to transferrin.

The MS spectrum of each peak was then analyzed and the corresponding deconvoluted spectrum determined to retrieve the proteins' MW. Figure 4b shows the MS spectrum relative to peak 3 eluting between 9.6 and 9.7 min and characterized by well-defined multicharged ions. Deconvoluted analysis (Figure 4c) showed an MW centered at 78 kDa and assigned to lactotransferrin. Analyzing all the peaks in the chromatogram (TIC - Fig. 4a) by using the same approach other proteins were identified, namely: cytochrome c (11.7KDa), proline-rich protein 9 (12.8KDa), β-2-microglobulin (13.6KDa), α-lactalbumin (16 Kda - active form 14KDa), β-lactoglobulin (isoform B 18.2KDa, isoform A 18.3KDa), serotransferrin (77.7KDa). Notably by using an ion trap analyzer or an Orbitrap system only proteins with an MW lower than 30 kDa were identified.

3.2 Affinity chromatography – IgG and IgG-depleted BC fractions

As shown in Figure 1 relative to 1D-gel electrophoresis, IgG are the main components of BC with a relative concentration as determined by densitometry analysis of more than 80%. Due to such a high content, IgG were found to interfere with the quantitative analysis of minor analytes such as IgM, β-lactoglobulin and α-lactalbumin (data not shown). For this reason, an affinity chromatography based on protein G was carried out to prepare an IgG depleted fraction to be used for the quantitative analysis of the minor protein components. Moreover, the IgG enriched fraction was used for quantitative analysis of IgG.

The affinity chromatogram displayed two peaks, the first, not retained, corresponding to the non-immunoglobulin protein fraction and the second to the IgG fraction (Figure 5).

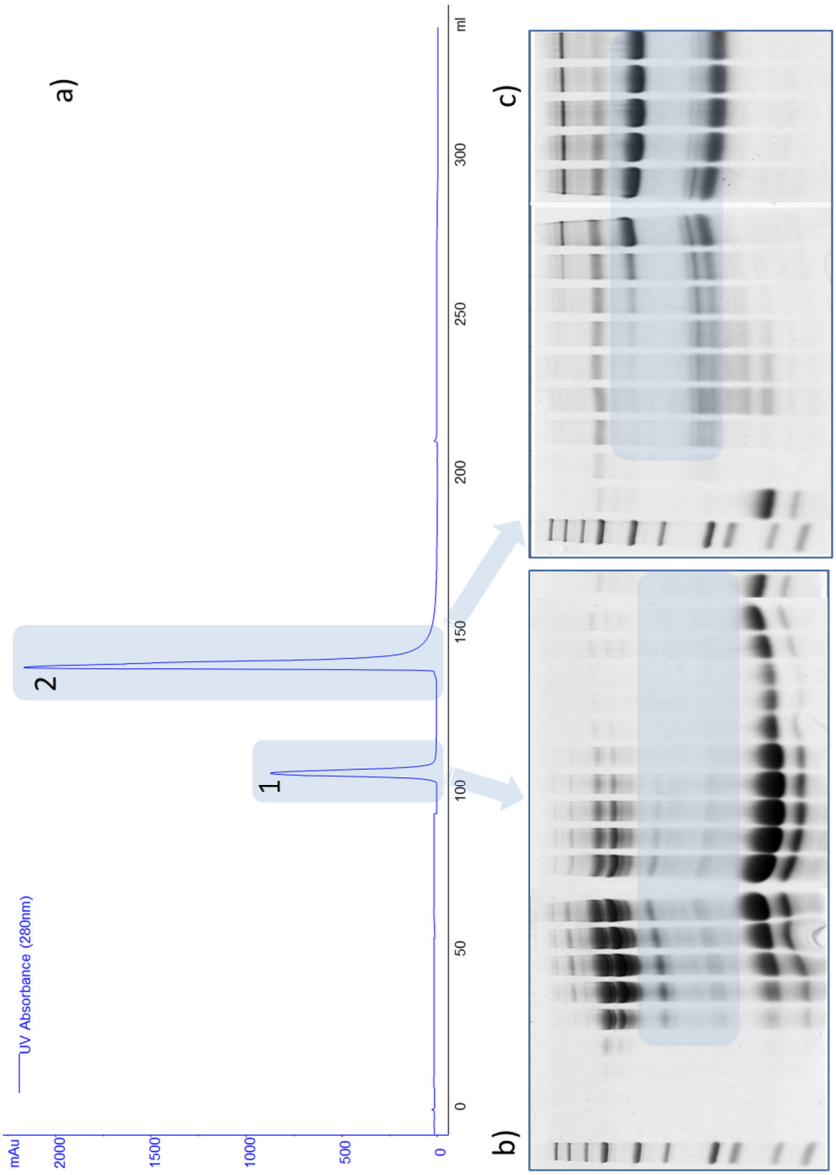


Figure 5 – IgG purification by affinity chromatography. Panel a) FPLC-UV chromatogram of BC4 (casein-depleted and defatted bovine colostrum) obtained by using protein G as affinity stationary phase; peaks 1 and 2 are attributed to the IgG-depleted and IgG fractions as demonstrated by the SDS-gel electrophoresis patterns. Panels b) and c) report the gel electrophoretic pattern of peak 1 and 2 run in reducing conditions. The characteristic bands at 25 and 50 KDa are well evident in peak 2 and only in negligible amounts in peak 1.

According to the relative abundance of IgG in colostrum (80%), the second peak was the most abundant. The electrophoretic patterns obtained for each collected fraction (Figure 5) indicate a good depletion of immunoglobulins, whose two characteristic bands in reducing conditions at 50 and 25 kDa are very weak in the aliquots eluted within the first peak, and clearly much more intense in the fractions of the second peak corresponding to the immunoglobulins.

3.3 IgG and IgM analysis by SEC-UV

After optimization of the chromatographic conditions, a suitable separation of IgG and IgM in BC was achieved by SEC-UV.

3.3.1 Immunoglobulin G analysis

Figure 6 a-b-c represents typical SEC-UV chromatograms relative to IgG standard at different concentration levels (LOD 0.05, MQC 1.00, HQC 4.00 mg mL^{-1}), spiked with TH as IS (0.1 mg mL^{-1}).

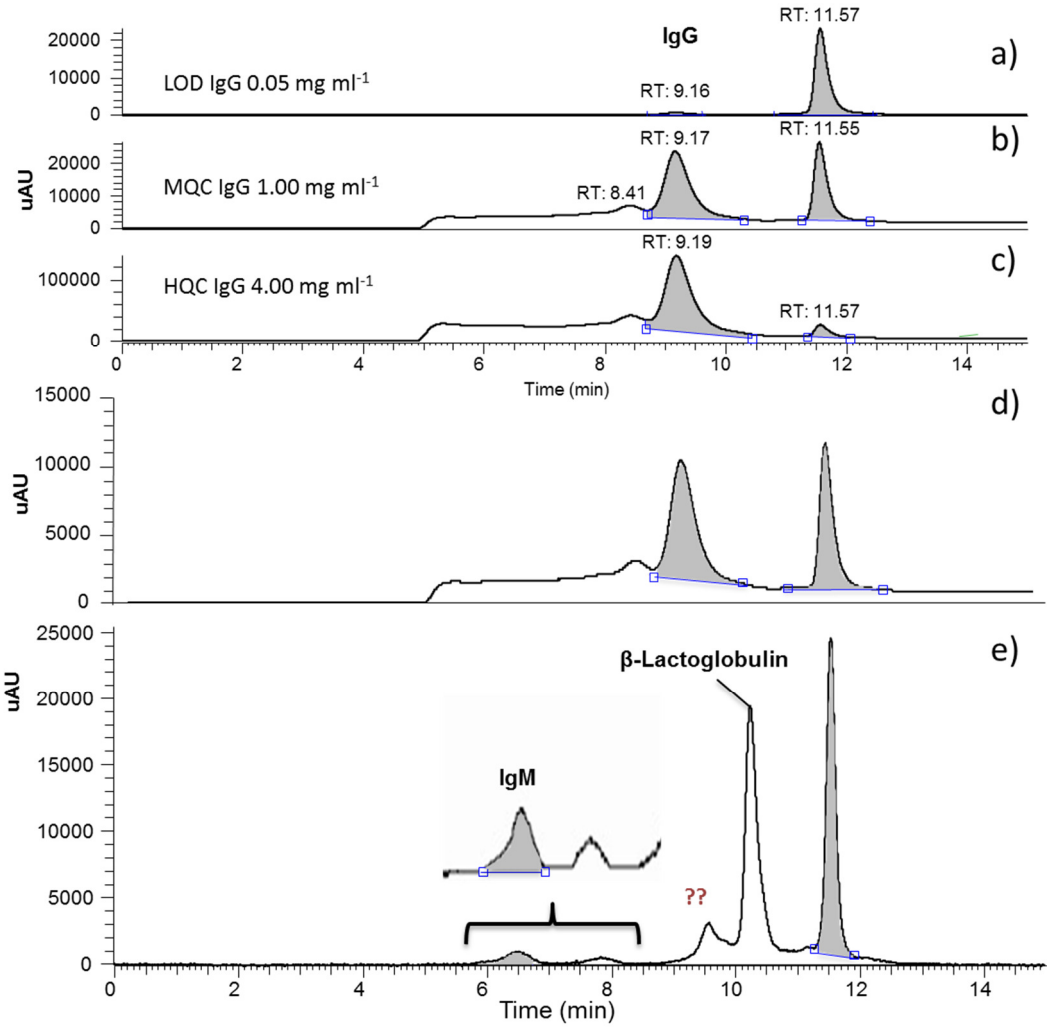


Figure 6 – IgG and IgM analysis by SEC-UV; Panels a), b), c) show the SEC-UV chromatogram of purified IgG at different concentrations (LOD 0.05 mg mL-1, MQC 1.00 mg mL -1, HQC 4.00 mg mL-1), using TH as internal standard; panel d) shows the SEC-UV chromatogram of IgG purified fraction from BC4 sample; panel e) shows the SEC-UV chromatogram of IgG depleted fraction of BC4; IgM peak elutes at 6.12±0.01 min.

IgG eluted at 9.15±0.01min, flanked by a less intense peak eluting at a lower retention time (8.4 min) and attributed to IgG aggregates (the 1D gel electrophoresis pattern in reducing condition of the collected peak shows the typical IgG subunits at 50 and 25

kDa). TH eluted at 11.57±0.02min. The calibration curve relative to the IgG monomer eluting at 9.15 min was built ($y = 2.1129x + 0.0012$, $R^2 = 0.9998$) using a concentration range of 0.05-4.00 mg mL^{-1} (Figure 7).

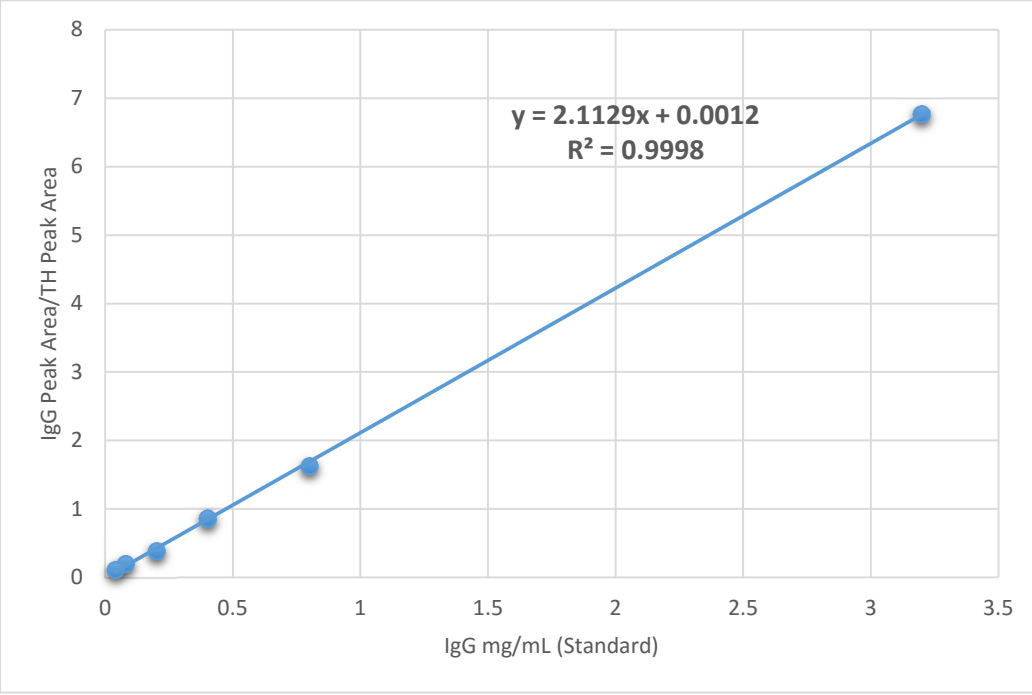


Figure 7 - IgG Calibration Curve

The LLOQ was evaluated by analyzing six replicates at the concentration of 0.1 mg mL^{-1}. The LOD was 0.05 mg mL^{-1}. As reported in Table 3, intra- and inter-day precisions were determined by analyzing six replicates of quality control samples (LQC= 0.1 mg/mL, MQC= 1.0 mg/mL and HQC = 4.0 mg/mL) on the same day and on three different days.

	IgG			IgM		
	LQC (LLOQ)	MQC	HQC	LQC (LLOQ)	MQC	HQC
Theoretical concentration [mg/mL]	0.100	1.000	8.000	0.125	0.375	0.900
INTRA-DAY						
Mean Concentration (n=6)	0.105	1.011	3.979	0.115	0.368	0.870
Precision - CV %	4.9	1.4	1.6	7.8	5.2	6.2
Accuracy - BIAS %	5.0	1.1	-0.5	-8.0	-1.9	-3.3
INTER-DAY						
Mean Concentration (n=6)	0.106	1.014	3.975	0.144	0.343	0.842
Precision - CV %	3.9	1.1	1.5	7.8	6.5	6.1
Accuracy - BIAS %	6.0	1.4	-0.6	15.2	-8.5	-6.4

Table 3 - IgG and IgM quantitative analysis by SEC-UV: precision and accuracy.

Satisfactory precision, accuracy and reproducibility were obtained: the intra- and inter-day precisions (CV%) were <5.0% and the accuracies (BIAS%) were <10%. No significant carry-over from the analyte or IS was observed in the double blank after subsequent injection of the highest calibration standard at the retention time of analyte.

The method was then applied to a commercially available decaseinated and defatted BC sample (sample BC4). The typical SEC-UV chromatogram (Figure 6d) displays the peak referred to TH at 11.56±0.02 min and the peak of IgG at 9.15±0.01min, flanked, as in the case of standard samples, by a minor peak at lower retention time and attributed to IgG aggregates. IgG content for the analyzed batch samples accounted for 77.6 mg/100 mg (mean content) of BC raw material (Table 4).

Sample	IgG mg IgG / 100mg BC	IgM mg IgM / 100mg BC	α-Lactalbumin mg α-Lac / 100mg BC	β-Lactoglobulin mg β-Lac / 100mg BC
IgG Eluted BC4 (batch A)	78.409	-	-	-
IgG Eluted BC4 (batch B)	76.869	-	-	-
IgG-depleted-BC4 (batch A)	-	1.460	0.248	1.483
IgG-depleted-BC4 (batch B)	-	1.910	0.358	1.865
IgG-depleted-BC4 (batch C)	-	2.000	0.252	1.681
AVERAGE ± DEV.ST.	**77.64 ± 1.08**	**1.79 ± 0.29**	**0.28 ± 0.06**	**1.67 ± 0.19**

Table 4 - IgG, IgM, α-lactalbumin and β-lactoglobulin contents measured in four different batches of BC4 (casein-depleted and defatted bovine colostrum).

3.3.2 Immunoglobulin M analysis

The IgM quantification by HPLC-SEC-UV was validated as reported for IgG (concentration levels: LQC 0.125, MQC 0.375, HQC 0.900 mg mL^{-1}) and using TH as IS (0.1 mg mL^{-1}). Typical retention time for IgM was about 6.1 min, and a typical calibration curve is shown in Figure 8 ($y = 2.9997x + 0.0057$, $R^2 = 0.9997$).

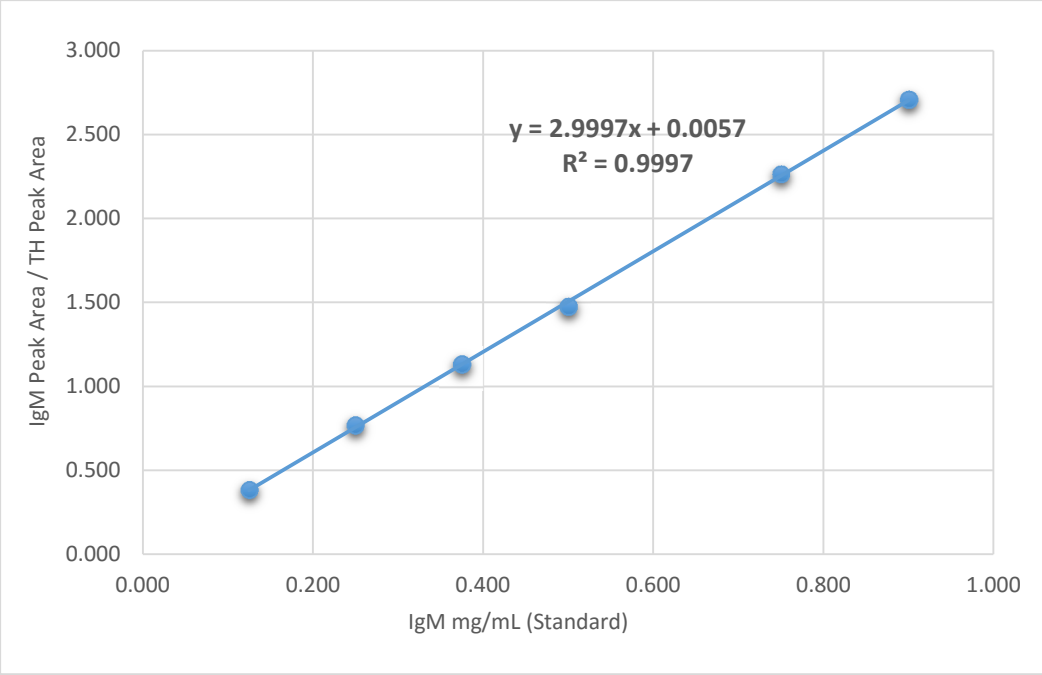


Figure 8 - IgM Calibration Curve

The LLOQ was evaluated by analyzing six replicates at a concentration of 0.125 mg mL^{-1}. The intra- and inter-day precisions (CV%) were <5.0% and the accuracies (BIAS%) were <20% (Table 3), confirming the repeatability and reproducibility of the method. The method was then applied to the analysis of IgM in a commercially available BC raw material (sample BC4). The corresponding SEC-UV chromatogram (Figure 6e) is characterized by a main peak at 11 min attributed to β-lactoglobulin as determined by comparing the RT of a genuine standard flanked by an unknown peak. The IgM peak eluted at 6.12±0.01 min. As shown in Table 4 the IgM content measured in different batches was of 1.7 mg corresponding to 2.5 % of the total immunoglobulin content.

3.4 β-lactoglobulin and α-lactalbumin analysis by LC-ESI/MS

As revealed by the protein MS identification on the 2D-gel analysis (Table 2), besides IgG, β-lactoglobulin (18KDa) and α-lactalbumin (14KDa) represent the most abundant whey proteins. Both proteins were analyzed by LC-ESI-MS and using ubiquitin (8.5KDa) as internal std. Suitable HPLC conditions were achieved by using a mixture of CH_3CN/H_2O (70/30 v/v) spiked with 0.01% HCOOH to improve efficiency and peak symmetry. Figure 9 reports the total ion current (TIC) of the standards mixture containing β-lactoglobulin A and B, and α-lactalbumin and ubiquitin as IS, revealing 3 peaks.

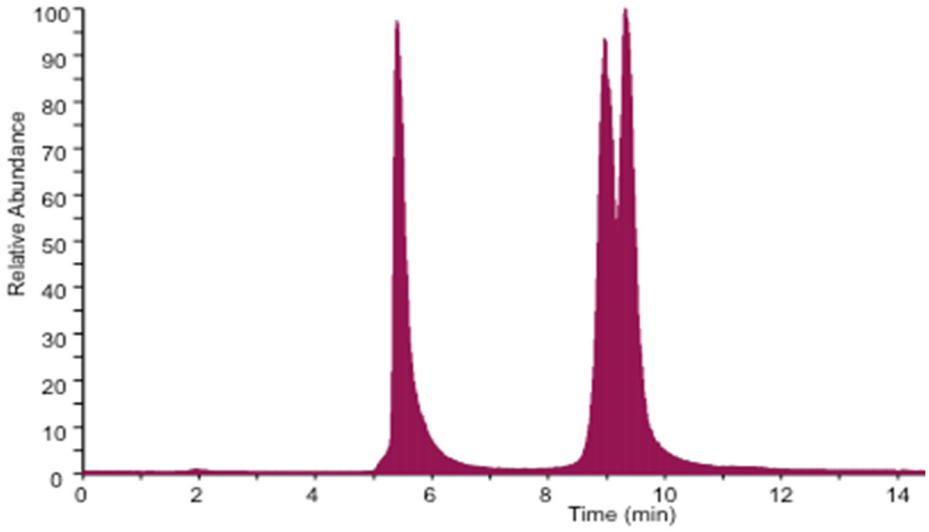


Figure 9 - β-lactoglobulin and α-lactalbumin analysis by LC-ESI/MS. Total ion current (TIC) of the standards mixture containing β-lactoglobulin A and B, α-lactalbumin and ubiquitin as IS.

The multi-charged ions and the corresponding deconvoluted spectra for each peak revealed that the first peak at 5.38±0.01min corresponds to ubiquitin (Figure 10a); the second one (RT 9.08±0.06 min), corresponds to the co-elution of the isoform B of β-Lactoglobulin and α-Lactalbumin (Figure 10b); the last one (RT 9.43±.05 min) is attributed to the isoform A of β-Lactoglobulin (Figure 10c).

