

An in depth proteomic analysis based on ProteoMiner, affinity chromatography and nano-HPLC–MS/MS to explain the potential health benefits of bovine colostrum[☆]

Alessandra Altomare^a, Elisa Fasoli^b, Mara Colzani^a, Ximena Maria Paredes Parra^a, Marina Ferrari^a, Francesco Cilurzo^a, Cristiano Rumio^c, Luca Cannizzaro^{a,d}, Marina Carini^a, Pier Giorgio Righetti^b, Giancarlo Aldini^{a,*}

^a Department of Pharmaceutical Sciences, Università degli Studi di Milano, Via Mangiagalli 25, 20133 Milano, Italy

^b Department of Chemistry, Materials and Chemical Engineering “Giulio Natta”, Polytechnic of Milan, Via Mancinelli 7, 20131 Milan, Italy

^c Department of Pharmacology and Biomolecular Sciences, Università degli Studi di Milano, Via Trentacoste 2, 20133 Milan, Italy

^d Systems Biology Center, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

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 potential 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 bioinformatic 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 when 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 advanced 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.

Keywords:

Bovine colostrum proteome

Low-abundance proteins

Combinatorial peptide ligand libraries

Wound healing

Mass spectrometry

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 newborn 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 potentially 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 defense of the gastro-intestinal (GI) tract [8–11]. BC also has growth-promoting effects on calf gut epithelial cells as well as 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].

[☆] Selected paper from conference of Recent Development in Pharmaceutical Analysis (RDPA2015), June 28–July 1, 2015, Perugia, Italy.

* Corresponding author.

E-mail address: giancarlo.aldini@unimi.it (G. Aldini).

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, favoring 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 exploration represents a difficult challenge because, similarly to other body fluids, the proteins are contained in a large dynamic range, spanning at least 9 orders of magnitude (from hundred mg ml⁻¹ for IgG to few pg for some growth factors) [2]. 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]. In the last five years, CPLL technology has been successfully used for the proteome investigation of human [20,21] sheep [22], goat [23] and cow's [20] milk and of human colostrum [21] but never of bovine colostrum. The CPLL technology is here 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

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 supplied by Bio-Rad Laboratories, Inc., Hercules CA.

β-mercaptoethanol, dithiothreitol (DTT), ammonium persulfate, 3-[3-cholamidopropyl dimethylammonio]-1-propanesulfonate (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). Commercially available ELISA kits for the quantitative measurement of bovine IgG (Bethyl Laboratories, Inc., Tema Rierca, Italy) and bovine lumican (MyBiosource, Valter Occhiena S.R.L., Italy) were used according to the manufacturer's instructions. Spectrofluorimetric readings were performed using a Wallac Victor 2 reader (PerkinElmer).

Pooled, 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 dilution with 3 volumes of demineralized water, the suspension obtained was introduced into a reactor (controlled continuous stirring), where it was heated at 35–36 °C for about 30 min. 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 (5x 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 cm² [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 3 mL 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. 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 2.5 h.

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 μ L of samples containing 20–25 μ g of proteins were mixed with 10 μ L of Laemmli sample buffer (containing 50 mM DTT in case of reducing conditions) and heated at 95 °C for 5 min. 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 μ g of each sample was 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 h) 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.

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 shaking for 10 min. 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 μ L 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 sulfate (100 μ L first elution, 80 μ L 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 μ L of MilliQ H₂O. Gel pieces were destained using 100 μ L of destaining solution (50% 25 mM NH₄HCO₃/50% ACN) for 10 min and then with 100 μ L of acetonitrile for other 10 min.