To selectively monitor the elution of the three analytes, selected ion chromatograms (SIC) were reconstituted by monitoring for each analyte the characteristic 3 most intense ions and in particular *m*/*z* 714.45, 779.36, 857.11 for ubiquitin, *m*/*z* 1223.77, 1310.79, 1412.09 for the isoform A of β-lactoglobuiln, *m*/*z* 1218.06, 1305.09, 1404.95 for the isoform B of β-lactoglobuiln and *m*/*z* 1418.51, 1575.44, 1772.32 for α-lactoalbumin.

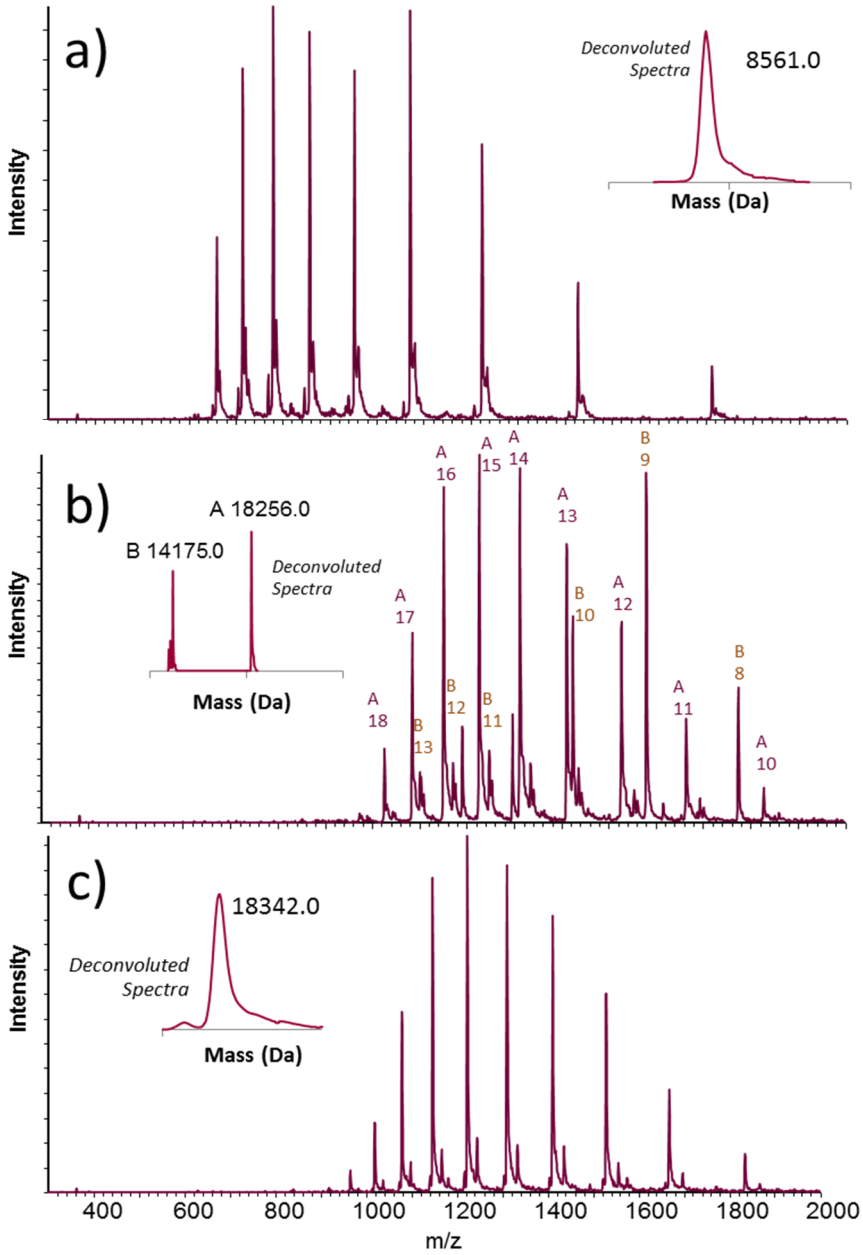


Figure 10 - β-lactoglobulin and α-lactalbumin analysis by LC-ESI/MS. Panel a) MS spectrum of the peak at 5.38±0.01min corresponding to ubiquitin; panel b) MS spectrum of the peak at 9.08±0.06 min, corresponding to the co-elution of the isoform B of β-Lactoglobulin and α-Lactalbumin; panel c) MS spectrum of the peak at RT 9.43±.05 min attributed to the isoform A of β-Lactoglobulin. The inlets of panels b, c and d show the deconvoluted spectra.

Calibration curves for β-lactoglobulin A/B and α-lactalbumin Figure 11 (panel a-b-c) ranged from 0.5 to 20.0 µM with R2= 0.9998 ±0.0001.

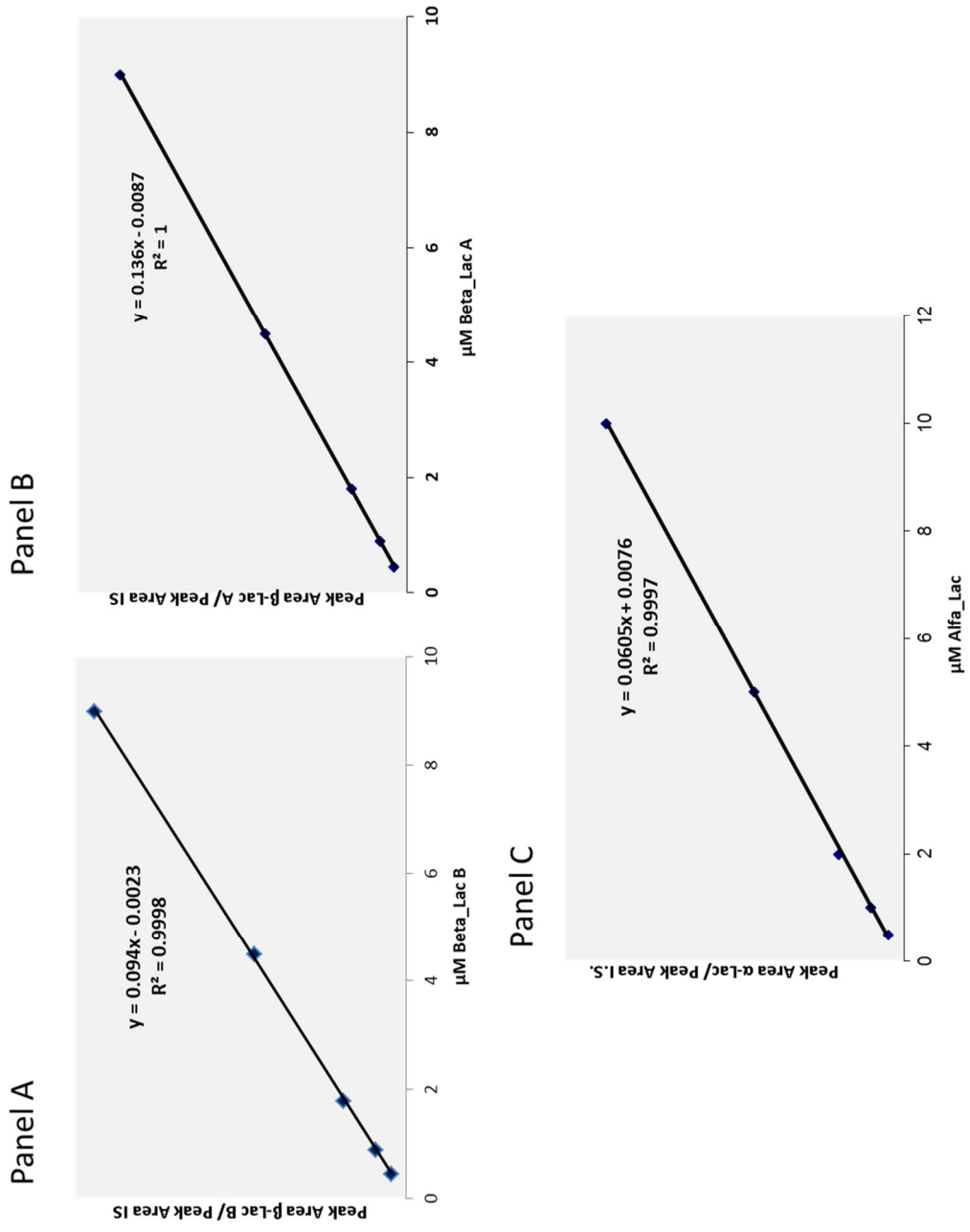


Figure 11 - Panel a) β-lactoglobulin B calibration curve; panel b) β-lactoglobulin A calibration curve; panel c) α-lactalbumin calibration curve.

Figure 12a shows the TIC relative to BC4 sample spiked with the IS. SIC traces relative to ubiquitin, α-Lactalbumin, β-Lactoglobulin A and B are displayed in panels b, c, d, e, respectively.

Intra- and inter-day precisions were determined by an assay of quality control samples at low, medium and high concentrations (six replicates each) on the same day and on three different days; precision and accuracy were within ±20% for LLOQ (LQC) and ±15% for all the other concentrations, indicating satisfactory accuracy and precision (Table 5).

	β-Lactoglobulin			α-Lactalbumin		
Molarity Level (µM)	**LQC (LLOQ)**	**MQC**	**HQC**	**LQC (LLOQ)**	**MQC**	**HQC**
Theoretical Concentration [µm]	0.500	2.000	20.000	0.500	2.000	20.000
INTRA-DAY						
Mean Molarity (n=6)	0.499	2.001	16.580	0.462	2.026	19.920
Precision - CV %	6.7	1.7	2.5	4.7	3.7	2.1
Accuracy - BIAS %	-0.2	0.05	-17.1	-7.6	1.3	-0.4
INTER-DAY						
Mean Molarity (n=6)	0.518	2.020	16.513	0.425	2.101	19.410
Precision - CV %	1.1	1.5	2.0	9.4	7.2	8.5
Accuracy - BIAS %	3.6	1.0	-17.4	-15.0	5.0	-2.9

Table 5 - β-lactoglobulin and α-lactalbumin quantitative analysis by HPLC-ESI-MS: precision and accuracy.

The method was then applied to the IgG depleted fraction of BC sample. Analytes were easily identified and quantitatively determined. Results are summarized in Table 4.

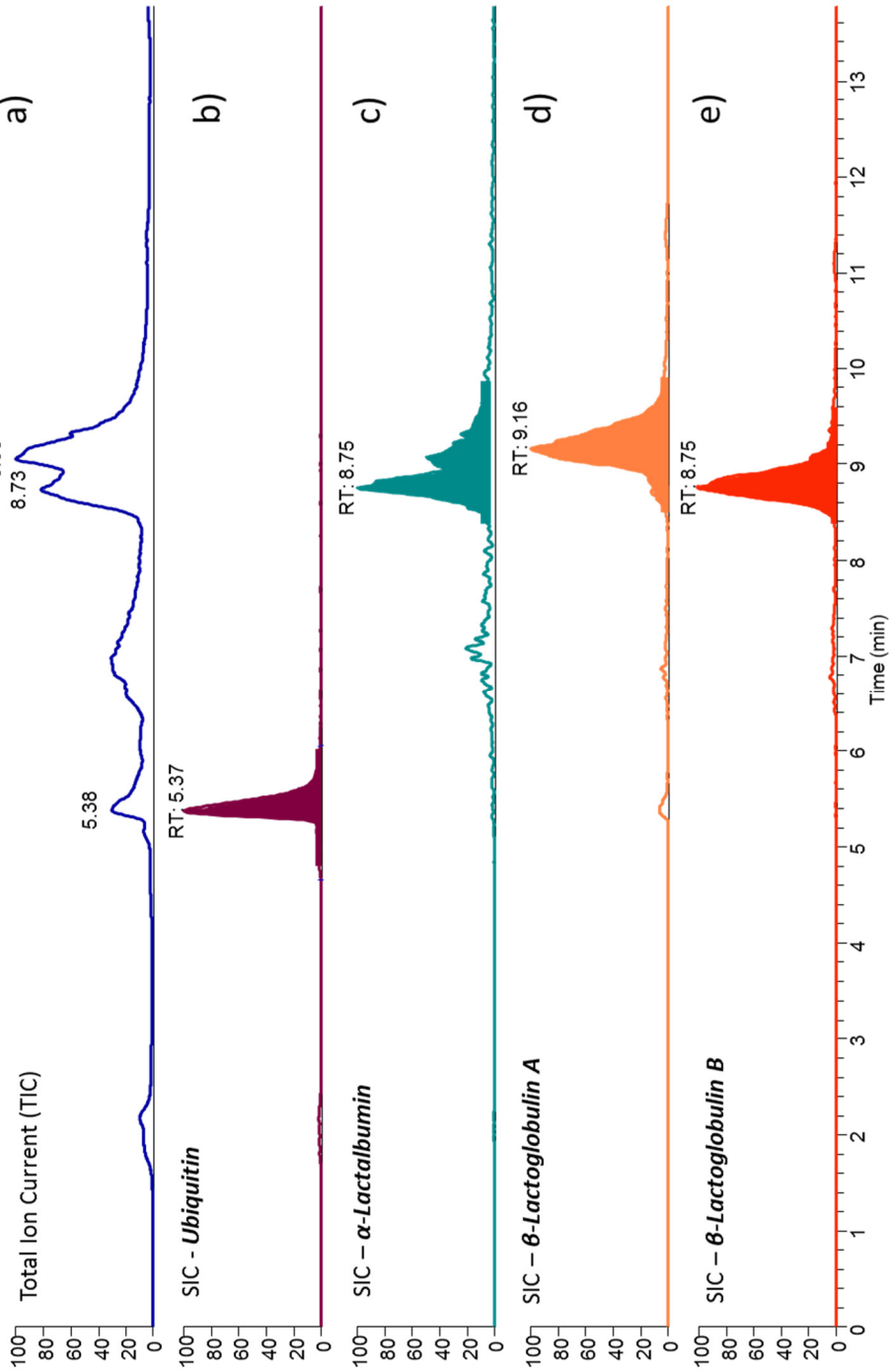


Figure 12 – β-lactoglobulin and α-lactalbumin analysis in BC4. panel a) TIC relative to BC4 sample spiked with the IS. SIC traces relative to ubiquitin, α-Lactalbumin, β-Lactoglobulin A and B are displayed in panels B, C, D, Eb, c, d, e, respectively.

3.5 Targeting bioactive proteins by ELISA assay

ELISA analyses were performed to measure the content of three functional proteins involved in wound healing as previously identified by proteomic analysis of BC conducted by our team: lumican (37KDa), IL-17 (14 kDa) and IGF-1 (7.5 kDa) [5]. Quantitative measurements of BC4 were carried out according to the suppliers' manuals. IL-17, Lumican and IGF-1 content was 2.25, 14.10 and 1.54 pg mg^{-1} calculated on the dry matter, respectively.

3.6. Protein Functionality

3.6.1 Lactoperoxidase Activity Assay - Spectrophotometric method

The LPO activity was measured by setting a calibration curve built by plotting the enzymatic activity, reported as the increase of absorbance at 413 nm min^{-1}, in respect to the enzyme standard concentrations (four levels) (Figure 13).

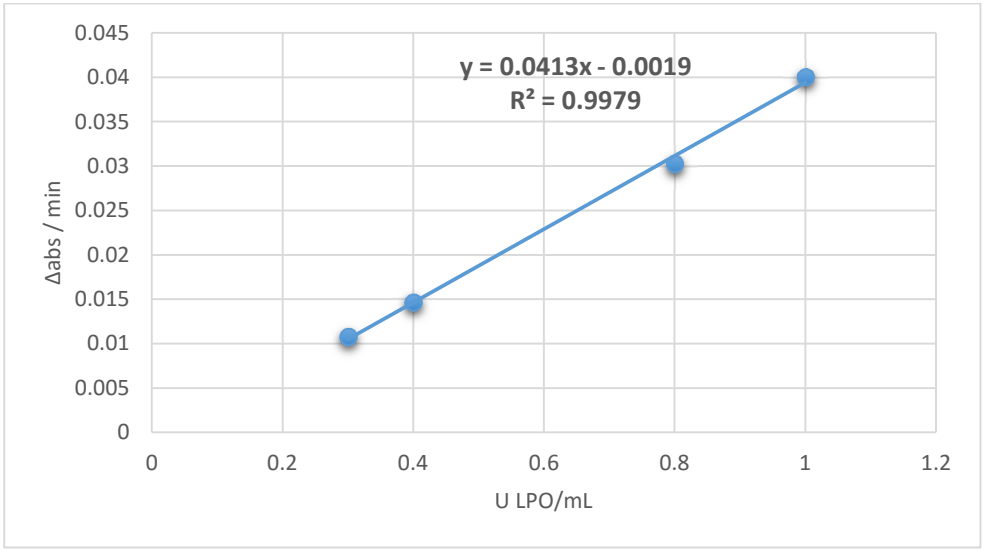
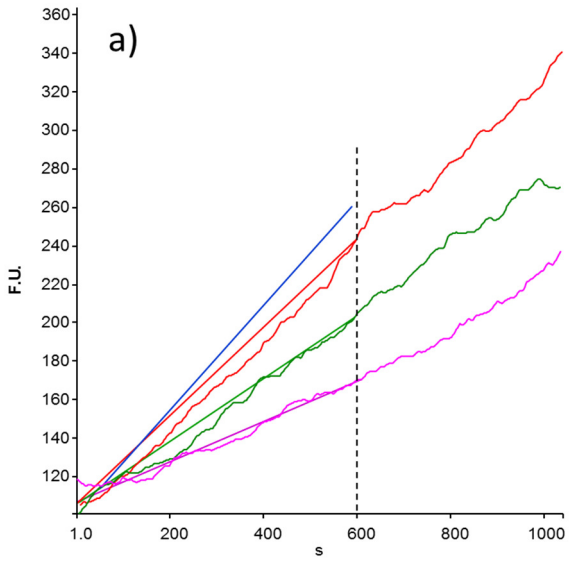


Figure 13 - LPO Activity calibration curve: the enzymatic activity is reported as the increase of absorbance at 413 nm per minute, in respect to the enzyme standard concentrations (four levels).

LPO activity was measured on BC, decaseinated/defatted BC and purified BC. Enzymatic activity in genuine BC accounted for 0.122±0.01 U mL^{-1}, which is much lower than in bovine milk (1.2 to 19.4 U mL^{-1}). However, considering that the minimum amount of LPO in milk required to activate the LPO system is 0.02U mL^{-1}, the activity found in untreated BC is enough to be effective. No LPO activity was measured in processed BC up to a concentration of 100 mg ml^{-1}.

3.6.2 Determination of Xanthine Dehydrogenase/Oxidase Activity - Fluorimetric Assay

The enzymatic activity of Xanthine Dehydrogenase/Oxidase (XDH/XO) was initially measured in BC0. An aliquot of 10 µl of colostrum (corresponding to 1.428 mg protein) was able to generate a well detectable increase of fluorescence, whose slope was automatically calculated by using the available arithmetic tool within the FL WinLab™ software (Perkin Elmer). The enzymatic activity was determined by using a quantitative method based on adding the standard of XDH/XO to BC0 at three different concentration levels (0.0005, 0.0010, and 0.0020 U mL^{-1}). Figure 14 shows the fluorescent time course and the calculated slopes of BCO (panel a) spiked in the absence and presence of the three concentrations of standard enzymes.



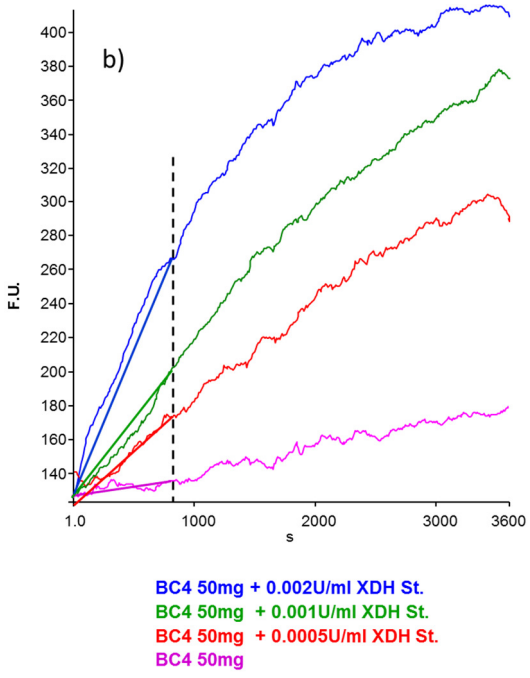


Figure 14 - Determination of Xanthine Dehydrogenase/Oxidase Activity by fluorescence - fluorescent time courses and calculated slopes of BCO (panel a) and BC4 (panel b) spiked in the absence and presence of the three concentrations of standard enzymes (0.002 U ml-1 blue line; 0.001 U ml-1 red; 0.0005 U ml-1 green). The XDH/XO activity in the tested samples was determined by setting y=0 in the y=mx+b equation determined by considering the linear calibration curve obtained by considering the slopes of the three concentration levels.

The XDH/XO activity in BC0 was determined by setting y=0 in the y=mx+b equation determined by considering the linear calibration curve obtained by considering the slopes of the three concentration levels (Figure 14b). A similar approach was then applied to BC samples. The enzymatic activity was determined in only two samples (BC4 and BC5) when dissolved at a final concentration of 50 mg (30 mg of protein mL^{-1}) and characterized by a significant difference in terms of slope. The other samples did not provide a detectable slope up to a concentration of 1.5 mg ml^{-1}. Higher concentrations were not tested due to the formation of a cloudy suspension due to the presence of caseins. For each sample the amount of the enzyme was determined by ELISA thus permitting the reporting of enzymatic activity in respect to the amount of the enzyme. As shown in Table 6 the ratio reduced between 10 and 100 times in processed colostrum in respect to frozen colostrum.

SAMPLE	mg protein tested (Bradford)	XDH Enzymatic activity (U/ml)	U XDH / mg protein	ng XDH / µg proteins (ELISA)	U XDH / mg XDH
Blank	0.000	0.000	-	-	-
BC0 (10µl)	1.428	1.41E-03	9.90E-04	1.169	**0.840**
BC4 (50mg)	30.501	1.07E-04	3.51E-06	1.317	**0.003**
BC5 (50mg)	31.131	2.20E-03	7.07E-05	1.174	**0.060**

Table 6 - XDH enzymatic activity and protein content measured in BC0, BC4 and BC5.