100 μ L of reducing solution (10 mM DTT in 50 mM NH₄HCO₃) were 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 μ L of alkylating solu-

tion (55 mM IAA in 50 mM NH₄HCO₃) were added to gel pieces and incubated in the dark for 45 min. 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 μ g of trypsin to each sample). After an overnight incubation at 37 °C, the tryptic mixtures were acidified with 1 μ L of formic acid and eluted using 70 μ L of 3% TFA/30%

ACN in H₂O MilliQ, which was added and incubated for 10 min 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 μ L of buffer A (0.1% HCOOH). 5 μ L of each sample were injected on a C18 column (HALO PicoFrit, 75 mM \times 10 cm, 2.7 μ m 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 μ L/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 = 5×10^5 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 1×10^4 counts) under the following conditions: centroid mode, isolation width = 2.5 m/z , AGC target = 1×10^4 and normalized CE = 35 eV. Dynamic exclusion was enabled for 45 s for ions already observed 3 times in 30 s. 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. Two technical replicates were used for each sample and the proteins obtained for each technical replication were merged and the redundancies deleted.

2.6. Data processing using different bioinformatic tools

Protein identification—Proteins was identified using the software Proteome Discoverer 1.3 (Thermo Scientific) via 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

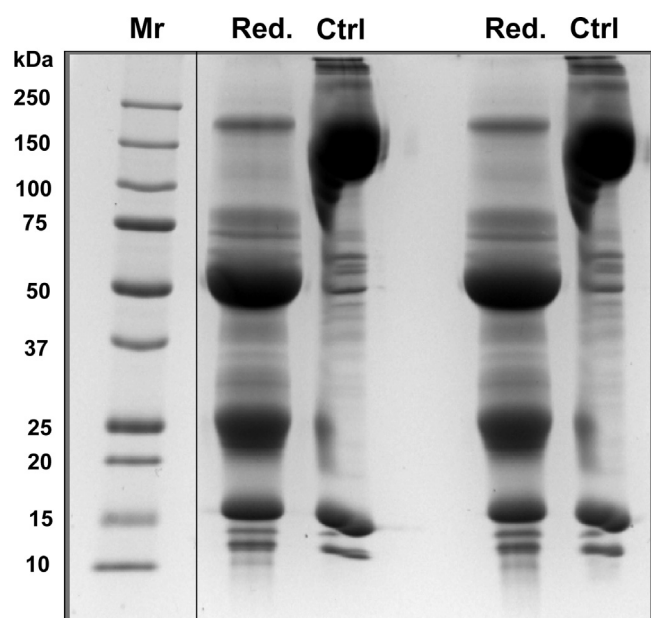


Fig. 1. SDS-PAGE profiling of colostrum proteins. Mr: molecular mass ladder. Samples: Ctrl: control; red: reduced with 50 mM DTT. The two samples to the left have been dialyzed so as to remove sugars and other low MW constituents; the two to the right are untreated. Staining with colloidal Coomassie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

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 according to the following biological processes: wound healing, bactericidal/bacteriostatic, neurogenesis, regulation of bone growth.

3. Results

3.1. SDS-PAGE analysis of bovine colostrum

Fig. 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 centered at about 150 kDa, which disappears under reducing conditions in favor 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), β -lactoglobulin (19.9 kDa), serum

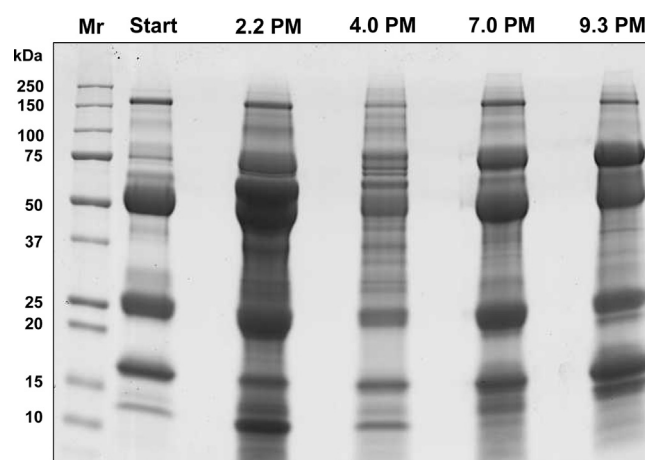


Fig. 2. SDS-PAGE profiling of colostrum proteins. Mr: molecular mass ladder. Samples: start (control, untreated colostrum), 2.2 PM, 4.0 PM, 7.0 PM and 9.3 PM (colostrum captured with CPLL beads at pH 2.2, 4.0, 7.0 and 9.3, respectively). In these last 4 lanes notice the marked reduction of the 50 and 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.