Discussion

Bovine colostrum components in addition to their well-established nutritional properties also exhibit specific biological activities that make bovine colostrum, and its derivatives, a valuable ingredient for several food supplements and health products [4, 16]. Most of the beneficial effects of BC have been ascribed to the proteins [7]. By using an in depth proteomic and bioinformatic approach, we recently identified more than 1200 proteins and among these we selected protein clusters that on the basis of their biological activity can explain some claimed health benefits. For instance, BC is used as a functional ingredient in products claimed to better help the wound healing process and we identified 93 proteins involved in the different steps of the wound healing process [5].

On the one hand, BC is an amazing natural resource of bioactive protein components, on the other it is a very complex and oxidizable matrix that requires high standardization and in depth quality profiling for its usage as ingredient in food supplement products, to guarantee safety and a reproducible efficacy. We here report a suitable analytical platform for the quality control of the protein components of raw materials.

Analysis starts with a 1D and 2D gel electrophoresis in both reducing and non-reducing conditions. This is quite an easy analysis that permits the establishment of the presence of high molecular aggregates as well of caseins that in some commercial products are intentionally removed. Densitometry analysis can also furnish a draft semi-quantitative analysis of the most abundant proteins including IgG, IgM, lactoalbumins and lactoglobulins.

The immunoglobulins fraction represents a very important antimicrobial component of BC. Ig antibodies express multifunctional activities, including complement

activation, bacterial opsonisation and agglutination, and act by binding to specific sites on the surfaces of the most infectious agents or products, either inactivating them or reducing infection [6]. In bovine colostrum, immunoglobulin G (IgG; subclasses IgG1 and IgG2) is the major immune component, and is accompanied by low levels of IgA and IgM [17].

As critically evaluated by Rathe et al. [4] evidence suggests that bovine colostrum has immunity-modulating capabilities that may be utilized in maintaining or improving host defences under different detrimental conditions or immune system exposures. Such effects are due to the Ig fraction and hence an accurate qualitative analysis is needed for this class of bioactive components. Several analytical methods have so far been reported for the quantitative analysis of Ig, including radial immunodiffusion, reversed-phase liquid chromatography, ion exchange chromatography, capillary electrophoresis, SEC, and ELISA, the latter representing the most widely used method [18]. ELISA kits have different positive aspects such as ease of use, sensitivity, speed of analysis and they do not require complex instruments or skilled operators. However, ELISA kits also have some drawbacks such as high cost and limited accuracy.

We here propose a validated SEC-UV method that measures both IgG and IgM. The method was found to be accurate and precise with a LOD/LLOQ of 0.05 mg mL^{-1}/0.10 mg mL^{-1}, and 0.125 mg mL^{-1}/0.250mg mL^{-1}, for IgG and IgM respectively. Analyses were carried out on two fractions prepared by an affinity chromatography based on protein G as stationary phase. IgG are analysed in the IgG enriched fraction while the IgM are analysed in the IgG depleted fraction. The SEC-UV method was found not to be sensitive enough for the measurement of IgA that can be measured by ELISA test.

The IgG depleted fraction is then analysed by HPLC-MS for measuring other major protein components. The identified intact proteins clearly depend on the type of MS analyser. The ion trap was limited because it was unable to detect proteins with a MW greater than 30 kDa which by contrast were easily detected by the TOF analyser. Besides the two isoforms of β-lactoglobulin A/B and the α-lactoalbumin easily identified by LC-ESI ion trap, LC-ESI-q-TOF system also identified intact lactotransferrin, cytocrome C, β2-microglobulin, bovine serum albumin, lactoperoxidase. Quantitative analysis of β-lactoglobulins A/B and the α-lactoalbumin were then carried out applying a validated HPLC-MS method by using the Ion Trap as mass analyser. Notably the number of proteins quantified by the intact protein assay can be significantly increased by using a TOF MS analyser. The intact protein analysis is also suitable to measure isoforms and post-translational modifications such as lactosylation.

The strength of the proposed method is the simultaneous analysis of both the isoforms of β-lactoglobulins and of α-lactoalbumin in colostrum samples. We used the IgG-depleted fractions for the α-lactalbumin and β-lactoglobulin content measurement, since immunoglobulins have been found to interfere with the MS analysis and reduce the chromatographic column performance. Although we set-up a method that enables the simultaneous quantification of the two isoforms of β-lactoglobulin, we here reported the values measured for the total content of the protein, since the standard of β-lactoglobulin we used is a not-defined mixture of both the isoforms. However, by assuming that the two isoforms have a similar ionic response a semi-quantitative measurement can be considered (in BC4 the isoforms A and B accounted for the 61.3 and 38.7 %, respectively).

Besides immunoglobulins and the other most abundant proteins, the quality control of BC should also target the bioactive proteins responsible for the claimed BC health properties. Selection of such bioactive proteins can be carried out by selecting them from the database of BC proteins recently identified by us using a proteomic and bioinformatics approach [5]. Sensitive and selective analytical methods are then used for their quantitative measurement, such as ELISA. As an example, three proteins involved in the re-epithelisation and wound healing process were selected: Lumican, IL-17 e IGF-1 and their quantitation was carried out by ELISA assay.

Biological potency is a very important parameter for the quality control of biological products, such as recombinant proteins. Such a test is well integrated with the chemical analysis which furnishes detailed data regarding the structural integrity of the protein (primary sequence, MW, disulphide bonds and some PTM) but it is limited when assessing the correct biological activity of the analyte. The biological activity would be of great value also for testing the efficacy of BC because the chemical analysis cannot clearly cover all the bioactive constituents and their intrinsic activity. Moreover, the conformational integrity of bioactive proteins is a parameter that should be carefully evaluated due to the fact that BC is subjected to several purification steps including the lyophilization that can affect conformational integrity and consequently protein function. However, *in vitro* tests to evaluate the biological activity of BC are far from being devised particularly because the biological activity of BC results from the combined action of several constituents and hence it cannot be easily surrogated to an *in vitro* test. Based on these facts we propose to measure the enzymatic activity of some representative enzymes, which are contained in colostrum, and for this purpose, we considered Lactoperoxidase and Xanthine Dehydrogenase/Oxidase. The enzymatic activity can be considered as an index of protein functional integrity;

in other words if bioactive proteins are denaturated, even partially, during any step of the process (colostrum conservation, transfer, processing, lyophilisation) then also the target enzymes are affected and this can be determined by measuring their activity. Lactoperoxidase was well detected in the raw material but not in any of the commercial products tested. By contrast, XDH/XO activity, besides being significantly measured in the raw material, was also determined in some commercial products and its activity was used as a marker of protein functional integrity. Since the overall XDH/XO activity is due to both the intrinsic enzyme activity and to the absolute amount and since this latter clearly depends on several uncontrolled variables, the amount of XDH/XO was determined by ELISA test, thus expressing the XDH/XO activity in respect to the absolute amount of the enzyme. XDH/XO activity in colostrum was well detectable reaching 0.8 U mg of protein. In processed colostrum the activity was found to be very variable since in a set of samples it was not detected and in another samples it was found reduced between 10 and 100 fold. This loss of activity clearly depends on several parameters including the process technology and such results clearly indicate that XDH/XO activity could be a useful parameter to set-up each step of the process and also to monitor the shelf-life of the powder.

Conclusions

In summary a comprehensive method for the quality control of the protein fraction of BC has been proposed and then applied to some commercial BC derivatives. The method permits, after a suitable fractionation by affinity chromatography, the quantitative determination of the Ig components as well as of the most abundant protein components including β-lactoglobulin and α-lactoalbumin. Functional proteins selected from the database on the basis of the health claim made are then

determined by ELISA. Finally, protein functional integrity is assayed by measuring the XDH/XO activity. In conclusion the here proposed method is suitable for the quality control of protein fractions of BC which are needed to guarantee safety and biological activity of such a complex matrix.

References

1. Godden, S., *Colostrum management for dairy calves.* Vet Clin North Am Food Anim Pract, 2008. **24**(1): p. 19-39.
2. Weaver, D.M., et al., *Passive transfer of colostral immunoglobulins in calves.* J Vet Intern Med, 2000. **14**(6): p. 569-77.
3. Senda, A., et al., *Changes in the bovine whey proteome during the early lactation period.* Anim Sci J, 2011. **82**(5): p. 698-706.
4. Rathe, M., et al., *Clinical applications of bovine colostrum therapy: a systematic review.* Nutr Rev, 2014. **72**(4): p. 237-54.
5. Altomare, A., et al., *Health benefits of bovine colostrum revealed by an in depth proteomic analysis*. 2015.
6. Korhonen, H., P. Marnila, and H.S. Gill, *Bovine milk antibodies for health.* Br J Nutr, 2000. **84 Suppl 1**: p. S135-46.
7. Sacerdote, P., et al., *Biological components in a standardized derivative of bovine colostrum.* J Dairy Sci, 2013. **96**(3): p. 1745-54.
8. Inagaki, M., et al., *Multiple-dose therapy with bovine colostrum confers significant protection against diarrhea in a mouse model of human rotavirus-induced gastrointestinal disease.* J Dairy Sci, 2013. **96**(2): p. 806-14.
9. Purup, S., et al., *Biological activity of bovine milk on proliferation of human intestinal cells.* J Dairy Res, 2007. **74**(1): p. 58-65.
10. Kelly, G.S., *Bovine colostrums: a review of clinical uses.* Altern Med Rev, 2003. **8**(4): p. 378-94.
11. Bölke, E., et al., *Preoperative oral application of immunoglobulin-enriched colostrum milk and mediator response during abdominal surgery.* Shock, 2002. **17**(1): p. 9-12.
12. Permyakov, E.A. and L.J. Berliner, *alpha-Lactalbumin: structure and function.* FEBS Lett, 2000. **473**(3): p. 269-74.
13. Dupont, C., et al., *Alpha-lactalbumin-enriched and probiotic-supplemented infant formula in infants with colic: growth and gastrointestinal tolerance.* Eur J Clin Nutr, 2010. **64**(7): p. 765-7.
14. Wernimont, S., et al., *Effect of an α-lactalbumin-enriched infant formula supplemented with oligofructose on fecal microbiota, stool characteristics, and hydration status: a randomized, double-blind, controlled trial.* Clin Pediatr (Phila), 2015. **54**(4): p. 359-70.
15. Shindler, J.S., R.E. Childs, and W.G. Bardsley, *Peroxidase from human cervical mucus. The isolation and characterisation.* Eur J Biochem, 1976. **65**(2): p. 325-31.
16. Bagwe, S., et al., *Bovine colostrum: an emerging nutraceutical.* J Complement Integr Med, 2015.
17. Verweij, J.J., A.P. Koets, and S.W. Eisenberg, *Effect of continuous milking on immunoglobulin concentrations in bovine colostrum.* Vet Immunol Immunopathol, 2014. **160**(3-4): p. 225-9.
18. Gapper, L.W., et al., *Analysis of bovine immunoglobulin G in milk, colostrum and dietary supplements: a review.* Anal Bioanal Chem, 2007. **389**(1): p. 93-109.

Part III

The secrets of Oriental panacea: Panax ginseng

Abstract

The Panax ginseng root proteome has been investigated via capture with combinatorial peptide ligand libraries (CPLL) at three different pH values. Proteomic characterization by SDS-PAGE and nLC-MS/MS analysis, via LTQ-Orbitrap XL, led to the identification of a total of 207 expressed proteins. This quite large number of identifications was achieved by consulting two different plant databases: Panax ginseng and Arabidopsis thaliana. The major groups of identified proteins were associated to structural species (19.2%), oxidoreductase (19,5%), dehydrogenases (7.6%) and synthases (9.0%). For the first time, an exploration of protein-protein interactions was performed by merging all recognized proteins and building an interactomic map, characterized by 196 nodes and 1554 interactions. Finally a peptidomic analysis was developed combining different in-silico enzymatic digestions to simulate the human gastrointestinal process: from 661 generated peptides, 95 were identified as possible bioactives and in particular 6 of them were characterized by antimicrobial activity. The present report offers new insight for future invesigations focused on elucidation of biological properties of Panax ginseng proteome and peptidome.

1. Introduction

Asian ginseng (*Panax ginseng C. A. Meyer)* has a history of herbal use going back over 5,000 years as described in Chinese traditional medicine textbooks. It is one of the most highly regarded of herbal medicines in the Orient, where it has gained a reputation for being able to promote health, general body vigour and also to prolong life [1-3]. The genus name *Panax* is derived from the Greek word meaning "panacea" or "all-healing": the species ginseng is said to mean "wonder of the world". Both terms refer to the medicinal virtues of the plant belonging to *Araliaceae* family. Ginseng is traditionally used as an aid during convalescence and as a prophylatictic to build resistance, to reduce susceptibility to illness and to promote health and longevity [4]. Panax ginseng is a slow-maturing perennial herb native to the mountain forests of northeastern China, Korea and Russia. Seven major species of ginseng are cultivated extensively in China, Japan, Korea, Russia, Canada and Wisconsis in the US. The three mostly studied species are: *Panax ginseng* (Asian ginseng), *Panax quinquefolius* (American ginseng) and *Panax japonicas* (Japanese ginseng). Ginseng usually starts flowering at its fourth year and the roots take four to six years to reach maturity. Therefore, productivity of *Panax ginseng* can be significantly affected by various environmental factors such as temperature, condition of soil, light intensity, content of water and diseases [5-10]. Ginseng root consists of dried main root, lateral roots and root hairs or “tailings” of *Panax ginseng* C. A. Meyer, used as a tonic to revitalize and replenish vital energy. The bioactivities of ginseng roots, which include hypotensive [11], anticancer [12-14], antioxidant [15], anti-inflammatory activities as well as improving impaired memory [16-18] and anti-cardiovascular diseases [19, 20], are due to the presence of main pharmacologically active

components like ginsenosides. To understand these effects, the components of ginseng have been subjected to extensive analysis. Its saponins, known as ginsenosides, are considered to be the main active pharmacological compunds in P. ginseng. The distribution of ginsenosides varies from species to species [2] and each ginsenoside has different pharmacological effects even if it could produce multiple effects in the same tissue [3].

Although many reports have been published regarding the pharmacological effects of ginsenosides, little is known about biochemical pathways involving different proteins. High-throughput and high-sensitive proteomic techniques, such as gel electrophoresis and mass spectrometry, allow the separation of complex protein mixtures, which makes possible to charcterize protein profiles in plant roots. Recently, proteomic studies have been performed on ginseng roots: the first proteomic profiles were provided by Lum et al. [21], subsequently improved by Kim et al. by identifying several high-abundance proteins [22]. Most of highly abundant proteins in ginseng root are root-specific RNase-like proteins that function as vegetative storage proteins for survival in the natural environment [23, 24]. Subsequently, proteomic analysis was used on different varieties of ginseng, such as *Panax ginseng* and *Panax quinquefolium* [25-27]. However, to date, no systemic research has been reported for a deep investigation focused to identify the entire ginseng root proteome.

Therefore, in this study, a proteomic approach was applied to identify proteins of *Panax ginseng* C. A. Meyer root, contributing to understand their role in metabolic processes. Previously, a method for sample preparation of dried roots was established [28] and a suitable technology, called combinatorial peptide ligand

libraries (CPLLs) was used to capture the entire proteome by enriching trace proteins and concomitantly reducing the concentration of abundant species [29]. Ginseng proteins were identified by nLC-MS/MS, using an Orbitrap mass spectrometer, in order to extensively map the proteome for a consequent exploration of protein functions via Gene Ontology analysis. Moreover, also an interactomic investigation was performed merging all the proteins, by exploiting the STRING v.9.1 software, which was integrated with the analysis of peptides, obtained after an in-silico human gastrointestinal digestion.

These results provide insight into the proteome and peptidome of *Panax ginseng* roots and contribute to a more comprehensive understanding of biological functions.

2. Materials and methods

2.1 Chemicals

ProteoMiner™ (combinatorial hexapeptide ligand library beads, CPLL), Laemmli buffer, 40% acrylamide/Bis solution, N,N,N',N'-tetramethylethylenediamine (TEMED), molecular mass standards and electrophoresis apparatus for one-dimensional electrophoresis were from Bio-Rad Laboratories, Inc., Hercules CA. β-mercaptoethanol, dithiothreitol (DTT), ammonium persulfate, 3-[3-cholamidopropyl dimethylammonio]-1-propanosulfonate (CHAPS), acetonitrile (ACN), trifluoroacetic acid (TFA), sodium dodecyl sulphate (SDS), iodoacetamide (IAA), formic acid (FA) and all other chemicals used all along the experimental work were current pure analytical grade products and purchased from Sigma-Aldrich S.r.l, Italy. Water with 0.1% formic acid and acetonitrile with 0.1% formic acid LC/MS grade were also

purchased from Sigma Aldrich S.r.l, Italy Complete protease inhibitor cocktail tablets and sequencing grade trypsin were from Roche Diagnostics (Basel, CH).

2.2 Panax ginseng root treatment

Dried powdered roots of *Panax ginseng* C. A. Meyer were used in this experiments. The dry radix was from Martin Bauer group (Vestenbergsgreuth, Germany). The drug was identified and ground by mill knives by the Phytochemistry Research Laboratory of Aboca s.p.a. (Sansepolcro, Arezzo, Italy).

Samples were stored at room temperature until required and they were prepared as described by Wang et al. [28]. Briefly *Panax ginseng* powder (3 g) was washed firstly with 20 mL of 10% TCA/ice-cold acetone, secondly with 20 mL of 80% MeOH/H_2O and finally with with 20 mL of 80% ice-cold acetone/H_2O. After drying at 25°C overnight, powder was incubated with 30 mL of extraction buffer (50 mM Tris-HCl pH=8.0, 200 mM DTT, 0.3% SDS, 1 protease inhibitor tablet for 50 mL buffer) overnight at room temperature. After centrifugation at 14000 rpm for 30 min, the supernatant was precipitated with 10% TCA solution for 1 h at -20°C. The protein pellet thus obtained was washed with ice-cold acetone to remove contaminants and resuspended into 60 mL buffer for CPLLs incubation (50 mM Tris-HCl pH=7.2, 50 mM NaCl), added with one tablet of proteases inhibitors cocktail. This solution was divided into three, twenty mL fractions: one of them was equilibrated into the same pH 7.2 buffer, the other two fractions were titrated, respectively, to pH 9.0 by addition of NaOH solution and to pH 2.2 by addition of 0.1% TFA and formic acid to mimic reverse phase conditions for the capture of hydrophobic proteins. All fractions were loaded onto 100 µL of CPLL beads at the three different pH values [30] and the

capture was performed batch-wise, gently rocking on a rotating platform for 2 hours. After that, the beads were rinsed twice with the incubation buffers, so as to remove any excess of non-adsorbed proteins (see the scheme of Figure 1).

Desorption was implemented by washing the beads twice (each time with 80 µL) with a boiling 4% SDS solution containing 20 mM DTT, 12.5 % (v/v) glycerol, 0.005% (m/v) bromophenol blue, and 62.5 mM Tris-HCl (pH 6.8) [31].

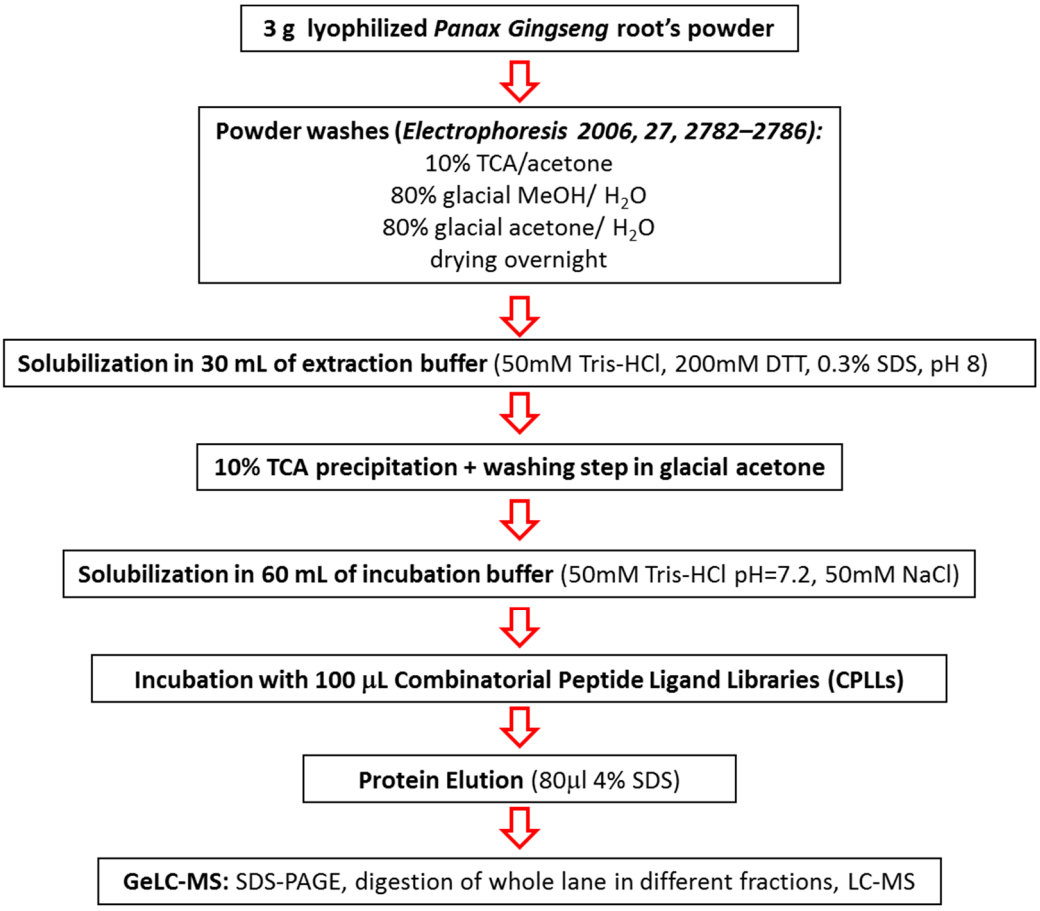


Figure 1 - Scheme of complete treatment protocol of Panax ginseng root's powder. The capture with CPLLs is performed at three pH values, namely 2.2, 7.2 and 9.0.

2.3 SDS-PAGE analysis and trypsin digestion

Ten µL of the above eluates (labelled as 2.2, 7.2 and 9.0) and ten µL of the non-treated sample (CTRL) were loaded onto an SDS-PAGE gel, composed by a 4% polyacrylamide stacking gel (125 mM Tris-HCl, pH 6.8, 0.1%, m/v, SDS) over a 12% resolving polyacrylamide gel (in 375 mM Tris-HCl, pH 8.8, 0.1%, m/v, SDS buffer). The cathodic compartment was filled with Tris-glycine buffer, pH 8.3, containing 0.1%, m/v, SDS whereas Tris buffer, at pH 8.8, was present at the anode. Electrophoresis was performed at 100 V until the dye front reached the bottom of the gel and at 150 V until the end of the separation. Staining and distaining were performed with Colloidal Coomassie Blue and 7% (v/v) acetic acid in water, respectively. For each lane of the gel, subsequent gel pieces in the range of 200-15 kDa were excised to perform in-gel digestion, accordingly to a standardized protocol [32]. Thirteen gel pieces were cut from each sample, rinsed with pure water, destained for 10 min with 50% acetonitrile/25 mM ammonium bicarbonate and incubated at 56 °C with 10 mM dithiothreitol in 50 mM ammonium bicarbonate for cysteine reduction. After 60 min incubation, the solution was discarded and the gel pieces were incubated with 55 mM iodoacetamide in 50 mM ammonium bicarbonate for 45 min, in the dark, for cysteine alkylation. After discarding the iodoacetamide solution, gel pieces were rinsed with 50 mM ammonium bicarbonate and incubated in 100% acetonitrile for dehydration. Each gel slice was incubated with 1 µg of sequencing grade trypsin (Roche) dissolved in 50 mM ammonium bicarbonate. After overnight digestion at 37°C, the solutions containing the digested peptides were collected. Gel pieces were incubated for 10 min with 30% acetonitrile, 3% trifluoroacetic acid, and the solution was collected and pooled with the initial peptide

mixture. Gel pieces were incubated for an additional 10 min with 100% acetonitrile, and the solution was pooled with the previous one. Eluted peptides were evaporated on a vacuum concentrator (Christ) and stored at -20°C.

2.4 Mass spectrometry analysis

The dried peptides were solubilized in 15 µL of 0.1% formic acid. Five µL of sample were injected on a C_{18} column (picoFrit column, C18 HALO, 90Å, 75 µm ID, 2.7 µm, 10.5 cm length, New Objective) by a constant flow rate of 0.4 µL/min delivered by a nano-chromatographic system (UltiMate 3000 RSLCnano System, Thermo Scientific). The separating gradient ramped linearly from 1% acetonitrile to 40% acetonitrile in 30 minutes; the eluting peptides were on-line sprayed in a LTQ-Orbitrap XL mass spectrometer by a nano-ESI source (all Thermo Scientific). Full scan mass spectra were acquired in the Orbitrap cell in the mass range 300 to 1500 *m/z* (positive polarity, profile mode, AGC = 5×10^5). The nine most intense ions (minimum charge state 2+, minimum intensity 10000 cps) were automatically selected and fragmented in the ion trap by collision-induced dissociation (CID). After two subsequent occurrences in less than 30 s, target ions already selected for fragmentation were dynamically excluded for 45 s.

2.5 Protein identification

The MS data were analyzed by the Proteome Discoverer software (v. 1.3.0.339 Thermo), using the Sequest algorithm. The database of proteins belonging to different types of Panax (640 entries, downloaded on 02 July 2014 from Uniprot) and the database of *Arabidopsis thaliana* proteins (31760 entries, downloaded on 18

February 2014 from UniProt) were used for spectra matching. Cysteine carbamidomethylation and methionine oxidation were set as variable modifications. Peptide mass tolerance was set to 10 ppm, fragment mass tolerance to 0.5 Da and maximum number of missed cleavages = 2. The false discovery rate (FDR) for peptide identification was set at 0.05; only proteins identified by at least two peptides were considered as genuine identifications and further analyzed.

2.6 Analysis of identified proteins

In order to describe the functional classes of identified proteins, a Gene Ontology (GO) analysis was performed by using the web available software QuickGo (www.ebi.ac.uk/QuickGo).

An *interactomics* map was built up by means of STRING (Search Tool for the Retrieval of Interacting Genes) v.9.1 software (http://stringdb.org/), set on *Arabidopsis thaliana* as organism database. This is a large database of known and predicted protein-protein interactions. Proteins were represented with nodes and the interactions with continuous lines to represent direct interactions (physical), while indirect ones (functional) were presented by broken lines. All the edges were supported by at least a reference from the literature or from canonical information stored in the STRING dataset. Cluster networks were created by using the K-means algorithm which is included in the STRING website and a value of 7 was selected for all the analyses. The pathways classification was done after the automatic enrichment in STRING, based on the information provided by the KEGG-Pathway Database.

2.7. Analysis of proteolytic peptides obtained by simulated gastrointestinal digestion

Bioactive peptides encrypted in the Panax proteome were predicted by combining different in-silico enzymatic digestions in order to simulate the human gastrointestinal process: pepsin (stomach) and intestinal enzymes (trypsin, chymotrypsin, elastase, carboxypeptidase A and B and aminopeptidases). In order to perform the simulated gastrointestinal digestion, a global protein database was created by merging all MS tabular report, obtained by searching MS spectra of initial protein extract and all CPLLs eluates, against an *Arabidopsis Thaliana* database. All the digestions were performed *in-silico* using the MS-Digest software which is included in the ProteinProspector v 5.10.1 website (http://prospector.ucsf.edu/prospector/mshome.htm). To evaluate the results, all the potential peptides were ranked by using the PeptideRanker software (http://bioware.ucd.ie/~testing/biowareweb/), using the N-to-1 neural network probability to predict which peptides could be more bioactive. In addition, all the potential bioactive peptides were compared with the CAMP database, which includes known antimicrobial bioactive peptides (http://www.bicnirrh.res.in/antimicrobial/).

3. Results

Figure 1 shows a complete scheme of protein extraction, CPLLs enrichment, proteins separation by SDS-PAGE electrophoresis and their identification by MS analysis.

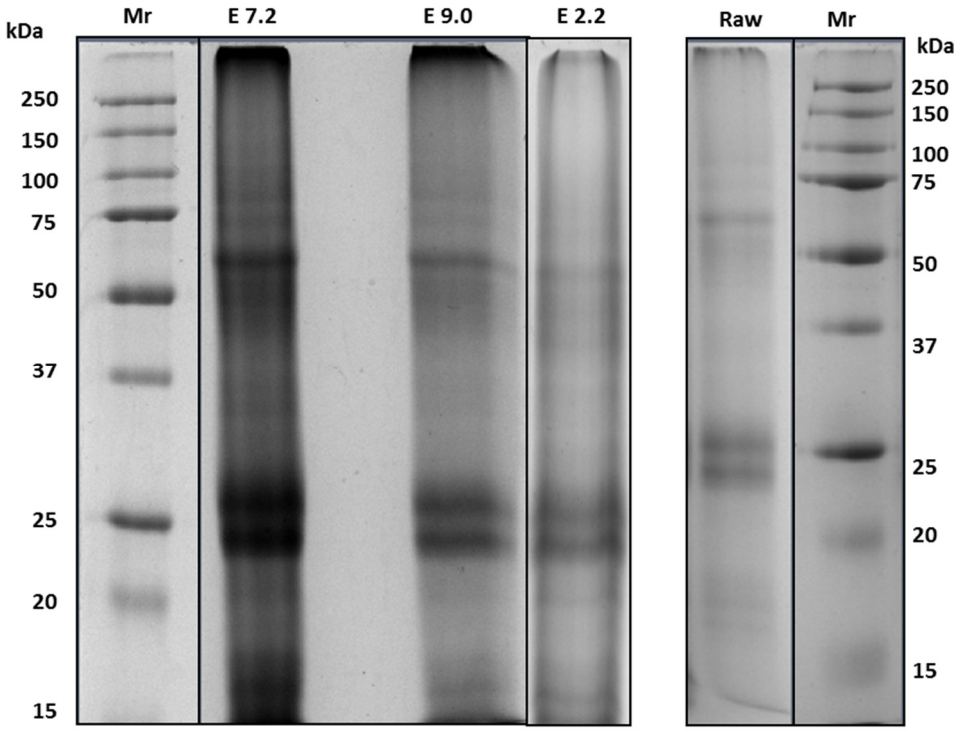


Figure 2 - SDS-PAGE separation of the untreated (Raw) sample vs. the eluates of CPLL capture at the three different pH values (tracks E2.2, E7.2 and E9.0). Each lane has been cut into segments, the proteins digested and sent to MS analysis. Mr: molecular mass ladder. Staining with micellar Coomassie blue.

In order to increase the knowledge on the *Panax ginseng* proteome, we have performed the most efficient protocol able to extract proteins from a dried root powder. After testing various extraction procedures reported in literature [21], we have applied the initial steps of the protocol described by Wang et al. [28], which required preliminary washes of dried powder in order to eliminate interferences and contaminants until the plant powder became colorless. The protein extraction was performed by using a sample buffer previously described [22]. The final combination of TCA and acetone is commonly used to precipitate proteins and to remove any

remained contaminants. By using a combination of different methods, we were able to extract about 300 mg of soluble proteins from 3 g of sample, quantified by the BioRad protein assay with BSA as standard [33]. The extracted protein quantity was appropriate for incubation with bead-based libraries of combinatorial peptide ligands (CPLLs), a protein enrichment technology which simultaneously dilutes high-abundance proteins and concentrates low-abundance ones [34]. The CPLLs technology was originally optimized for use with human biological samples [35], but it was recently adopted with other sample types like plants and foodstuffs within the recommend total protein range (>50 mg of proteins for 100 μL beads volume) [36-38]. So in order to aim to a very large protein discovery, a sample overloading, as well as a capture at three pH values, are recommended.

The CPLLs treatment has contributed to increase protein capture, as seen by the SDS-PAGE profiling in figure 2 where the untreated sample (Raw) is characterized by three major protein bands, while CPLLs beads eluates from the three pH values (lanes E 2.2, E 7.2 and E 9.0) exhibit additional bands, particularly increased in the regions of the gel corresponding to 75-30 KDa and 20-15 KDa. In fact, while the electrophoretic profile of the raw extract has displayed only 3 or 4 evident protein bands, probably corresponding to high-abundance proteins, the CPLLs eluates profiles have revealed more intense and much more resolved protein bands, referred both to high-abundance proteins and to low-abundance ones. The efficiency and the potentiality of CPLLs treatment has been demonstrated by mass spectrometry analysis.

Figure 3 shows the number of proteins identified by matching the experimental spectra to two different databases: *Arabidopsis thaliana* and *Panax ginseng*.

Database	Database entries	Raw extract	CPLLs pH 7.2	CPLLs pH 2.2	CPLLs pH 9.0	Total ID
Arabidopsis Thaliana	31706	108	83	51	96	**152**
Panax Ginseng	640	42	44	38	31	**55**

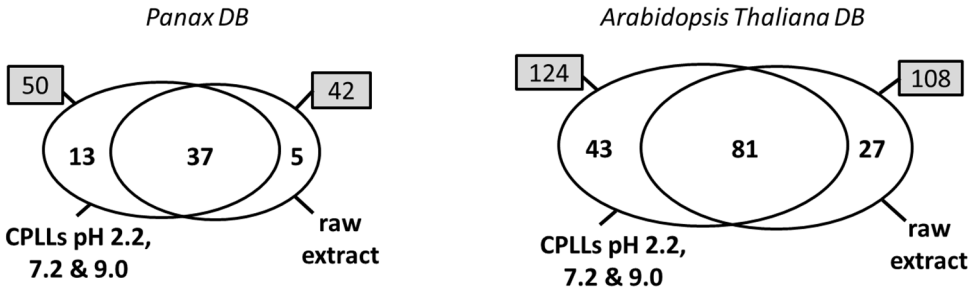


Figure 3 - MS data of the identified proteins. Upper panel: table of protein IDs as obtained in a raw extract and all CPLLs eluates by using both Panax ginseng and Arabidopsis thaliana databases, respectively. The total unique IDs were 152 for Arabidopsis thaliana database and 55 for Panax ginseng. Lower panel: Venn diagrams comparing the number of identified proteins in all CPLLs eluates with those recognized in the raw extract, considering both databases. In reagard to the Panax ginseng db, 37 proteins were found both in untreated and eluted samples, while 13 were specifically of CPLLs eluates. Moreover a quite high number of identifications was specific of each beads capture (e.g. 8 specific proteins for CPLLs treatment at pH 7,2). Also for Arabidopsis Thaliana db 81 were common identifications from all samples, while 43 were exclusive proteins of three eluates. The higher number of 27 specific proteins was obtained after beads capture at pH 9,0.

The use of two different databases was prompted by the necessity to increase the number of identified proteins. In fact while the *Panax ginseng* database is most specific, it's characterized by a large number of unreviewed entries. Moreover, the lack of complete DNA sequencing cause the incompleteness of the *Panax ginseng* proteome. The choice of *Arabidopsis thaliana* database is due to many reasons: first of all its database is complete and characterized by a higher number of entries than the *Panax ginseng* database (31706 vs. 640); in fact, the full proteome of *Arabidopsis*

thaliana is available at the UniProt depository as reference proteome. Secondly, *Arabidopsis thaliana* is commonly used as model organism for studying plant sciences and biology, after its complete genetic mapping and sequencing in 2000. The table displays the IDs obtained in the untreated sample (raw extract) and in all CPLLs eluates: many species are recognized in all samples so, in order to know the correct number of unique gene products, the last column reports the total distinct IDs without redundancies. The Venn diagrams, in the lower panel of figure 3, compare the identified proteins in CPLLs eluates vs. those recognized in raw extracts considering both types of databases. 13 and 43 more proteins, belonging to the *Panax ginseng* and the *Arabidopsis thaliana* database respectively, could be identified only after CPLLs treatment. Finally the different contributions of each CPLLs incubation were considered for a deep investigation of Panax proteome: for both databases, CPLLs treatments at different pH values were successful for capture of specific proteins, undetectable applying only the standard protocol. Concerning the *Panax ginseng* database, the most performant incubation condition was at physiological pH, able to capture 8 unique ginseng's proteins. Considering the *Arabidopsos thaliana* database, the beads incubation at pH 9,0 has identified the higher number of specific proteins (27) than other two pH conditions (15 at pH 7,2 and 9 at pH 2,2).

Figure 4 depicts a pie chart of Gene Ontology (GO) analysis applied to all 207 identified species in order to understand /categorize the belonging protein classes.

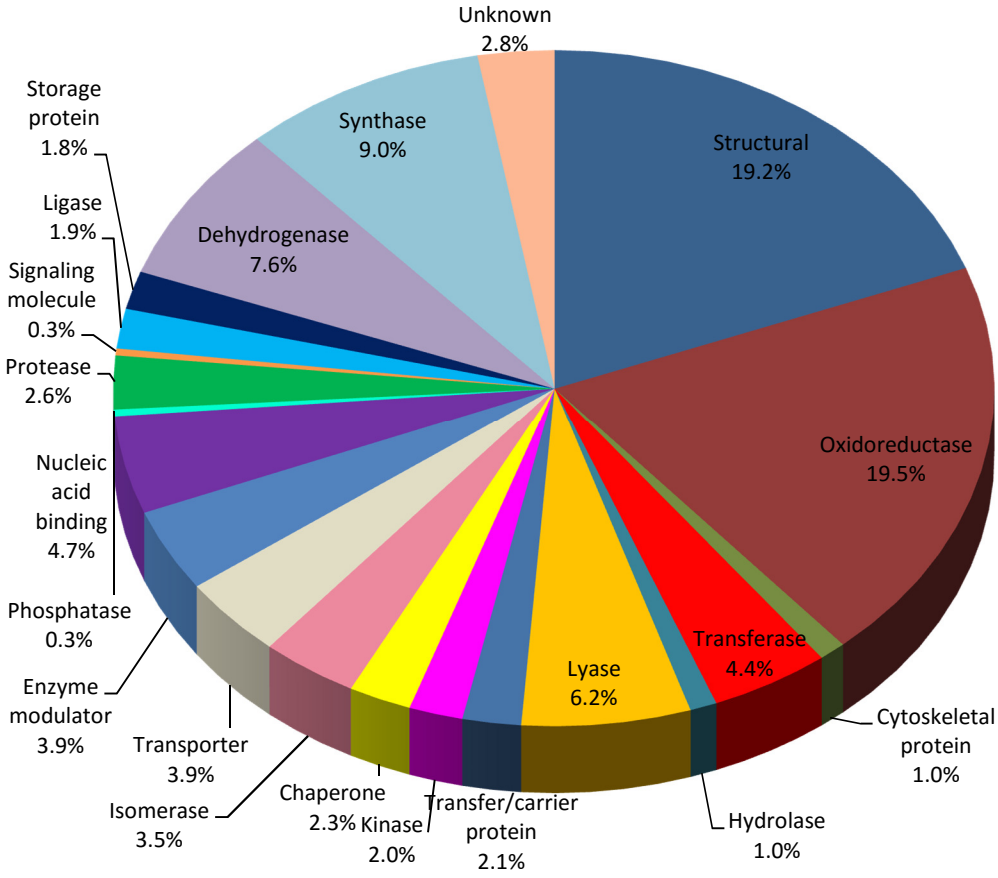


Figure 4 - Biological funtional categorization of 207 unique proteins expressed in ginseng dried root powder via GO analysis. The pie chart shows the percentage distribution of reported GO terms.

Figure 5 displays the general workflow for the analysis of the peptides obtained after an *in-silico* human gastrointestinal digestion: the software simulates digestion of pepsin, trypsin, chymotrypsin, elastase, carboxipeptidases (A and B) and aminopeptidases, using the MS-Digest program.

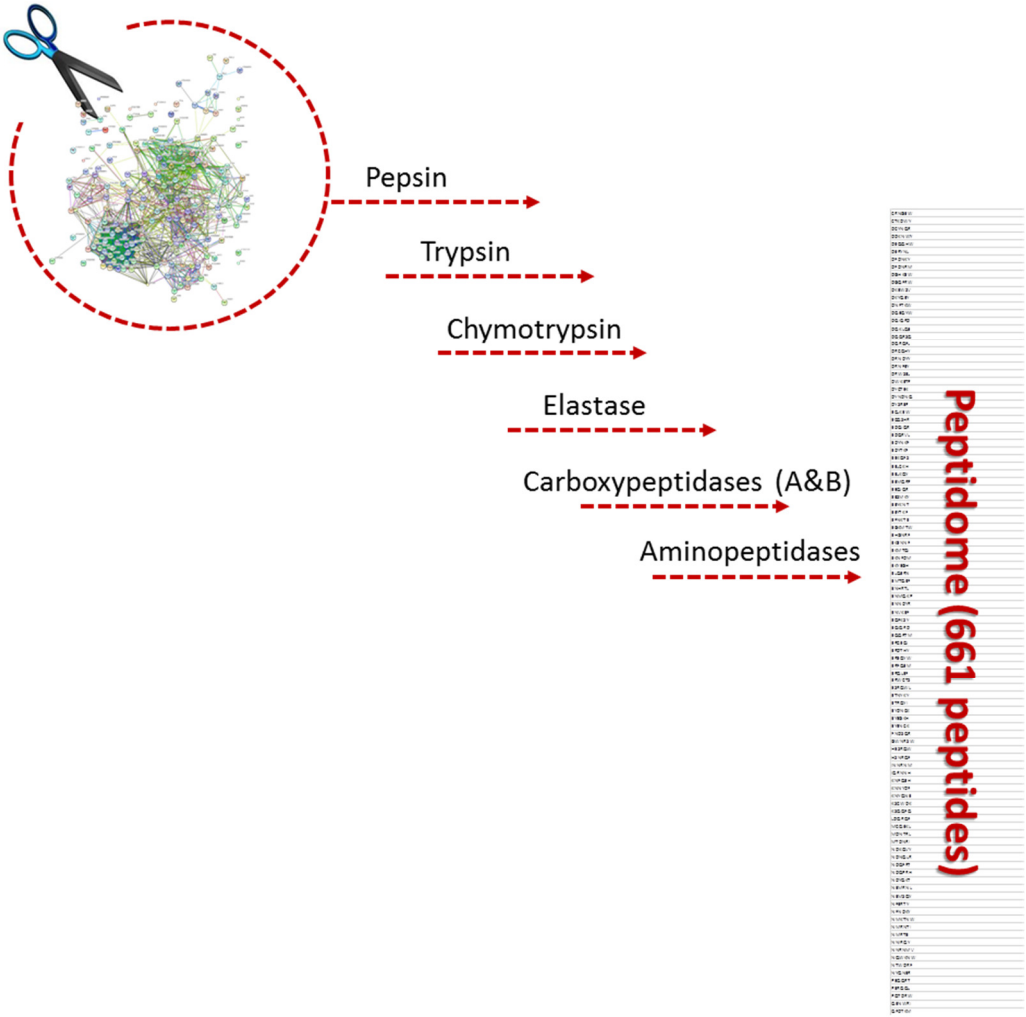


Figure 5 - In-silico simulation scheme of human gastro-intestinal digestion, performed on all proteins identified via the Arabidopsis thaliana database. All the listed digestive enzymes generated a total number of 661 peptides.

All generated peptides (661) were ranked on the basis of their predicted bioactivity by software the PeptideRanker, which was able to select 95 bioactive peptides focusing our attention on their potential antimicrobial action as reported in Figure 6.