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 whey proteins) [24]. The two lines are duplicated since they represent dialyzed vs. non-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 shown in Fig. 2 represents 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). 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 Fig. S1 shows the pI/Mr distribution of the visible species in the control (left panel) versus the eluate from CPLL beads after pH 4.0 capture (right panel). Many more spots are visible, especially in the acidic map region and in the Mr 10–40 kDa range.

Supplementary material related to this article is found, in the online version, at <http://dx.doi.org/10.1016/j.jpba.2016.01.013>.

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

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 frac-

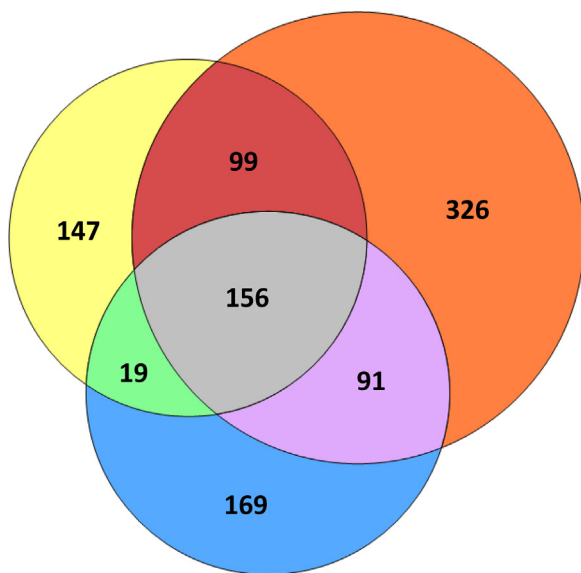


Fig. 3. Venn diagrams showing the separate contributions to protein discovery of the three different CPLL treatments at pH 4.0, 7.0 and 9.3. Yellow: colostrum captured with CPLL beads at pH 4.0 (421 proteins); blue: colostrum captured with CPLL beads at pH 9.3 (435 proteins); Orange: colostrum captured with CPLL beads at pH 7.0 (672 proteins). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tion 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 efficiency of the depletion process was then confirmed by measuring the IgG content by enzyme immunoassay (ELISA assay) of the fractions collected and lyophilized; the IgG content in the depleted fraction was less than 5% of the total amount of IgG as determined before the depletion.

IgG depleted colostrum was treated with combinatorial peptide ligand libraries (CPLL or PM, ProteoMiner) under different pH conditions (pH 2.2, 4.0, 7.0, 9.3).

3.4. Protein identification by MS analysis

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 PAGE slices, were performed.

The Venn diagrams on Fig. 3 shows the number of proteins identified by the CPLL treatments 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. 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 Fig. 4, in which the contribution of an additional CPLL treatment able to capture at pH 2.2 (Fig. 4—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 [26]). 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). In all cases, an average of only 20% detected proteins are in common between the control colostrum and each of the four treatments.

Fig. 4—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 in identification rate is obtained. This huge increase is likely due to low- to very-low abundance proteins present in colostrum, whose signal had been largely masked by the handful of proteins (5–6 major components), which alone constituted >95% of the protein mass in colostrum.

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 Fig. 4—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). Here too the overlap between the two samples is 20–25%.

By comparing the number of proteins identified in the different conditions (Fig. 4) it should be observed 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. 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 IDs. This number was reduced to 634 proteins upon setting high peptide confidence as filter (Peptide RANK 1) (Table S1). This list of proteins was used as input for functional analyses, as described above.

Supplementary material related to this article is found, in the online version, at <http://dx.doi.org/10.1016/j.jpba.2016.01.013>.