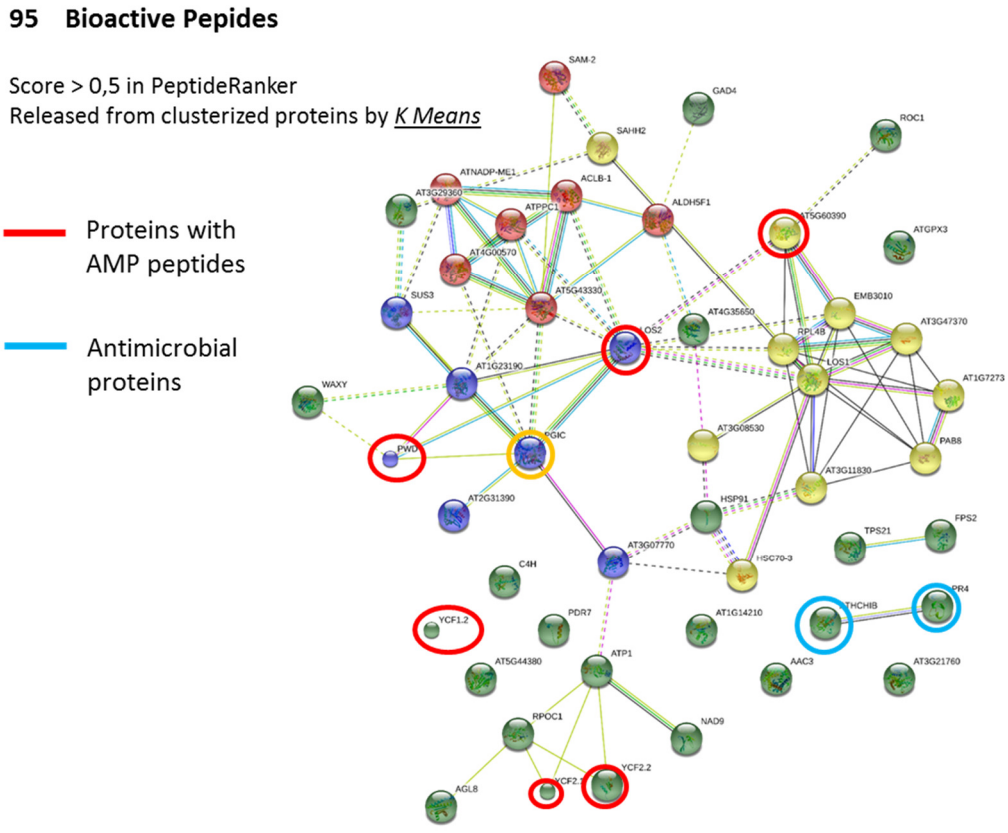


Figure 6 - Representation of 95 proteins releasing bioactive peptides, clusterized by KMeans; bioactive peptides have been ranked on the bases of their a Peptideranker score >0.5 calculated by the corresponding software. All nodes in interactomic map were proteins, showing different activities like antimicrobial or defensive functions.

4. Discussion

4.1. Protein identification

The deep proteome identification was performed by using two different databases, *Arabidopsis thaliana* and *Panax ginseng*, in order to increase the number of panax species discovered. In fact a total number of 206 proteins were identified by merging all unique IDs obtained considering both databases. Our results have substantially improved a previous study of Kim *et al.* [22], who recognized only 17 proteins using MALDI-TOF MS. The low rate of identified species was probably due to lack of ginseng genome DNA sequence database. Also for this reason we have decided to perform a search by considering the databank of plant species mostly used as model organism. In order to increase protein identification, Kim *et al.* have performed a BLAST search of amino acids sequences using the ginseng EST databases, finally recognizing 87 unique gene products. Our quite large discovery has been made possible thanks to very sensitive high resolution and high mass accuracy of the Orbitrap mass spectrometer, as well as to the use of CPLLs, which have enhanced by more tha 20% the final discovery. Indeed, as reported in the lower Venn diagram of Figure 3, the adopted strategy has been indispensable to capture 13 more proteins, via identification by the *Panax* db, and 43 more proteins, via the *Arabidospis Thaliana* db. Even if these numbers are not too high, we believe that the CPLLs technology has contributed to proteome knowledge of this perennial plant whose genome is not fully sequenced. Moreover the strategy to use different pHs incubations has contributed to capture specific unique gene products, setting up the most performant

conditions for a more efficient interaction between ginseng's proteins and hexapeptides chains.

4.2. Protein functional data

To gain further knowledge of the *Panax ginseng* proteome functionality, the 206 identified proteins were analysed by Gene Ontology, ascertaining their molecular function, and all GO categories were described in a pie chart of Figure 4. The most enriched protein classes were: structural (19.2%) and oxidoreductase (19,5%). The first category was probably a consequence of tissue type considered in this study. In fact roots are a very simple plant tissue, consisting of a central vascular system surrounded by a large storage parenchyma cells on the outside protective layer. So this sample type has justified the discovery of large amount of proteins with structural function. In fact, for example, proteins of cytoskeleton, like tubulin alpha-3 and tubulin beta-4 identified by MS analysis (TBA3_ARATH, TBB4_ARATH), and pectins are involved in the cell wall structure and the changes in structure/chemistry of cell walls directly affect the ripening and senescence of plants [40-41]. Moreover oxidoreductases are enzymes able to catalyse the transfer of electrons from one molecule to another one and they are normally involved in different metabolic processes like signalling and regulation of growth [39], transport activities and defense against pathogens [42]. Another large percentage of identified proteins belongs to the class of dehydrogenases (7.6%), enzymes able to interconvert alcohols into aldehydes or ketones with the reduction of nicotinamide adenine dinucleotide (NAD^+ to NADH), like our identification NADH dehydrogenase iron-sulphur protein 3 (NDUS3_ARATH). They are expressed at low levels in roots of

young plants grown on agar and at a high level in lack of oxygen or water and in low temperatures [43]. In addition, a particular dehydrogenase, the glyceraldehyde 3-phosphate dehydrogenase, identified via MS in our *Panax* extracts (Q6VAL5_PANGI, D0VFU1_PANGI), is involved in glycolysis, a pathway able to break starch into glucose. We speculate that this result could indicate that ginseng mainly absorbs energy as vegetative storage for ginseng root survival, which is destroyed through glycolysis to release the energy required for root maturation in slow-growth period [24]. Another major functional class of proteins is represented by synthases (9.0%), which catalyzes numerous synthesis processes. For example, the overexpression of Panax ginseng squalene synthase causes a remarkable increase of phytosterols as well as ginsenoside contents [44]. Proteins belonging to the synthesis category such as granule-bond starch synthase I identified in our samples (Q6XY46_9APIA), increase with ginseng root age, indicating that protein synthesis enhances the ginseng maturation at different metabolic levels [9].

In order to investigate protein-protein interactions, a proteome-scale interaction network was created by merging all the proteins identified for the *Panax ginseng* proteome by using STRING software (Figure 7). The databases include interactions from published literature describing experimentally studied interactions, as well as those from genome analysis by using several well-established methods based on domain fusion, phylogenetic profiling and gene neighborhood concepts. The whole *Arabidopsis Thaliana* genome was selected as a reference set, correlating the new results with the previous MS analysis. This network obtained represents the first comprehensive interactomics map for the *Panax ginseng* proteome and provides an interesting framework for navigating through the proteome.

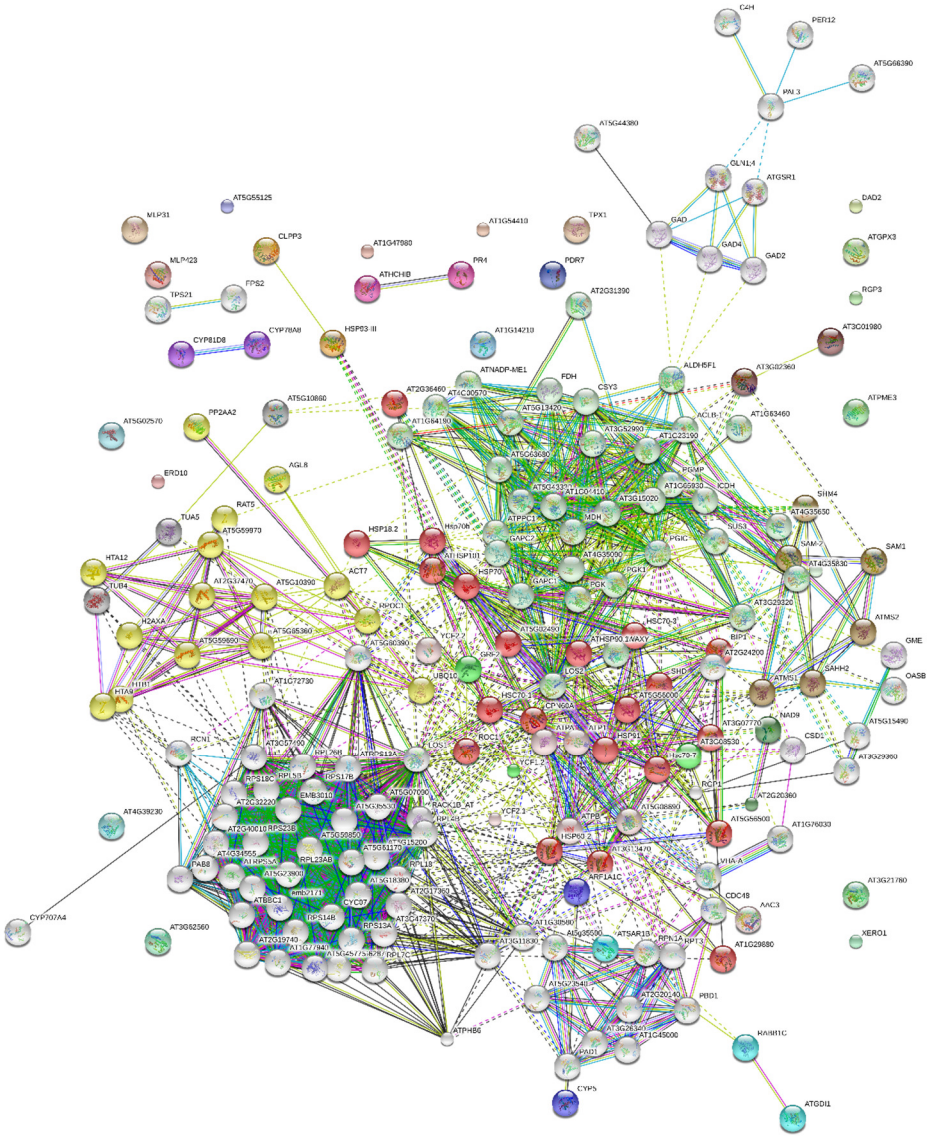


Figure 7 - interactomics map realized by STRING v.9.1 software, showing a protein-protein interaction network formed by 196 nodes and 1554 interactions. Clusters realized applying the algorithm MCL (factor 2).

The topological analysis of this network has demonstrated several sparsely connected sub-networks, including that one corresponding to the ribosomal proteins with highly connected interactions (Figure 7). The KEGG-Pathway Enrichment in STRING has revealed the presence of the canonical pathways: energy metabolism, carbohydrate and lipid metabolism and nucleotide and amino acid metabolism (Table 1).

KEGG PATHWAY	N° of genes
Metabolic pathways 65.9%	
Energy metabolism	
Carbon fixation in photosynthetic organisms	**10**
Oxidative phosphorylation	**8**
Nitrogen metabolism	**3**
Photosynthesis	**2**
Sulfur metabolism	**1**
Carbohydrate and lipid metabolism	
Carbohydrate metabolism	
Citrate cycle (TCA cycle)	**10**
Pyruvate metabolism	**8**
Glycolysis / Gluconeogenesis	**8**
Starch and sucrose metabolism	**8**
Glyoxylate and dicarboxylate metabolism	**7**
Pentose phosphate pathway	**6**
Butanoate metabolism	**4**
Ascorbate and aldarate metabolism	**3**
Pentose and glucuronate interconversions	**2**
Galactose metabolism	**2**
Fructose and mannose metabolism	**1**
Sugar metabolism	
Amino sugar and nucleotide sugar metabolism	**8**

Fatty acid metabolism	
Arachidonic acid metabolism	**1**
Glycan metabolism	
N-Glycan biosynthesis	**1**
Terpenoid backbone biosynthesis	**1**
Sesquiterpenoid biosynthesis	**1**
Other terpenoid biosynthesis	
Limonene and pinene degradation	**2**
Carotenoid biosynthesis	**1**
Nucleotide and amino acid metabolism	
Purine metabolism	**4**
Pyrimidine metabolism	**1**
Amino acid metabolism	
Alanine, aspartate and glutamate metabolism	**6**
Cysteine and methionine metabolism	**5**
beta-Alanine metabolism	**3**
Phenylalanine metabolism	**4**
Arginine and proline metabolism	**2**
Glycine, serine and threonine metabolism	**1**
Tyrosine metabolism	**1**
Tryptophan metabolism	**1**
Other amino acid metabolism	
Glutathione metabolism	**6**
Taurine and hypotaurine metabolism	**3**
Selenocompound metabolism	**2**
Cyanoamino acid metabolism	**1**
Metabolism of cofactors and vitamins	
One carbon pool by folate	**1**
Secondary metabolism	
Biosynthesis of secondary metabolites	
Phenylpropanoid biosynthesis	**4**
Stilbenoid, diarylheptanoid and gingerol biosynthesis	**3**
Flavonoid biosynthesis	**1**

Genetic Information Processing 27.4%	
Translation	
Ribosome	**27**
mRNA surveillance pathway	**3**
RNA transport	**2**
Aminoacyl-tRNA biosynthesis	**1**
Transcription	
Spliceosome	**5**
RNA polymerase	**1**
Folding, sorting and degradation	
Protein processing in endoplasmic reticulum	**11**
Proteasome	**8**
RNA degradation	**3**
Cellular Processes 5.8%	
Transport and catabolism	
Endocytosis	**6**
Peroxisome	**4**
Phagosome	**3**
Organismal Systems 0.9%	
Environmental adaptation	
Plant-pathogen interaction	**2**

Table 1 – The KEGG-Pathway enrichment, performed by STRING software, has revealed the presence of canonical pathways: metabolic pathways (65.9%), genetic information processes (27.4%), cellular transport and catabolism (5.8%) and environment adaptation (0.9%). the numbers of involved genes are reported for all subcategories of enriched cell pathways.

Among them the metabolic processes are the most relevant biological processes, representing the 65.9% of global dataset. KEGG pathway data have confirmed the proteins identifications by MS analysis: most of interactions involved proteins of energy metabolisms, like oxodoreductases recognized previously, of carbohydrate

metabolism (glycolysis) such as glyceraldehyde 3-phosphate dehydrogenase. Considering another pathway map class, the *Genetic Information Processes* has included 27 ribosomal proteins, which represented the top protein class of interactomics map. In conclusion, the investigation of the protein network, realized by using MS- proteomics data, has allowed the global protein–protein interactomics inspection of *Panax ginseng* proteome, providing a first reference map for further proteomics studies.

4.3. Analysis of proteolytic peptides

In order to explore the possibility to identify bioactive peptides, encrypted in the global *Panax ginseng* proteome, the physiologic enzymatic digestion of proteins was simulated in-silico by sequential hydrolysis with pepsin (stomach) and intestinal enzymes (trypsin, chymotrypsin, elastase, carboxypeptidase A and B and aminopeptidases). Despite restricted parameters such as one missed cleavages and a minimum of 6 residues per peptide, smaller peptides were not considered in the study because their sequence might not be unique and may belong to several different proteins. This model could not be completely prognostic because it didn't consider a lot of individual physiological factors, like pH variations, peptidase/protein ratio and interactions with other molecules. The predicted peptidome resulted into 661 different peptides, as shown in Figure 5, and it was subsequently investigated by using different databases that include known bioactive peptides. By using PeptideRanker (http://bioware.ucd.ie/~testing/ biowareweb/) the complete list of potential bioactive peptides was ranked using the N-to-1 neural network probability, which predicts more bioactive peptides (95). Among them, with a score higher than

0.5 (6–11 residues), the majority corresponded to peptides encrypted in the nucleotide/nucleoside phosphate binding proteins (Figure 6).

Antimicrobial peptides (AMPs), characterized by score < –0.251, were identified using the CAMP (Collection of Anti-Microbial Peptides) database (http://www.bicnirrh.res.in/antimicrobial/) and applying the DAC score (Discriminate Analysis Classifier score). Table 2 despicts 6 potential bioactive peptides with anti-microbial activities.

Peptide Ranker_ Score	Peptides Sequence	AMP Probability (CAMP_Analysis)	Accession Number	Protein Description	Protein Family	Function
0.73	NQWKNW	0.989	P56785	Full=Protein TIC 214; AltName: Full=Translocon at the inner envelope membrane of chloroplasts 214;	Belongs to the TIC214 family The Chloroplast Envelope Protein Translocase (CEPT or Tic-Toc)	Involved in protein precursor import into chloroplasts. May be part of an intermediate translocation complex acting as a protein-conducting channel at the inner envelope. Science 339:571-574(2013)
0.5	KNNYDF	0.905	P56785	Full=Protein TIC 214; AltName: Full=Translocon at the inner envelope membrane of chloroplasts 214;		
0.73	RPDRNF	0.88	P56786	Protein Ycf2 (chloroplast)	Belongs to the Ycf2 family	This family consists of chloroplast encoded Ycf2, which is around 2000 residues in length. The function of Ycf2 is unknown, though it may be an ATPase. Its retention in reduced chlorplast geneomes of non-photosynthetic plants, and transformation experiments in tobacco indicate that it has an essential function which is probably not related to photosynthesis.
0.83	QNEWGW	0.774	P25696	Full=Bifunctional enolase 2/transcriptional activator; AltName: Full=2-phospho-D-glycerate hydro-lyase 2; AltName: Full=2-phosphoglycerate dehydratase 2;	Belongs to the enolase family	Multifunctional enzyme that acts as an enolase involved in the metabolism and as a positive regulator of cold-responsive gene transcription.
0.81	ICCCNKM	0.699	A8MSE8	Elongation factor 1-alpha	Belongs to the TRAFAC class translation factor GTPase superfamily. Classic translation factor GTPase family. EF-Tu/EF-1A subfamily.	This protein promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis.
0.85	GWNRSW	0.603	Q9STV0	Alpha-glucan water dikinase 2	Belongs to the PEP-utilizing enzyme family	Mediates the incorporation of phosphate into alpha-glucan, mostly at the C-6 position of glucose units

Table 2 – It lists 6 peptides characterized by antimicrobial function (AMP), reporting all parameters like PeptideRanker score, AMP probability,peptide sequence and length. Also other parameters, referred to the original proteins, like accession number, protein family and function, are shown.

Two peptides belong to traslocon protein of chloroplast membrane, involved in protein precursor import to chloroplasts [45]. RPDRNF is a peptide of chloroplast protein Ycf2 previously identified by MS analysis, whose function is still unknown even if it seems to be a ATPase [46]. Other two peptides (QNEWGW and ICCCNKM) derived from proteins involved in cell metabolism: in particular the second one derives from elongation factor 1-alpha recognized in Panax samples and it seems to be involved in protein biosynthesis inside ribosomes. The last peptide, GWNRSW, belongs to alpha-glucan water dikinase 2 protein, an enzyme of PEP-utilizing family. Although all these potential bioactive peptides need to be validated by further bioactivity assays using corrspondent synthetic peptides, we believe that the computational methods are useful for a preliminary peptidomic analysis. In fact they are fast and low cost alternatives, able to predict and reduce the number of potential targets to be investigated. In addition, bioinformatics-driven tools provide useful insights not achievable in human or animal model studies.

5. Conclusions

Up to the present the lack of complete knowledge of *Panax ginseng* proteome has contributed to a defective characterization of ginseng proteins composition and functions. Despite numerous efforts to sequence *Panax ginseng* genome [47, 48], most research has focused on ginsenosides investigations [49-51]. This work provides new insight into the proteome of the Chinese traditional medical plant root, coupling the power of the Combinatorial Peptide Ligand Library (CPLL) technology, the SDS-PAGE separation and the high resolution nLC-ESI MS/MS. We have successfully identified 55 proteins belonging to *Panax ginseng* database and 152

belonging to *Arabidopsis thaliana* database, for a total of 207 identifications. The GO software has permitted a classification of proteins in order to understand the role and the function of the identified species. The four major classes are represented by oxidoreductase (19.5%), structural compound (19.2%), synthase (9.0%) and dehydrogenase (7.6%). In order to explore proteins interactions useful to understand the *Panax ginseng* biology, an interactomics map was built, thus obtaining a complex grid formed by 196 nodes connected via 1554 interactions. The protein-protein interaction network was simplified in a table, reporting the pathways classification, which has confirmed the previous GO analysis with a prevalence (65.9%) of proteins involved in metabolic pathways. Our proteomic research has concluded with a novelty in *Panax ginseng* studies: a peptidomic investigation by simulating *in-silico* the enzymatic protein digestion in the human gastrointestinal tract. A total of 661 different peptides were generated, but only 95 could be considered as potentially bioactive, including 6 with a predicted anti-microbial function. This approach could be an useful novelty to evaluate a possible antimicrobial activity expressed by proteolytic peptides, while normally antioxidant and antimicrobial actions are mainly associated to ginsenosides present in *Panax ginseng* root. For example recently fresh root extracts, from 4 years old *Panax ginseng* plants, have been explored for the synthesis of silver and gold nanoparticles, characterized by potential antimicrobial functions against pathogenic microorganisms [52-54]. In conclusion our proteomic and peptidomic analysis aimed to provide a deep investigation of *Panax ginseng* components for a better understanding of ginseng metabolism connected with beneficial effects, which have contributed to the diffusion as the most valuable traditional Asian medicine.

References

1. Lee FC. Facts about Ginseng, the Elixir of Life. Hollyn International Corp., Elizabeth, NJ, 1992.
2. Huang KC. The Pharmacology of Chinese Herbs. CRC Press, Boca Raton, FL, 1999.
3. A.S. Attele, J.A. Wu, C.-S. Yuan. Ginseng Pharmacology. Biochem. Pharmacol., 1999; 58:1685-1693.
4. Gillis CN, *Panax ginseng* pharmacology: A nitric oxide link? Biochem. Pharmacol., 1997, 54**:** 1–8.
5. You J., Liu X., Zhang B., Xie Z., Hou Z., Yang Z., Seasonal changes in soil acidity and related properties in ginseng artificial bed soils under a plastic shade. J. Ginseng Res. 2015, 39: 81-88.
6. Kim C., Choo G. C., Cho H. S., Lim J. T., Soil properties of cultivation sites for mountain-cultivated ginseng at local level. J. Ginseng Res. 2015, 39: 76-81.
7. Kim Y. K., Jeon J. N., jang M. G., Oh J. Y., Kwon W. S., Jung S. K., Yang D. C., Ginsenoside profiles and related gene expression during foliation on *Panax*. J. Ginseng Res. 2014, 38: 66-72.
8. Lee J. H., Lee J. S., Kwon W. S., Kang J. Y., Lee D. Y., In J. G., Kim Y. S., Seo J., Baeg J. S., Chang I. M., Grainger K. Characteristics of Korean ginseng varieties of Gumpoong, Sunun, Sunpoong, Sunone, Cheongsun and Sunhyang. J. Ginseng Res. 2015, 39: 91-104.
9. Ma R., Sun L., Chen X., Jiang R., Sun H., Zhao D. Proteomic changes in different growth periods of ginseng roots. Plant Phys. and Biochem. 2013, 67: 20-32.
10. Nam M. H., Heo E. J., Kim J. Y., Kim S., Kwon K. H., Se J. B., Kwon O., Yoo J. S., Park Y. M. Proteome analysis of the response of Panax ginseng C. A. Meyer leaves to high light: use of electrospray ionization quadrupole-time of flight massa spectrometry and expressed sequence tag data. Proteomics 2003, 3: 2351-2367.
11. B.X. Wang, Q.L. Zhou, M. Yang, Y. Wang, Z.Y. Cui, Y.Q. Liu, I. Takashi, Hypo- glycemic activity of ginseng glycopeptides, Acta Pharmacol. Sin. 2003, 24: 50-54.
12. J.H. Kang, K.H. Song, J.K. Woo, M.H. Park, M.H. Rhee, C. Choi, S.H. Oh, Ginse- noside Rp1 from Panax ginseng exhibits anti-cancer activity by down- regulation of the IGF-1R/Akt pathway in breast cancer cells, Plant Food Hum. Nutr. 2011, 66: 298-305.
13. Sharma J., Goyal P. K. Chemoprevention of chemical-induced skin cancer by *Panax ginseng* root extract. J. Ginseng Res. 2015, 39: 265-273.
14. Kim S. J., Kim A. K., Anti-breast cancer activity of fine black ginseng (*Panax ginseng* Meyer) and ginsenoside Rg5. J. Ginseng Res. 2015, 39: 125-134.
15. T. Yokozawa, E. Dong, H. Watnabe, H. Oura, H. Kashiwagi, Increase of active oxygen in rats after nephrectomy is suppressed by ginseng saponin. Phyt- other. Res. 1996, 10: 569-572.
16. Y. Wanga, R. Jiangb, G. Lic, Y. Chenb, H. Luob, Y. Gaob, Q. Gao, Structural and enhanced memory activity studies of extracts from Panax ginseng root, Food Chem. 2010, 119: 969-973.

17. J.L. Reay, A.B. Scholey, D.O. Kennedy, Panax ginseng (G115) improves aspects of working memory performance and subjective ratings of calmness in healthy young adults, Hum. Psychopharm. Clin. 2010, 25: 462-471.
18. Baek K. S., Hong Y. D., Kim Y., Sung N. Y., Yang S., Lee K. M., Park J. Y., Park J. S., Rho H. S., Shin S. S., Cho J. Y. Anti-inflammatory activity of AP-SF, a ginsenoside-enriched fraction, from Korean ginseng. J. Ginseng Res. 2015, 39: 155-161.
19. Lee C. H., Kim J. H., A review on the medical potentials of ginseng and gensenosides on cardiovascular diseases. J. Ginseng Res. 2014, 38: 161-166.
20. Liang X., Chen X., Liang Q., Zhang H., Hu P., Wang Y., Luo G. Metabonomic study of chinese medicine *shuanglong* formula as an effective treatment for myocardial infarction in rats. J. Prot. Res. 2011, 10: 790-799.
21. J.H. Lum, K.L. Fung, P.Y. Cheung, M.S. Wong, C.H. Lee, F.S. Kwok, M.C. Leung, et al., Proteome of oriental ginseng Panax ginseng C.A. Meyer and the potential to use it as an identification tool, Proteomics 2002, 2: 1123-1130.
22. Kim SI, Kim JY, Kim EA, Kwon KH, Kim KW, Cho K, Lee JH, Nam MH, Yang DC, Yoo JS, Park YM Proteome analysis of hairy root from Panax ginseng C.A. Meyer using peptide fingerprinting, internal sequencing and expressed sequence tag data. Proteomics. 2003, 3: 2379-2392.
23. J.Y. Yoon, B.H. Ha, J.S. Woo, Y.H. Lim, K.H. Kima, Purification and characterization of a 28-kDa major protein from ginseng root. Comp. Biochem. Physiol. Part B 2002, 132: 551-557.
24. S.I. Kim, S.M. Kweon, E.A. Kim, J.Y. Kim, S. Kim, J.S. Yoo, Y.M. Park, Characterization of RNase-like major storage protein from the ginseng root by proteomic approach, J. Plant Physiol. 2004, 161: 837-845.
25. M.H. Nam, S.I. Kim, J.R. Liu, D.C. Yang, Y.P. Lim, K.H. Kwon, J.S. Yoo, Y.M. Park, Proteomic analysis of Korean ginseng (Panax ginseng C.A. Meyer). J. Chromatogr. B. 2005, 815: 147-155.
26. L. Sun, X. Lei, R. Ma, R. Jiang, D. Zhao, Y. Wang, Two-dimensional gel electrophoresis analysis of different parts of Panax quinquefolius L. root. Afr. J. Biotechnol. 2011, 10: 17023-17029.
27. L. Sun, P. Ma, R. Ma, X. Lei, X. Chen, C. Qi, Two-dimensional electrophoresis analysis for different parts of Panax ginseng C.A. Meyer root. J. Huazhong Norm. Univ. (Nat. Sci.), 2010, 44: 639-643.
28. Wang W., Vignani R., Scali M., Cresti M., A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. Electrophoresis 2006, 27: 2782-2786.
29. Fasoli E., Righetti P. G., Proteomics of fruits and beverages, Curr. Opin. Food Science 2015, 4:76–85.
30. Fasoli E, Farinazzo A, Sun CJ, Kravchuk AV, Guerrier L, Fortis F, Boschetti E, Righetti PG. Interaction among proteins and peptide libraries in proteome analysis: pH involvement for a larger capture of species. J Proteomics 2010, 73:733-742.
31. Candiano G, Dimuccio V, Bruschi M, Santucci L, Gusmano R, Boschetti E, Righetti, PG, Ghiggeri GM. Combinatorial peptide ligand libraries for urine proteome analysis: investigation of different elution systems. Electrophoresis 2009, 30:2405-2411.

32. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., Mann, M. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 1996, 379, 466– 469.
33. Bradford M. M. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 1976, 7: 248-254.
34. Righetti PG, Fasoli E, Boschetti E. Combinatorial peptide ligand libraries: the conquest of the 'hidden proteome' advances at great strides. Electrophoresis. 2011;32:960-966.
35. U. Restuccia, E. Boschetti, E.Fasoli, F. Fortis, L. Guerrier, A. Bachi, AV. Kravchuk, P.G. Righetti. pI-based fractionation of serum proteomes versus anion exchange after enhancement of low-abundance proteins by means of peptide libraries. J Proteomics. 2009, 72:1061-1070.
36. E. Boschetti; L. Bindschedler; C. Tang; E. Fasoli; P.G. Righetti. Combinatorial peptide ligand libraries and plant proteomics: a winning strategy at a price. Journal of Chromatography A, 2009, 1216:1215-1222.
37. Fasoli E, Righetti PG. The peel and pulp of mango fruit: A proteomic samba. Biochim Biophys Acta, 2013;1834:2539-2545.
38. Righetti PG, Esteve C, D'Amato A, Fasoli E, Luisa Marina M, Concepción García M. A sarabande of tropical fruit proteomics: Avocado, banana, and mango. Proteomics 2015,15:1639-1645.
39. Robinson N. G., Procter C. M., Connoly E. L., Guerinot M. L. A ferric-chelate reductase for iron uptake from soil. Nature 1999, 397:694-697.
40. Brummell D. A., Harpster M. H. Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. Plant Mol. Biol. 2011,47:311-340.
41. Brummell D. A., Cin V. D., Crisosto C. H., Labavitch J.M. Cell wall metabolism during maturation, ripening and senescence of peach fruit. J. Exp. Bot. 2004,55:2029-2039.
42. Bandaranayake P.C.G., Filappova T., Tomilov A., Tomilova N. B., Jamison-McClung D., Ngo Q., Inoue K., Yoder J.I. A single-electron reducing quinone oxidoreductase is necessary to induce haustorium development in the roor parasitic plant *Triphysaria*. The Plant Cell 2010,22:1404-1419.
43. Chung H. J., Ferl R.J. Arabidopsis alcohol dehydrogenase expression in both shoots and roots is conditionated by root growth environment. Plant Phys. 1999,121: 429-436.
44. Lee M. H., Jeong J.H., Seo J.W., Shin C.G., Kim Y.S., In J.G., Yang D.C., Yi J.S., Choi Y.E. Enhanced triterpene and phytosterol biosynthesis in Pnax ginseng overexpressing squalene synthase gene. Plant Cell Physiol. 2004,45:976-984.
45. Drescher A, Ruf S, Calsa T Jr, Carrer H, Bock R. The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. Plant J. 2000,22:97-104.
46. Kikuchi S, Bédard J, Hirano M, Hirabayashi Y, Oishi M, Imai M, Takase M, Ide T, Nakai M. Uncovering the protein translocon at the chloroplast inner envelope membrane. Science. 2013, 339(6119):571-574.
47. Kim K, Lee SC, Lee J, Kim NH, Jang W, Yang TJ. The complete chloroplast genome sequence of *Panax quinquefolius* (L.). Mitochondrial DNA. 2015, 10: 1-2.

48. Jayakodi M, Lee SC, Lee YS, Park HS, Kim NH, Jang W, Lee HO, Joh HJ, Yang TJ. Comprehensive analysis of Panax ginseng root transcriptomes. BMC Plant Biol. 2015,15:138.
49. Kim SJ, Kim AK. Anti-breast cancer activity of Fine Black ginseng (Panax ginseng Meyer) and ginsenoside Rg5. J Ginseng Res. 2015, 39:125-134.
50. [50] Park JY, Choi P, Kim T, Ko H, Kim HK, Kang KS, Ham J. Protective Effects of Processed Ginseng and Its Active Ginsenosides on Cisplatin-Induced Nephrotoxicity: In Vitro and in Vivo Studies. J Agric Food Chem. 2015, 63:5964-9.
51. Jun YL, Bae CH, Kim D, Koo S, Kim S. Korean Red Ginseng protects dopaminergic neurons by suppressing the cleavage of p35 to p25 in a Parkinson's disease mouse model. J Ginseng Res. 2015, 39:148-154.
52. Shin HS, Yu M, Kim M, Choi HS, Kang DH. Renoprotective effect of red ginseng in gentamicin-induced acute kidney injury. Lab Invest. 2014,94:1147-1160.
53. Singh P, Kim YJ, Wang C, Mathiyalagan R, El-Agamy Farh M, Yang DC. Biogenic silver and gold nanoparticles synthesized using red ginseng root extract, and their applications. Artif Cells Nanomed Biotechnol. 2015, 23:1-6.
54. Singh P, Kim YJ, Wang C, Mathiyalagan R, Yang DC. The development of a green approach for the biosynthesis of silver and gold nanoparticles by using Panaxginseng root extract, and their biological applications. Artif Cells Nanomed Biotechnol. 2015, 14:1-8.

Supplementary data

Table S1: Lists of the 152 unique gene products identified by the use of Arabidopsis thaliana database, of 55 proteins identified by the use of Panax ginseng database and of 207 unique identifications considering both searches.

	Description	GENE_ID
1	PR10-2 OS=Panax ginseng PE=3 SV=1 - [E9M221_PANGI]	E9M221_PANGI
2	PR10-3 (Fragment) OS=Panax ginseng PE=2 SV=1 - [E9M220_PANGI]	E9M220_PANGI
3	Granule-bound starch synthase I (Fragment) OS=Panax zingiberensis GN=GBSSI PE=4 SV=1 - [Q6XY46_9APIA]	Q6XY46_9APIA
4	Cytoplasmic ribosomal protein S13 OS=Panax ginseng PE=2 SV=1 - [Q9MAV9_PANGI]	Q9MAV9_PANGI
5	Putative uncharacterized protein OS=Panax ginseng PE=2 SV=1 - [Q20BN0_PANGI]	Q20BN0_PANGI
6	RNase-like major storage protein OS=Panax notoginseng GN=PMP PE=2 SV=1 - [X2D3J9_9APIA]	X2D3J9_9APIA
7	60S ribosomal protein L17-like protein (Fragment) OS=Panax ginseng PE=2 SV=1 - [A9QMA3_PANGI]	A9QMA3_PANGI
8	Dehydrin 4 OS=Panax ginseng GN=Dhn4 PE=2 SV=1 - [Q1HGF4_PANGI]	Q1HGF4_PANGI
9	60S ribosomal protein L27a OS=Panax ginseng PE=2 SV=1 - [Q9MAW6_PANGI]	Q9MAW6_PANGI
10	Glutamine synthetase-like (Fragment) OS=Panax quinquefolius PE=2 SV=1 - [A9QMA5_PANQU]	A9QMA5_PANQU
11	Ribosomal protein S4 OS=Panax ginseng PE=2 SV=1 - [Q401B8_PANGI]	Q401B8_PANGI
12	ADP-ribosylation factor-like protein OS=Panax ginseng PE=2 SV=1 - [Q20BN2_PANGI]	Q20BN2_PANGI
13	Dehydrin 2 OS=Panax ginseng GN=Dhn2 PE=4 SV=1 - [Q1HGF6_PANGI]	Q1HGF6_PANGI
14	Pathogenesis-related protein 4 OS=Panax ginseng PE=2 SV=1 - [W8FMB6_PANGI]	W8FMB6_PANGI
15	Sesquiterpene synthase OS=Panax ginseng GN=STS PE=2 SV=1 - [S5RMQ6_PANGI]	S5RMQ6_PANGI
16	ATP synthase subunit alpha OS=Panax ginseng GN=atpA PE=3 SV=1 - [O47412_PANGI]	O47412_PANGI
17	Glutamate decarboxylase OS=Panax ginseng GN=GAD PE=2 SV=1 - [D3JX88_PANGI]	D3JX88_PANGI
18	Glutathione peroxidase OS=Panax ginseng PE=2 SV=1 - [U5NNF1_PANGI]	U5NNF1_PANGI
19	Dehydrin 1 OS=Panax ginseng GN=Dhn1 PE=2 SV=1 - [Q1HGF7_PANGI]	Q1HGF7_PANGI

20	ATP-dependent Clp protease proteolytic subunit OS=Panax ginseng GN=clpP PE=3 SV=1 - [CLPP_PANGI]	CLPP_PANGI
21	Class 1 chitinase OS=Panax ginseng GN=Chi-1 PE=2 SV=1 - [C1K2M3_PANGI]	C1K2M3_PANGI
22	60S ribosomal protein L31 OS=Panax ginseng GN=RPL31 PE=2 SV=1 - [RL31_PANGI]	RL31_PANGI
23	Glutathione peroxidase OS=Panax ginseng PE=2 SV=1 - [U5NNJ1_PANGI]	U5NNJ1_PANGI
24	Glycosyltransferase OS=Panax notoginseng PE=4 SV=1 - [F4YF65_9APIA]	F4YF65_9APIA
25	Catalase OS=Panax ginseng GN=Cat1 PE=2 SV=1 - [A9Z0Q0_PANGI]	A9Z0Q0_PANGI
26	Ribosomal protein L31 OS=Panax ginseng PE=2 SV=1 - [Q20BM6_PANGI]	Q20BM6_PANGI
27	Thioredoxin (Fragment) OS=Panax ginseng PE=2 SV=1 - [A9QMA1_PANGI]	A9QMA1_PANGI
28	ATP synthase subunit beta, chloroplastic OS=Panax ginseng GN=atpB PE=3 SV=1 - [ATPB_PANGI]	ATPB_PANGI
29	PgMADS protein1 OS=Panax ginseng GN=PgMADS1 PE=2 SV=1 - [F2ZBV8_PANGI]	F2ZBV8_PANGI
30	ATP synthase subunit alpha, chloroplastic OS=Panax ginseng GN=atpA PE=3 SV=1 - [ATPA_PANGI]	ATPA_PANGI
31	Cytochrome P450 OS=Panax notoginseng PE=2 SV=1 - [F4YF81_9APIA]	F4YF81_9APIA
32	Cytochrome P450 OS=Panax notoginseng PE=2 SV=1 - [F4YF78_9APIA]	F4YF78_9APIA
33	Peroxidase OS=Panax ginseng PE=2 SV=1 - [Q401B7_PANGI]	Q401B7_PANGI
34	Cytochrome P450 CYP73A100 OS=Panax ginseng PE=2 SV=1 - [C7A10_PANGI]	C7A10_PANGI
35	Cytochrome P450 (Fragment) OS=Panax notoginseng PE=2 SV=1 - [F4YF83_9APIA]	F4YF83_9APIA
36	Glycosyltransferase OS=Panax notoginseng PE=2 SV=1 - [F4YF66_9APIA]	F4YF66_9APIA
37	Farnesyl diphosphate synthase OS=Panax quinquefolius GN=FPS PE=2 SV=1 - [G0T3G2_PANQU]	G0T3G2_PANQU
38	Putative membrane protein ycf1 OS=Panax ginseng GN=ycf1 PE=3 SV=1 - [YCF1_PANGI]	YCF1_PANGI
39	Pleiotropic drug resistance transporter 3 OS=Panax ginseng GN=PDR3 PE=2 SV=1 - [U5KP94_PANGI]	U5KP94_PANGI
40	Protein Ycf2 OS=Panax ginseng GN=ycf2-A PE=3 SV=1 - [YCF2_PANGI]	YCF2_PANGI
41	DNA-directed RNA polymerase subunit beta' OS=Panax ginseng GN=rpoC1 PE=3 SV=2 - [RPOC1_PANGI]	RPOC1_PANGI
42	Glyceraldehyde-3-phosphate dehydrogenase (Fragment) OS=Panax ginseng PE=3 SV=1 - [D0VFU1_PANGI]	D0VFU1_PANGI
43	Glyceraldehyde-3-phosphate dehydrogenase (Fragment) OS=Panax notoginseng PE=2 SV=1 - [Q3LRW8_9APIA]	Q3LRW8_9APIA

44	Glyceraldehyde-3-phosphate dehydrogenase OS=Panax ginseng PE=2 SV=1 - [Q6VAL5_PANGI]	Q6VAL5_PANGI
45	Major latex-like protein OS=Panax ginseng GN=mlp151 PE=2 SV=1 - [B5THI3_PANGI]	B5THI3_PANGI
46	Heat shock protein 90-like protein (Fragment) OS=Panax quinquefolius PE=2 SV=1 - [A9LIX4_PANQU]	[A9LIX4_PANQU]
47	Superoxide dismutase [Cu-Zn] OS=Panax ginseng GN=SODCC PE=2 SV=1 - [SODC_PANGI]	SODC_PANGI
48	Peroxiredoxin OS=Panax ginseng PE=2 SV=1 - [D7RWP9_PANGI]	D7RWP9_PANGI
49	Ribonuclease-like storage protein OS=Panax ginseng PE=1 SV=2 - [RN28_PANGI]	RN28_PANGI
50	Ribonuclease 1 OS=Panax ginseng PE=1 SV=1 - [RNS1_PANGI]	RNS1_PANGI
51	Ribonuclease 2 OS=Panax ginseng PE=1 SV=1 - [RNS2_PANGI]	RNS2_PANGI
52	S-adenosylmethionine synthase OS=Panax ginseng PE=2 SV=1 - [B3GLJ1_PANGI]	B3GLJ1_PANGI
53	Cyclophilin ABH-like protein OS=Panax ginseng PE=2 SV=1 - [A5YU16_PANGI]	A5YU16_PANGI
54	Short-chain alcohol dehydrogenase OS=Panax ginseng PE=1 SV=1 - [B8YDG5_PANGI]	B8YDG5_PANGI
55	Histone H4 OS=Arabidopsis thaliana GN=At1g07660 PE=1 SV=2 - [H4_ARATH]	H4_ARATH
56	Polyubiquitin 10 OS=Arabidopsis thaliana GN=UBQ10 PE=1 SV=2 - [UBQ10_ARATH]	UBQ10_ARATH
57	60S ribosomal protein L23 OS=Arabidopsis thaliana GN=emb2171 PE=2 SV=1 - [F4J3P1_ARATH]	F4J3P1_ARATH
58	Histone H2B.1 OS=Arabidopsis thaliana GN=At1g07790 PE=1 SV=3 - [H2B1_ARATH]	H2B1_ARATH
59	40S ribosomal protein S13-1 OS=Arabidopsis thaliana GN=RPS13A PE=2 SV=1 - [RS131_ARATH]	RS131_ARATH
60	40S ribosomal protein S19-3 OS=Arabidopsis thaliana GN=RPS19C PE=2 SV=1 - [RS193_ARATH]	RS193_ARATH
61	Histone H3.2 OS=Arabidopsis thaliana GN=HTR2 PE=1 SV=2 - [H32_ARATH]	H32_ARATH
62	GTP-binding protein SAR1B OS=Arabidopsis thaliana GN=SAR1B PE=2 SV=1 - [SAR1B_ARATH]	SAR1B_ARATH
63	Putative Sar1 GTP binding protein OS=Arabidopsis thaliana GN=AT3G62560 PE=2 SV=1 - [Q8VYP7_ARATH]	Q8VYP7_ARATH
64	40S ribosomal protein S14-2 OS=Arabidopsis thaliana GN=RPS14B PE=1 SV=1 - [RS142_ARATH]	RS142_ARATH
65	40S ribosomal protein S17-2 OS=Arabidopsis thaliana GN=RPS17B PE=2 SV=3 - [RS172_ARATH]	RS172_ARATH
66	Histone H2B.4 OS=Arabidopsis thaliana GN=At2g37470 PE=1 SV=3 - [H2B4_ARATH]	H2B4_ARATH
67	Probable mediator of RNA polymerase II transcription subunit 37c OS=Arabidopsis thaliana GN=MED37C PE=1 SV=1 - [MD37C_ARATH]	MD37C_ARATH