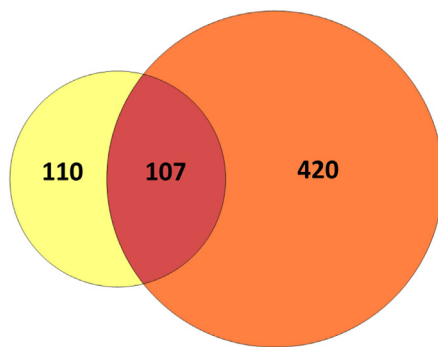
3.5. Descriptive proteomic analysis

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 (Fig. 5—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% could not be annotated. Regarding the molecular function (Fig. 5—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 proteins with catalytic activity. Whereas, with regard to intra/extracellular localization (Fig. 5—panel C), we 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%).

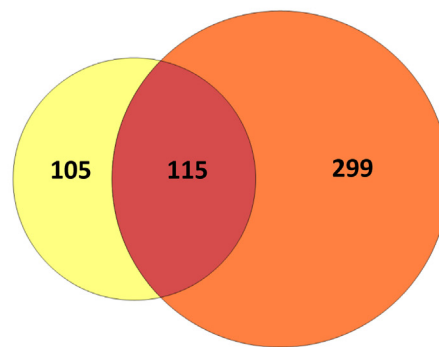
3.6. Functional proteomic analysis

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 func-

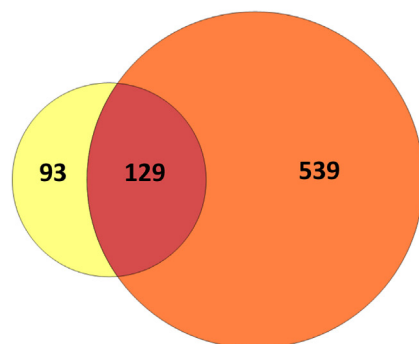
A. Untreated Colostrum Vs. Colostrum captured with CPLL beads at pH 2.2



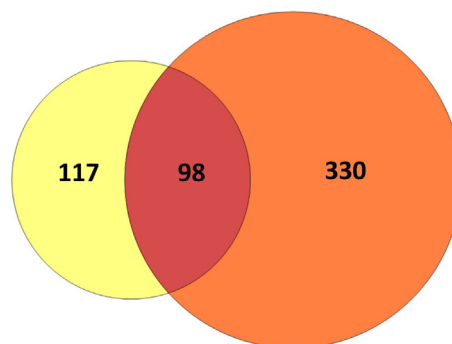
B. Untreated Colostrum Vs. Colostrum captured with CPLL beads at pH 4.0



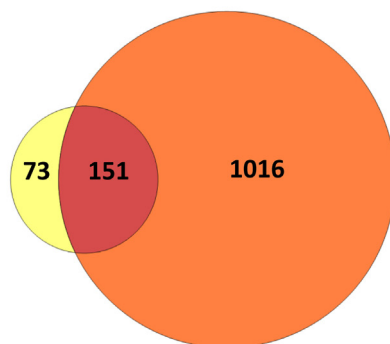
C. Untreated Colostrum Vs. Colostrum captured with CPLL beads at pH 7.0



D. Untreated Colostrum Vs. Colostrum captured with CPLL beads at pH 9.3



E. Untreated Colostrum Vs. Colostrum captured with CPLL beads



F. IgG-depleted Colostrum Vs. IgG-Depleted Colostrum captured with CPLL beads

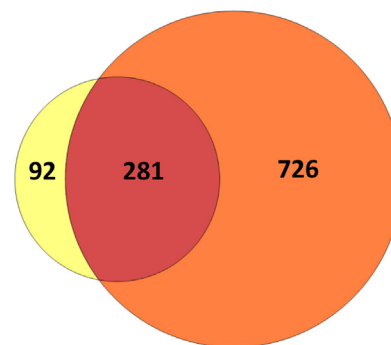


Fig. 4. Venn diagrams showing the separate contributions to protein discovery of four different CPLL treatments at pH 2.2 (panel A), 4.0 (panel B), 7.0 (panel C) and 9.3 (panel D) as compared to the proteome discovery of untreated colostrum samples; panel E summarizes the overall findings in all CPLL 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).

tional information obtained from others databases, namely Gene Ontology, KEGG, Pfam and 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 (Fig. S2).