68	60S ribosomal protein L23a-2 OS=Arabidopsis thaliana GN=RPL23AB PE=3 SV=1 - [A8MS83_ARATH]	A8MS83_ARATH
69	Probable histone H2A variant 3 OS=Arabidopsis thaliana GN=At1g52740 PE=1 SV=1 - [H2AV3_ARATH]	H2AV3_ARATH
70	40S ribosomal protein S9-1 OS=Arabidopsis thaliana GN=AT5G15200 PE=3 SV=1 - [B3H7J6_ARATH]	B3H7J6_ARATH
71	Probable histone H2AXa OS=Arabidopsis thaliana GN=At1g08880 PE=1 SV=1 - [H2AXA_ARATH]	H2AXA_ARATH
72	40S ribosomal protein S15a-1 OS=Arabidopsis thaliana GN=RPS15AA PE=2 SV=2 - [R15A1_ARATH]	R15A1_ARATH
73	Histone H2A.6 OS=Arabidopsis thaliana GN=RAT5 PE=1 SV=1 - [H2A6_ARATH]	H2A6_ARATH
74	40S ribosomal protein S5-2 OS=Arabidopsis thaliana GN=RPS5B PE=2 SV=2 - [RS52_ARATH]	RS52_ARATH
75	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3 OS=Arabidopsis thaliana GN=NAD9 PE=1 SV=2 - [NDUS3_ARATH]	NDUS3_ARATH
76	Glyceraldehyde-3-phosphate dehydrogenase GAPC1, cytosolic OS=Arabidopsis thaliana GN=GAPC1 PE=1 SV=2 - [G3PC1_ARATH]	G3PC1_ARATH
77	Eukaryotic initiation factor 4A-3 OS=Arabidopsis thaliana GN=TIF4A-3 PE=1 SV=1 - [IF4A3_ARATH]	IF4A3_ARATH
78	Probable mediator of RNA polymerase II transcription subunit 37c OS=Arabidopsis thaliana GN=MED37D PE=1 SV=2 - [MD37D_ARATH]	MD37D_ARATH
79	Phosphoglycerate kinase 1, chloroplastic OS=Arabidopsis thaliana GN=PGK1 PE=1 SV=1 - [PGKH1_ARATH]	PGKH1_ARATH
80	Mediator of RNA polymerase II transcription subunit 37a OS=Arabidopsis thaliana GN=MED37A PE=1 SV=1 - [MD37A_ARATH]	MD37A_ARATH
81	Probable mediator of RNA polymerase II transcription subunit 37e OS=Arabidopsis thaliana GN=MED37E PE=1 SV=3 - [MD37E_ARATH]	MD37E_ARATH
82	Histone H2B.9 OS=Arabidopsis thaliana GN=At5g02570 PE=1 SV=3 - [H2B9_ARATH]	H2B9_ARATH
83	60S ribosomal protein L30-2 OS=Arabidopsis thaliana GN=RPL30B PE=3 SV=1 - [RL302_ARATH]	RL302_ARATH
84	Heat shock 70 kDa protein 3 OS=Arabidopsis thaliana GN=HSP70-3 PE=1 SV=1 - [HSP7C_ARATH]	HSP7C_ARATH
85	Probable histone H2A.4 OS=Arabidopsis thaliana GN=At5g02560 PE=1 SV=1 - [H2A4_ARATH]	H2A4_ARATH
86	60S ribosomal protein L11-2 OS=Arabidopsis thaliana GN=RPL11B PE=2 SV=2 - [RL112_ARATH]	RL112_ARATH
87	60S ribosomal protein L18-2 OS=Arabidopsis thaliana GN=RPL18B PE=1 SV=2 - [RL182_ARATH]	RL182_ARATH

88	ADP-ribosylation factor 1 OS=Arabidopsis thaliana GN=ARF1 PE=1 SV=2 - [ARF1_ARATH]	ARF1_ARATH
89	AAA-type ATPase family protein OS=Arabidopsis thaliana GN=AT1G45000 PE=4 SV=1 - [B3H533_ARATH]	B3H533_ARATH
90	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase 1 OS=Arabidopsis thaliana GN=MS1 PE=1 SV=1 - [METE1_ARATH]	METE1_ARATH
91	Elongation factor 1-alpha OS=Arabidopsis thaliana GN=AT5G60390 PE=3 SV=1 - [A8MSE8_ARATH]	A8MSE8_ARATH
92	40S ribosomal protein S3-3 OS=Arabidopsis thaliana GN=RPS3C PE=1 SV=1 - [RS33_ARATH]	RS33_ARATH
93	Heat shock protein 90-1 OS=Arabidopsis thaliana GN=HSP90-1 PE=1 SV=3 - [HS901_ARATH]	HS901_ARATH
94	Tubulin beta-4 chain OS=Arabidopsis thaliana GN=TUBB4 PE=2 SV=2 - [TBB4_ARATH]	TBB4_ARATH
95	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase 2 OS=Arabidopsis thaliana GN=MS2 PE=1 SV=1 - [METE2_ARATH]	METE2_ARATH
96	40S ribosomal protein S18 OS=Arabidopsis thaliana GN=RPS18A PE=1 SV=1 - [RS18_ARATH]	RS18_ARATH
97	Phosphoglycerate kinase OS=Arabidopsis thaliana GN=T8K14.3 PE=2 SV=1 - [Q9SAJ4_ARATH]	Q9SAJ4_ARATH
98	Heat shock 70 kDa protein 5 OS=Arabidopsis thaliana GN=HSP70-5 PE=2 SV=1 - [HSP7E_ARATH]	HSP7E_ARATH
99	Adenosylhomocysteinase 2 OS=Arabidopsis thaliana GN=SAHH2 PE=1 SV=1 - [SAHH2_ARATH]	SAHH2_ARATH
100	20S proteasome alpha subunit PAD1 OS=Arabidopsis thaliana GN=PAD1 PE=4 SV=1 - [Q2V3Q0_ARATH]	Q2V3Q0_ARATH
101	Catalase OS=Arabidopsis thaliana GN=CAT2 PE=2 SV=1 - [F4JM86_ARATH]	F4JM86_ARATH
102	Chaperone protein ClpB1 OS=Arabidopsis thaliana GN=CLPB1 PE=1 SV=2 - [CLPB1_ARATH]	CLPB1_ARATH
103	40S ribosomal protein S23-2 OS=Arabidopsis thaliana GN=RPS23B PE=2 SV=2 - [RS232_ARATH]	RS232_ARATH
104	ATP synthase subunit alpha, mitochondrial OS=Arabidopsis thaliana GN=ATPA PE=1 SV=2 - [ATPAM_ARATH]	ATPAM_ARATH
105	Cell division control protein 48 homolog A OS=Arabidopsis thaliana GN=CDC48A PE=1 SV=1 - [CD48A_ARATH]	CD48A_ARATH
106	Pyruvate kinase OS=Arabidopsis thaliana GN=AT3G52990 PE=2 SV=1 - [A8MR07_ARATH]	A8MR07_ARATH
107	60S ribosomal protein L7-3 OS=Arabidopsis thaliana GN=RPL7C PE=2 SV=1 - [RL73_ARATH]	RL73_ARATH
108	Heat shock protein 90-4 OS=Arabidopsis thaliana GN=HSP90-4 PE=2 SV=1 - [HS904_ARATH]	HS904_ARATH
109	60S ribosomal protein L13-3 OS=Arabidopsis thaliana GN=RPL13D PE=2 SV=1 - [RL133_ARATH]	RL133_ARATH

110	18.1 kDa class I heat shock protein OS=Arabidopsis thaliana GN=HSP18.1 PE=2 SV=1 - [HS181_ARATH]	HS181_ARATH
111	60S ribosomal protein L13 OS=Arabidopsis thaliana GN=BBC1 PE=3 SV=1 - [A8MQA1_ARATH]	A8MQA1_ARATH
112	40S ribosomal protein S6-2 OS=Arabidopsis thaliana GN=EMB3010 PE=2 SV=1 - [F4KGU2_ARATH]	F4KGU2_ARATH
113	60S ribosomal protein L26-2 OS=Arabidopsis thaliana GN=RPL26B PE=2 SV=1 - [RL262_ARATH]	RL262_ARATH
114	Cytosolic isocitrate dehydrogenase [NADP] OS=Arabidopsis thaliana GN=CICDH PE=2 SV=1 - [ICDHC_ARATH]	ICDHC_ARATH
115	Peptidyl-prolyl cis-trans isomerase OS=Arabidopsis thaliana GN=CYP5 PE=1 SV=1 - [A8MS66_ARATH]	A8MS66_ARATH
116	Peroxidase 72 OS=Arabidopsis thaliana GN=PER72 PE=1 SV=1 - [PER72_ARATH]	PER72_ARATH
117	Ras-related protein RABB1c OS=Arabidopsis thaliana GN=RABB1C PE=2 SV=1 - [RAB1C_ARATH]	RAB1C_ARATH
118	6-phosphogluconate dehydrogenase, decarboxylating 1, chloroplastic OS=Arabidopsis thaliana GN=At1g64190 PE=2 SV=1 - [6PGD1_ARATH]	6PGD1_ARATH
119	40S ribosomal protein S16-3 OS=Arabidopsis thaliana GN=AT5G18380 PE=3 SV=1 - [A8MRX2_ARATH]	A8MRX2_ARATH
120	6-phosphogluconate dehydrogenase, decarboxylating 3 OS=Arabidopsis thaliana GN=At3g02360 PE=2 SV=1 - [6GPD3_ARATH]	6GPD3_ARATH
121	Serine hydroxymethyltransferase 4 OS=Arabidopsis thaliana GN=SHM4 PE=1 SV=1 - [GLYC4_ARATH]	GLYC4_ARATH
122	Probable fructokinase-1 OS=Arabidopsis thaliana GN=At2g31390 PE=2 SV=1 - [SCRK1_ARATH]	SCRK1_ARATH
123	26S proteasome non-ATPase regulatory subunit 14 homolog OS=Arabidopsis thaliana GN=RPN11 PE=1 SV=1 - [PSDE_ARATH]	PSDE_ARATH
124	ATP-citrate synthase beta chain protein 1 OS=Arabidopsis thaliana GN=ACLB-1 PE=2 SV=1 - [ACLB1_ARATH]	ACLB1_ARATH
125	Proteasome subunit beta type-2-A OS=Arabidopsis thaliana GN=PBD1 PE=1 SV=1 - [PSB2A_ARATH]	PSB2A_ARATH
126	V-type proton ATPase catalytic subunit A OS=Arabidopsis thaliana GN=VHA-A PE=1 SV=1 - [VATA_ARATH]	VATA_ARATH
127	ATP synthase subunit alpha OS=Arabidopsis thaliana GN=AT2G07698 PE=2 SV=1 - [F4IMB5_ARATH]	F4IMB5_ARATH
128	40S ribosomal protein S25-3 OS=Arabidopsis thaliana GN=RPS25D PE=3 SV=2 - [RS253_ARATH]	RS253_ARATH
129	26S proteasome regulatory subunit 4 homolog B OS=Arabidopsis thaliana GN=RPT2B PE=1 SV=1 - [PRS4B_ARATH]	PRS4B_ARATH

130	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A beta isoform OS=Arabidopsis thaliana GN=PP2AA2 PE=1 SV=2 - [2AAB_ARATH]	2AAB_ARATH
131	40S ribosomal protein S20-2 OS=Arabidopsis thaliana GN=RPS20B PE=2 SV=1 - [RS202_ARATH]	RS202_ARATH
132	Heat shock 70 kDa protein 7, chloroplastic OS=Arabidopsis thaliana GN=HSP70-7 PE=1 SV=1 - [HSP7G_ARATH]	HSP7G_ARATH
133	60S ribosomal protein L7a-2 OS=Arabidopsis thaliana GN=RPL7AB PE=1 SV=1 - [RL7A2_ARATH]	RL7A2_ARATH
134	40S ribosomal protein S3a-2 OS=Arabidopsis thaliana GN=RPS3AB PE=2 SV=3 - [RS3A2_ARATH]	RS3A2_ARATH
135	Glutamate decarboxylase 4 OS=Arabidopsis thaliana GN=GAD4 PE=2 SV=1 - [DCE4_ARATH]	DCE4_ARATH
136	S-adenosylmethionine synthase 2 OS=Arabidopsis thaliana GN=SAM2 PE=1 SV=1 - [METK2_ARATH]	METK2_ARATH
137	Phosphoenolpyruvate carboxylase 1 OS=Arabidopsis thaliana GN=PPC1 PE=1 SV=1 - [CAPP1_ARATH]	CAPP1_ARATH
138	UDP-arabinopyranose mutase 1 OS=Arabidopsis thaliana GN=RGP1 PE=1 SV=1 - [RGP1_ARATH]	RGP1_ARATH
139	14-3-3-like protein GF14 omega OS=Arabidopsis thaliana GN=GRF2 PE=2 SV=2 - [14332_ARATH]	14332_ARATH
140	Glutamine synthetase cytosolic isozyme 1-4 OS=Arabidopsis thaliana GN=GLN1-4 PE=1 SV=1 - [GLN14_ARATH]	GLN14_ARATH
141	UDP-arabinopyranose mutase 3 OS=Arabidopsis thaliana GN=RGP3 PE=1 SV=2 - [RGP3_ARATH]	RGP3_ARATH
142	Ribosomal protein S5/Elongation factor G/III/V family protein OS=Arabidopsis thaliana GN=AT3G12915 PE=4 SV=1 - [F4JB05_ARATH]	F4JB05_ARATH
143	AT2G44100 protein OS=Arabidopsis thaliana GN=GDI1 PE=2 SV=1 - [B9DFX1_ARATH]	B9DFX1_ARATH
144	Citrate synthase 3, peroxisomal OS=Arabidopsis thaliana GN=CSY3 PE=2 SV=1 - [CISY3_ARATH]	CISY3_ARATH
145	UDP-glucose 6-dehydrogenase 2 OS=Arabidopsis thaliana GN=UGD2 PE=1 SV=1 - [UGDH2_ARATH]	UGDH2_ARATH
146	Proteasome subunit beta type-5-B OS=Arabidopsis thaliana GN=PBE2 PE=1 SV=1 - [PSB5B_ARATH]	PSB5B_ARATH
147	Uncharacterized protein OS=Arabidopsis thaliana GN=AT1G47980 PE=2 SV=1 - [F4HWR0_ARATH]	F4HWR0_ARATH
148	At1g30580 OS=Arabidopsis thaliana GN=T5I8.3 PE=2 SV=1 - [Q9SA73_ARATH]	Q9SA73_ARATH
149	26S protease regulatory subunit 6B homolog OS=Arabidopsis thaliana GN=RPT3 PE=1 SV=1 - [PRS6B_ARATH]	PRS6B_ARATH
150	HSP90-like protein GRP94 OS=Arabidopsis thaliana GN=SHD PE=2 SV=1 - [F4JQ55_ARATH]	F4JQ55_ARATH

151	GDP-mannose 3,5-epimerase OS=Arabidopsis thaliana GN=At5g28840 PE=1 SV=1 - [GME_ARATH]	GME_ARATH
152	V-type proton ATPase subunit B1 OS=Arabidopsis thaliana GN=VHA-B1 PE=2 SV=2 - [VATB1_ARATH]	VATB1_ARATH
153	40S ribosomal protein S4-2 OS=Arabidopsis thaliana GN=AT5G07090 PE=2 SV=1 - [F4K5C7_ARATH]	F4K5C7_ARATH
154	Peroxisomal isocitrate dehydrogenase [NADP] OS=Arabidopsis thaliana GN=ICDH PE=1 SV=1 - [ICDHX_ARATH]	ICDHX_ARATH
155	ADP,ATP carrier protein 3, mitochondrial OS=Arabidopsis thaliana GN=AAC3 PE=1 SV=1 - [ADT3_ARATH]	ADT3_ARATH
156	NAD-dependent malic enzyme 2, mitochondrial OS=Arabidopsis thaliana GN=NAD-ME2 PE=1 SV=1 - [MAO2_ARATH]	MAO2_ARATH
157	Cysteine synthase, chloroplastic/chromoplastic OS=Arabidopsis thaliana GN=OASB PE=1 SV=2 - [CYSKP_ARATH]	CYSKP_ARATH
158	Clp ATPase OS=Arabidopsis thaliana GN=HSP93-III PE=3 SV=1 - [F4JF64_ARATH]	F4JF64_ARATH
159	Guanine nucleotide-binding protein subunit beta-like protein B OS=Arabidopsis thaliana GN=RACK1B PE=2 SV=1 - [GPLPB_ARATH]	GPLPB_ARATH
160	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial OS=Arabidopsis thaliana GN=At2g20360 PE=2 SV=2 - [NDUA9_ARATH]	NDUA9_ARATH
161	TCP-1/cpn60 chaperonin family protein OS=Arabidopsis thaliana GN=AT3G11830 PE=2 SV=1 - [F4J7H2_ARATH]	F4J7H2_ARATH
162	Heat shock protein 60-2 OS=Arabidopsis thaliana GN=HSP60-2 PE=2 SV=1 - [F4IVR2_ARATH]	F4IVR2_ARATH
163	Glutamate decarboxylase 2 OS=Arabidopsis thaliana GN=At1g65960 PE=2 SV=1 - [Q56W28_ARATH]	Q56W28_ARATH
164	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform OS=Arabidopsis thaliana GN=PP2AA1 PE=1 SV=1 - [2AAA_ARATH]	2AAA_ARATH
165	Probable phosphoglucomutase, cytoplasmic 1 OS=Arabidopsis thaliana GN=At1g23190 PE=2 SV=2 - [PGMC1_ARATH]	PGMC1_ARATH
166	Succinate-semialdehyde dehydrogenase, mitochondrial OS=Arabidopsis thaliana GN=ALDH5F1 PE=1 SV=2 - [SSDH_ARATH]	SSDH_ARATH
167	60S ribosomal protein L4-1 OS=Arabidopsis thaliana GN=AT3G09630 PE=4 SV=1 - [Q2V3X4_ARATH]	Q2V3X4_ARATH

168	Alpha-glucan phosphorylase 1 OS=Arabidopsis thaliana GN=PHS1 PE=1 SV=1 - [PHS1_ARATH]	PHS1_ARATH
169	40S ribosomal protein S2-4 OS=Arabidopsis thaliana GN=RPS2D PE=2 SV=1 - [RS24_ARATH]	RS24_ARATH
170	60S ribosomal protein L5-2 OS=Arabidopsis thaliana GN=RPL5B PE=2 SV=3 - [RL52_ARATH]	RL52_ARATH
171	Sucrose synthase 3 OS=Arabidopsis thaliana GN=SUS3 PE=1 SV=1 - [SUS3_ARATH]	SUS3_ARATH
172	Polyadenylate-binding protein 8 OS=Arabidopsis thaliana GN=PAB8 PE=1 SV=1 - [PABP8_ARATH]	PABP8_ARATH
173	Phenylalanine ammonia-lyase 3 OS=Arabidopsis thaliana GN=PAL3 PE=1 SV=3 - [PAL3_ARATH]	PAL3_ARATH
174	Aconitate hydratase 1 OS=Arabidopsis thaliana GN=ACO1 PE=1 SV=2 - [ACO1_ARATH]	ACO1_ARATH
175	26S proteasome non-ATPase regulatory subunit 2 homolog A OS=Arabidopsis thaliana GN=RPN1A PE=1 SV=2 - [PSD2A_ARATH]	PSD2A_ARATH
176	Aldolase-type TIM barrel family protein OS=Arabidopsis thaliana GN=T22N19_70 PE=2 SV=1 - [Q9LYR4_ARATH]	Q9LYR4_ARATH
177	Aminopeptidase family protein OS=Arabidopsis thaliana GN=AT2G24200 PE=2 SV=1 - [F4INR3_ARATH]	F4INR3_ARATH
178	Pyruvate kinase OS=Arabidopsis thaliana GN=AT5G63680 PE=3 SV=1 - [Q9FFP6_ARATH]	Q9FFP6_ARATH
179	Tubulin alpha-3 chain OS=Arabidopsis thaliana GN=TUBA3 PE=2 SV=2 - [TBA3_ARATH]	TBA3_ARATH
180	UDP-glucose 6-dehydrogenase 3 OS=Arabidopsis thaliana GN=UGD3 PE=1 SV=1 - [UGDH3_ARATH]	UGDH3_ARATH
181	60S acidic ribosomal protein P0-1 OS=Arabidopsis thaliana GN=RPP0A PE=1 SV=1 - [RLA01_ARATH]	RLA01_ARATH
182	NADP-dependent malic enzyme 1 OS=Arabidopsis thaliana GN=NADP-ME1 PE=1 SV=1 - [MAOP1_ARATH]	MAOP1_ARATH
183	Clathrin heavy chain 2 OS=Arabidopsis thaliana GN=CHC2 PE=1 SV=1 - [CLAH2_ARATH]	CLAH2_ARATH
184	Isocitrate dehydrogenase [NAD] regulatory subunit 3, mitochondrial OS=Arabidopsis thaliana GN=IDH3 PE=1 SV=1 - [IDH3_ARATH]	IDH3_ARATH
185	Glucose-6-phosphate isomerase, cytosolic OS=Arabidopsis thaliana GN=PGIC PE=2 SV=1 - [G6PI_ARATH]	G6PI_ARATH
186	Glycine--tRNA ligase 1, mitochondrial OS=Arabidopsis thaliana GN=GLYRS-1 PE=2 SV=1 - [SYGM1_ARATH]	SYGM1_ARATH
187	HEAT SHOCK PROTEIN 89.1 OS=Arabidopsis thaliana GN=Hsp89.1 PE=2 SV=1 - [F4JFN3_ARATH]	F4JFN3_ARATH
188	Phosphoglucomutase, chloroplastic OS=Arabidopsis thaliana GN=PGMP PE=1 SV=2 - [PGMP_ARATH]	PGMP_ARATH

189	Pectinesterase/pectinesterase inhibitor 3 OS=Arabidopsis thaliana GN=PME3 PE=2 SV=2 - [PME3_ARATH]	PME3_ARATH
190	Heat shock 70 kDa protein 14 OS=Arabidopsis thaliana GN=HSP70-14 PE=2 SV=1 - [HSP7O_ARATH]	HSP7O_ARATH
191	At1g56070/T6H22_13 OS=Arabidopsis thaliana GN=At1g56075 PE=2 SV=1 - [Q9ASR1_ARATH]	Q9ASR1_ARATH
192	Alpha-glucan water dikinase 2 OS=Arabidopsis thaliana GN=GWD2 PE=2 SV=3 - [GWD2_ARATH]	GWD2_ARATH
193	Chaperonin 60 subunit alpha 1, chloroplastic OS=Arabidopsis thaliana GN=CPN60A1 PE=1 SV=2 - [CPNA1_ARATH]	CPNA1_ARATH
194	Chaperonin 60 subunit beta 2, chloroplastic OS=Arabidopsis thaliana GN=CPN60B2 PE=1 SV=1 - [CPNB2_ARATH]	CPNB2_ARATH
195	Chaperonin 60 subunit beta 3, chloroplastic OS=Arabidopsis thaliana GN=CPN60B3 PE=1 SV=1 - [CPNB3_ARATH]	CPNB3_ARATH
196	Malate dehydrogenase OS=Arabidopsis thaliana GN=mMDH2 PE=3 SV=1 - [A8MQK3_ARATH]	A8MQK3_ARATH
197	Prohibitin-6, mitochondrial OS=Arabidopsis thaliana GN=PHB6 PE=1 SV=1 - [PHB6_ARATH]	PHB6_ARATH
198	Malate dehydrogenase, cytoplasmic 1 OS=Arabidopsis thaliana GN=At1g04410 PE=1 SV=2 - [MDHC1_ARATH]	MDHC1_ARATH
199	Malate dehydrogenase, cytoplasmic 2 OS=Arabidopsis thaliana GN=At5g43330 PE=2 SV=1 - [MDHC2_ARATH]	MDHC2_ARATH
200	Malate dehydrogenase, chloroplastic OS=Arabidopsis thaliana GN=At3g47520 PE=1 SV=1 - [MDHP_ARATH]	MDHP_ARATH
201	Bifunctional enolase 2/transcriptional activator OS=Arabidopsis thaliana GN=ENO2 PE=1 SV=1 - [ENO2_ARATH]	ENO2_ARATH
202	ATP synthase subunit beta-1, mitochondrial OS=Arabidopsis thaliana GN=At5g08670 PE=1 SV=1 - [ATPBM_ARATH]	ATPBM_ARATH
203	Fructose-bisphosphate aldolase OS=Arabidopsis thaliana GN=AT2G36460 PE=2 SV=1 - [Q9SJQ9_ARATH]	Q9SJQ9_ARATH
204	Proteasome subunit alpha type-6-A OS=Arabidopsis thaliana GN=PAA1 PE=1 SV=2 - [PSA6A_ARATH]	PSA6A_ARATH
205	Actin-7 OS=Arabidopsis thaliana GN=ACT7 PE=1 SV=1 - [ACT7_ARATH]	ACT7_ARATH

206	NAD(P)H oxidoreductase, isoflavone reductase-like protein OS=Arabidopsis thaliana GN=T22F8.130 PE=2 SV=1 - [Q9T030_ARATH]	Q9T030_ARATH
207	Formate dehydrogenase, mitochondrial OS=Arabidopsis thaliana GN=FDH1 PE=1 SV=1 - [FDH_ARATH]	FDH_ARATH
208	CBS domain-containing protein CBSX3, mitochondrial OS=Arabidopsis thaliana GN=CBSX3 PE=1 SV=1 - [CBSX3_ARATH]	CBSX3_ARATH

Table S2: Lists of 95 potential bioactive peptides, generated by in-silico enzymatic digestion.

PeptideRanker Score	Peptides Sequence	Peptides Lenght	Class	AMP Probability (CAMP_Analysis)	Accession Number	Protein Description
0.93	**CCSEFGW**	**7**	NAMP	0.081	P19171	Basic endochitinase B
0.92	**DGQFRW**	**6**	NAMP	0.001	Q9M111	Sucrose synthase 3
0.89	**NGFCDRM**	**7**	NAMP	0.001	P92549	ATP synthase subunit alpha, mitochondrial
0.85	**NTWDRF**	**6**	NAMP	0.003	Q9STY6	40S ribosomal protein S20-2
0.85	**GWNRSW**	**6**	AMP	0.603	Q9STV0	Alpha-glucan water dikinase 2
0.83	**QNEWGW**	**6**	AMP	0.774	P25696	Bifunctional enolase 2/transcriptional activator
0.82	**QQCDRF**	**6**	NAMP	0.048	Q8L7K9	NAD-dependent malic enzyme 2, mitochondrial
0.82	**DNCCMQL**	**7**	NAMP	0.135	O81796	Isocitrate dehydrogenase [NAD] regulatory subunit 3, mitochondrial
0.81	**YQNRGCF**	**7**	NAMP	0.056	Q0WLB5	Clathrin heavy chain 2
0.81	**NSFRGDF**	**7**	NAMP	0.003	Q9LSY8	UDP-glycosyltransferase 71B2
0.81	**ICCCNKM**	**7**	AMP	0.699	A8MSE8	Elongation factor 1-alpha
0.8	**NGTCRWL**	**7**	NAMP	0.019	P17562	S-adenosylmethionine synthase 2
0.79	**EKNDFCF**	**7**	NAMP	0.208	P56785	Protein TIC 214
0.78	**GGCDKQGFP**	**9**	NAMP	0.168	P51430	40S ribosomal protein S6-2
0.78	**GGCDKQGF**	**8**	NAMP	0.041	P51430	40S ribosomal protein S6-2
0.75	**SKEQGFF**	**7**	NAMP	0.003	P56786	Protein Ycf2
0.75	**GGQRQFNGM**	**9**	NAMP	0.088	O49447	ADP,ATP carrier protein 3, mitochondrial
0.74	**QQCDRFI**	**7**	NAMP	0.054	Q8L7K9	NAD-dependent malic enzyme 2, mitochondrial

0.74	**DFDNRM**	**6**	NAMP	0.003	O65719	Heat shock 70 kDa protein 3
0.73	**RWCSCNI**	**7**	NAMP	0.128	Q9LK36	Adenosylhomocysteinase 2
0.73	**RPDRNF**	**6**	AMP	0.88	P56786	Protein Ycf2
0.73	**QNTWGQF**	**7**	NAMP	0.115	Q7PC86	ABC transporter G family member 35
0.73	**NQWKNW**	**6**	AMP	0.989	P56785	Protein TIC 214
0.73	**MCQGGDF**	**7**	NAMP	0.002	P34790	Peptidyl-prolyl cis-trans isomerase CYP18-3
0.73	**GQSKGYGF**	**8**	NAMP	0.201	Q9FXA2	Polyadenylate-binding protein 8
0.73	**DNFTKW**	**6**	NAMP	0.026	Q2V3X4	60S ribosomal protein L4-1
0.73	**CSCGDRPI**	**8**	NAMP	0.01	Q9MAH0	Phosphoenolpyruvate carboxylase 1
0.72	**QNRGCFN**	**7**	NAMP	0.497	Q0WLB5	Clathrin heavy chain 2
0.72	**QFNNGTW**	**7**	NAMP	0.029	Q7XA90	At1g14210
0.71	**NDNQLRF**	**7**	NAMP	0.28	Q9MAQ0	Probable granule-bound starch synthase 1, chloroplastic/amyloplastic
0.71	**DKSFDCL**	**7**	NAMP	0.005	Q84UU4	Alpha-humulene/(-)-(E)-beta-caryophyllene synthase
0.7	**CVQQCDRF**	**8**	NAMP	0.016	Q8L7K9	NAD-dependent malic enzyme 2, mitochondrial
0.69	**NGNDFRY**	**7**	NAMP	0.006	P92994	Trans-cinnamate 4-monooxygenase
0.69	**MGGCDKQGF**	**9**	NAMP	0.024	P51430	40S ribosomal protein S6-2
0.69	**GQMDRTF**	**7**	NAMP	0.006	P56786	Protein Ycf2
0.69	**FNKDTGF**	**7**	NAMP	0.01	Q8L7K9	NAD-dependent malic enzyme 2, mitochondrial
0.69	**DGSHNRW**	**7**	NAMP	0.001	Q9STV0	Alpha-glucan water dikinase 2
0.67	**NNCQSNCWGS**	**10**	NAMP	0.01	P43082	Hevein-like preproprotein
0.67	**GQTCDWI**	**7**	NAMP	0.003	Q0WLB5	Clathrin heavy chain 2
0.67	**CSPNEKW**	**7**	NAMP	0.003	Q0WLB5	Clathrin heavy chain 2

0.66	**NNCQSNCW**	**8**	NAMP	0.106	P43082	Hevein-like preproprotein
0.66	**GDLGCQGM**	**8**	NAMP	0.003	O82191	NADP-dependent malic enzyme 1
0.66	**EGRSGCGF**	**8**	NAMP	0.003	Q84UU4	Alpha-humulene/(-)-(E)-beta-caryophyllene synthase
0.66	**CDSQKGCCPP**	**10**	NAMP	0.42	Q7XA90	At1g14210
0.66	**CDSQKGCCP**	**9**	NAMP	0.454	Q7XA90	At1g14210
0.65	**PQTDRW**	**6**	NAMP	0.432	Q9LK36	Adenosylhomocysteinase 2
0.65	**DGTQNEWGW**	**9**	NAMP	0.186	P25696	Bifunctional enolase 2/transcriptional activator
0.65	**CCSQYGY**	**7**	NAMP	0.019	P43082	Hevein-like preproprotein
0.65	**CCCNKMD**	**7**	NAMP	0.133	A8MSE8	Elongation factor 1-alpha
0.64	**DQRQFL**	**6**	NAMP	0.116	P56786	Protein Ycf2
0.63	**RSGGHDF**	**7**	NAMP	0.002	Q9FKV0	Berberine bridge enzyme-like protein
0.63	**DDQGINF**	**7**	NAMP	0.059	Q9SID0	Probable fructokinase-1
0.63	**DCYNQRSF**	**8**	NAMP	0.003	P19171	Basic endochitinase B
0.62	**RGCFNEL**	**7**	NAMP	0	Q0WLB5	Clathrin heavy chain 2
0.62	**KSCWDK**	**6**	NAMP	0.049	Q7PC86	ABC transporter G family member 35
0.62	**ESCFRDP**	**7**	NAMP	0.001	P56786	Protein Ycf2
0.61	**DFGNENCL**	**8**	NAMP	0.012	Q9S7C0	Heat shock 70 kDa protein 14
0.6	**SEHGNRF**	**7**	NAMP	0.005	Q95748	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3
0.6	**GNQSFRI**	**7**	NAMP	0.04	Q0WRS4	Putative heat shock protein
0.59	**SQCRQM**	**6**	NAMP	0.006	Q9SAK4	Succinate-semialdehyde dehydrogenase, mitochondrial
0.59	**FDSRGNP**	**7**	NAMP	0.006	P25696	Bifunctional enolase 2/transcriptional activator
0.58	**GQPDSRNM**	**8**	NAMP	0.012	P43082	Hevein-like preproprotein

0.58	**FCDQTNL**	**7**	NAMP	0.205	Q7XA90	At1g14210
0.58	**EDFKCSW**	**7**	NAMP	0	Q43315	Farnesyl pyrophosphate synthase 2
0.58	**CCCNKMDA**	**8**	NAMP	0.256	A8MSE8	Elongation factor 1-alpha
0.57	**RGSGGNML**	**8**	NAMP	0.002	P56786	Protein Ycf2
0.57	**QQSDFR**	**6**	NAMP	0.05	P56785	Protein TIC 214
0.57	**NRCQSMM**	**7**	NAMP	0	P34795	Glucose-6-phosphate isomerase, cytosolic
0.57	**GQQCGRQGGG**	**10**	NAMP	0.002	P43082	Hevein-like preproprotein
0.57	**EGRSGCGFH**	**9**	NAMP	0.002	Q84UU4	Alpha-humulene/(-)-(E)-beta-caryophyllene synthase
0.57	**DSQKGCCPP**	**9**	NAMP	0.26	Q7XA90	At1g14210
0.56	**DSQKGCCP**	**8**	NAMP	0.15	Q7XA90	At1g14210
0.56	**DACDRNL**	**7**	NAMP	0.001	Q9C522	ATP-citrate synthase beta chain protein 1
0.55	**RNCDPNGP**	**8**	NAMP	0.003	Q9ASR1	elongation factor EF-2-like protein LOS1
0.55	**EFGQTCDW**	**8**	NAMP	0.001	Q0WRS4	Putative heat shock protein
0.55	**DCSQFDN**	**7**	NAMP	0.013	P56786	Protein Ycf2
0.54	**GGNMRDSF**	**8**	NAMP	0.003	Q9SF16	T-complex protein 1 subunit eta
0.54	**EHGNRF**	**6**	NAMP	0.005	Q95748	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3
0.53	**SGDGRYY**	**7**	NAMP	0	O49299	Probable phosphoglucomutase, cytoplasmic 1
0.53	**QCRQMSM**	**7**	NAMP	0.001	Q9SAK4	Succinate-semialdehyde dehydrogenase, mitochondrial
0.53	**GSEGFQKL**	**8**	NAMP	0.034	Q9C522	ATP-citrate synthase beta chain protein 1
0.53	**DNCRENMM**	**8**	NAMP	0	Q9ZPS3	Glutamate decarboxylase 4
0.53	**CTKDWY**	**6**	NAMP	0.002	Q9STV0	Alpha-glucan water dikinase 2
0.53	**AGQQCGRQGGG**	**11**	NAMP	0.005	P43082	Hevein-like preproprotein

0.52	**NGKNTCPL**	**8**	NAMP	0.012	O22850	Probable glutathione peroxidase 3, mitochondrial
0.52	**GGSCFQKDI**	**9**	NAMP	0.057	Q9LIA8	UDP-glucose 6-dehydrogenase 2
0.52	**ESRQWL**	**6**	NAMP	0.009	Q43315	Farnesyl pyrophosphate synthase 2
0.52	**CRNGEW**	**6**	NAMP	0	P57106	Malate dehydrogenase, cytoplasmic 2
0.51	**QFDTKM**	**6**	NAMP	0.006	Q9CAI7	Eukaryotic initiation factor 4A-3
0.51	**FCEQCGV**	**7**	NAMP	0.015	P56763	DNA-directed RNA polymerase subunit beta'
0.5	**QNRGCFNEL**	**9**	NAMP	0.039	Q0WLB5	Clathrin heavy chain 2
0.5	**KNNYDF**	**6**	AMP	0.905	P56785	Protein TIC 214
0.5	**GWCGNTEP**	**8**	NAMP	0.003	P19171	Basic endochitinase B
0.5	**DSCMERI**	**7**	NAMP	0	Q38876	Agamous-like MADS-box protein AGL8

Conclusions

The present study provides new insight into the proteome of two amazing sources of bioactive proteins (*Bovine Colostrum* and *Panax Ginseng*), that could be exploited in the nutraceutical and pharmaceutical filed, coupling the power of the Combinatorial Peptide Ligand Library (CPLL) technology, the SDS-PAGE separation and the high resolution nLC-ESI MS/MS. The potentialities of ProteoMiner technique deserve special mention, for its capability in greatly expanding the visibility of low- to very-low abundance species, up to now unexplored. Overall, it has been underlined the importance of proteomics in the pharmaceutical field and in particular for elucidating the biological effects of complex protein matrices.

Moreover, the proteomic research has concluded with a novelty in *Panax ginseng* studies: a peptidomic investigation by simulating in-silico the enzymatic protein digestion in the human gastrointestinal tract, to evaluate a possible antimicrobial activity expressed by proteolytic peptides.

This work contributes not only to **mapping the *Bovine Colostrum* and *Panax Ginseng* proteome**, but also to setting-up a suitable analytical platform aimed at achieving an high standardization and an in-depth **quality profiling of bovine colostrum** in the manufacture of food supplements.

Such an approach also permits the exploration of novel biological functions of *Bovine Colostrum* and *Panax Ginseng* by searching in the "*homemade*" database for the presence of proteins characterized by innovative functions.

This knowledge can be used in the future for further "explorations" of different natural sources of bioactive macromolecules, known to be effective in several claimed healthy effects.

Scientific Pubblication & Communication

Published Article G. Trevisan, S. Materazzi, C. Fusi, A. Altomare, G. Aldini, M. Lodovici, R. Patacchimi, P. Geppetti, R. Nassini. Novel Therapeutic Strategy to Prevent Chemotherapy-Induced Persistent Sensory Neuropathy by TRPA1 Blockade. *Cancer Research* **2013**, *73*, 3120-3131.

Published Article L. Regazzoni, L. Del Vecchio, A. Altomare, K.-J. Yeum, D. Cusi, F. Locatelli, M. Carini, G. Aldini. Human Serum Albumin Cysteinylation is increased in End Stage Renal Disease Patients and Reduced by Hemodialysis: Mass Spectrometry Studies. *Free Radical Research* **2013**, *47*, 172-180.

Published Article L. Bertoletti, L. Regazzoni, A. Altomare, R. Colombo, M. Colzani, G. Vistoli, L. Marchese, M. Carini, E. De Lorenzi, G. Aldini. Advanced glycation end products of beta2-microglobulin in uremic patients as determined by high resolution mass spectrometry. *J Pharm Biomed Anal.* **2014** Mar;91:193-201.

Published Article D. Garzon, A. Ariza, L. Regazzoni, R. Clerici, A. Altomare, F.R. Sirtori, M. Carini, M.J. Torres, D. Pérez-Sala, G. Aldini. Mass spectrometric strategies for the identification and characterization of human serum albumin covalently adducted by amoxicillin: ex vivo studies. *Chem Res Toxicol.* **2014** Sep 15;27(9):1566-74.

Published Article G. Aldini, M.R. Domingues, C.M. Spickett, P. Domingues, A. Altomare, F.J. Sánchez-Gómez, C.L. Oeste, D. Pérez-Sala. Protein lipoxidation: Detection strategies and challenges. *Redox Biol.* **2015** May 21;5:253-266.

Published Article S. Stegen, B. Stegen, G. Aldini, A. Altomare, l. Cannizzaro, M. Orioli, S. Gerlo, L. Deldicque, M. Ramaekers, P. Hespel, W. Derave. Plasma carnosine, but not muscle carnosine, attenuates high-fat diet-induced metabolic stress. *Appl Physiol Nutr Metab.* **2015** Sep;40(9):868-76.

Published Article

M. Colzani, A. Altomare, M. Caliendo, G. Aldini, P.G. Righetti, E. Fasoli. The secrets of Oriental panacea: Panax ginseng. *J Proteomics.* **2016** Jan 1;130:150-9.

Manuscript submitted under revision

A. Altomare, E. Fasoli, M. Colzani, X. M. Paredes Parra, M. Ferrari, F. Cilurzo, C. Rumio, L. Cannizzaro, M. Carini, P. G. Righetti, G. Aldini. Health benefits of bovine colostrum revealed by an in depth proteomic analysis. *Journal of Pharmaceutical and Biomedical Analysis.*

Manuscript submitted under revision

A. Altomare, Luca L. Regazzoni, Marica M. Orioli, Ximena X. Maria M. Paredes Parra, Francesca F. Selmin, Marina M. Carini and Giancarlo G. Aldini. Set-up and application of an analytical approach for the quality control of purified colostrum as food supplement. *Journal of Pharmaceutical and Biomedical Analysis.*

Abstract in Conference Proceedings

L. Bertoletti, L. Regazzoni, A. Altomare, R. Colombo, G. Vistoli, E. De Lorenzi, G. Aldini. Covalent Modification of Beta2-microglobulin by Reactive Carbonyl Species: High Resolution Mass Spectrometric Studies. *Book of Abstracts of the 5th Meeting NPCF - New Perspectives in Pharmaceutical Chemistry* (Trieste, March 28-30, 2011).

Abstract in Conference Proceedings

L. Bertoletti, L. Regazzoni, A. Altomare, M. Lo Monte, R. Colombo, G. Vistoli, M. Carini, E. De Lorenzi, G. Aldini. Covalent Modification of Beta2-microglobulin induced by Reactive Carbonyl Species: High Resolution Mass Spectrometric and Molecular Modelling Studies. *Book of Abstracts of 14th International Meeting on RDPA - RECENT DEVELOPMENTS IN PHARMACEUTICAL ANALYSIS* (Pavia, September 21-24, 2011).

Oral Communication A. Altomare, E. Fasoli, M. Colzani, M. Carini, P. G. Righetti, G. Aldini. Advanced analytical approach for the quality control of purified bovine colostrum as food supplement. *Book of Abstracts of International Meeting on RDPA - RECENT DEVELOPMENTS IN PHARMACEUTICAL ANALYSIS* (Perugia, June 28- July 1, 2015).

Oral Communication A. Altomare. Health benefits of bovine colostrum revealed by an in depth proteomic analysis. *Book of Abstracts of International Meeting on SSPA - Summer School in Pharmaceutical Analysis* (Rimini, Spetember 16-18, 2015).

Poster Session L. Bertoletti, M. Colzani, A. Altomare, M. Carini, R. Colombo, L. Marchese, E. De Lorenzi, G. Aldini. Covalent Modification of Beta2-microglobulin by Reactive Carbonyl Species: High Resolution Mass Spectrometric Studies. *Book of Abstracts of 24th International Symposium on Pharmaceutical and Biomedical Analysis - PBA* (Bologna, 30 June – 3 July, 2013).