Supplementary material related to this article is found, in the online version, at <http://dx.doi.org/10.1016/j.jpba.2016.01.013>.

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 het-

erogeneous composition of colostrum, which is not associated with a well-defined cellular system but is probably linked to many cellular targets.

The functional analysis was then focused on the wound healing process, which was selected since there are experimental data reporting the colostrum capacity to induce, enhance and promote wound healing in vitro and in different animal models [13,27,28,29]. To do this, the bovine colostrum database was filtered by considering the main physiological steps involved in the wound healing process such as: hemostasis, inflammation, proliferation and tissue re-modelling. Ninety three proteins involved in the wound healing process were identified (Appendix B) and were

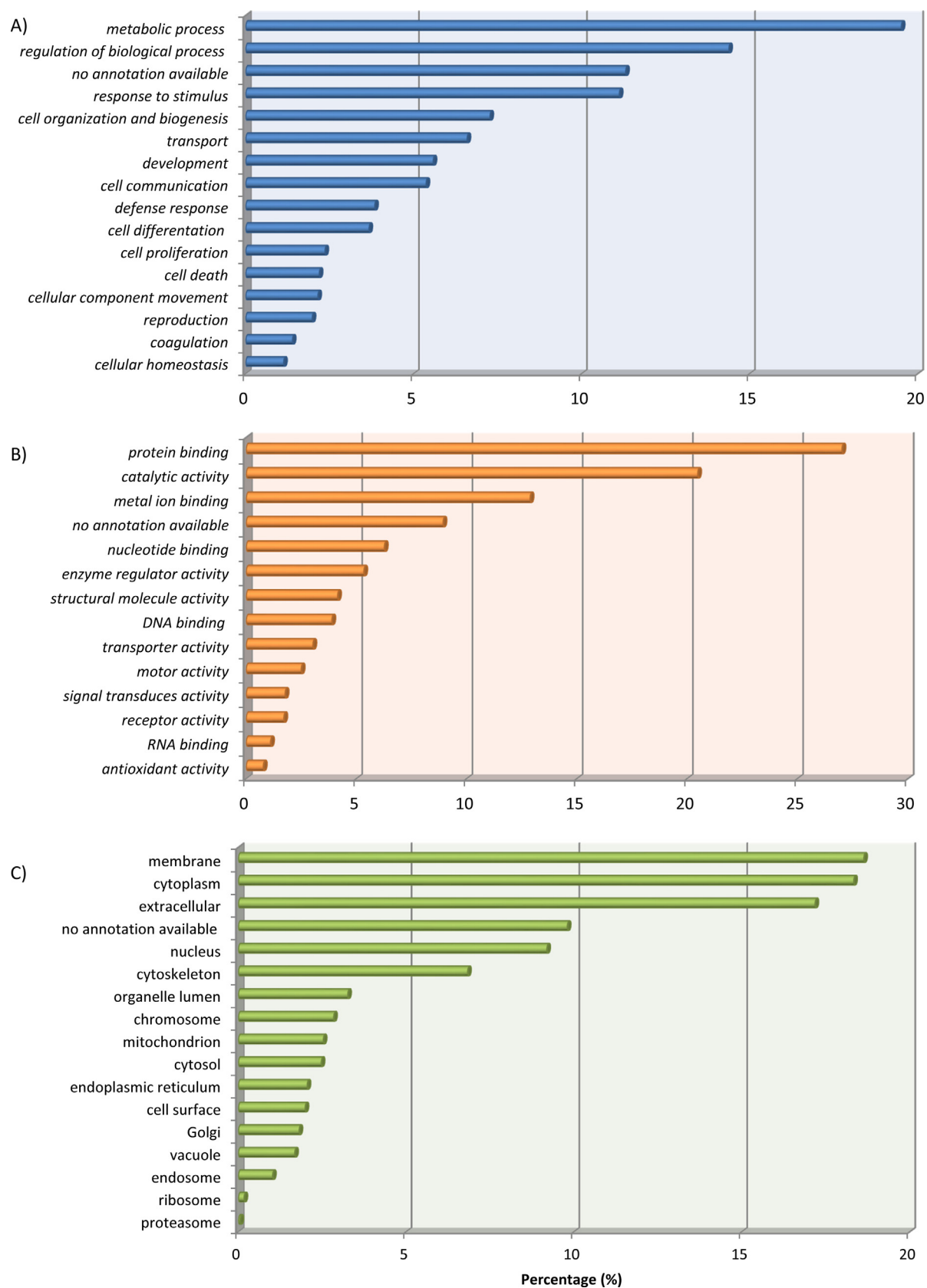


Fig. 5. BC protein distribution based on the biological process (panel A), molecular function (panel B), and cellular localization (panel C), obtained by using enrichment tool by proteome discoverer.

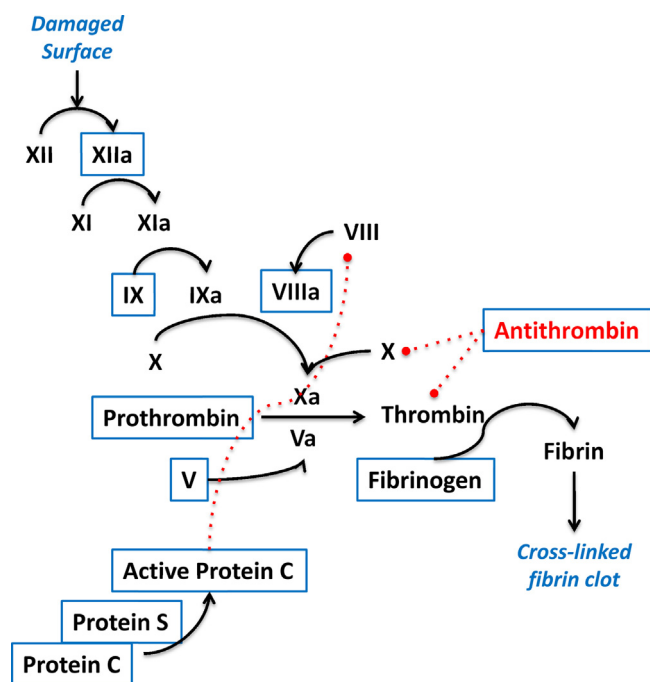


Fig. 6. Haemostasis pathway; the gene products highlighted by a blue colour have been detected in bovine colostrum.

clustered on the basis of their biological functions as follows: 17 are involved in the hemostasis process, 9 in the inflammatory process, 9 in angiogenesis, 15 in extracellular matrix organization and 13 in regulation of epithelial cell proliferation; 5 proteins are involved in collagen synthesis and 25 in epithelial remodeling. 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, Fig. 6 shows the involvement of the identified proteins in the molecular events leading to hemostasis, 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.

The BC proteome can then be used to search for novel potential health applications of BC or to identify protein constituents potentially active in a certain disease. An approach that could be used to reach this goal would consist of generating by literature/patent searching a list of proteins potentially useful as exogenous agents which is then matched with the BC database. As an example, a list of proteins active in wound healing and re-epithelialization was generated by searching proteins promoting wound healing and re-epithelialization in PubMed and 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 to enhance re-epithelialization in human corneal diseases, namely lumican, galectin-3 and vitronectin. The content of lumican in BC (as determined by ELISA) ranged between 10 and 30 $\mu\text{g mg}^{-1}$.

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 [5,31]. 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 identification of unknown compounds and consequently do not add to what is already known of the molecular mechanisms of BC. To overcome such limit, we here propose an off-target method based on proteomic analysis which 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 additional sample preparation 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-established “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) so as to modulate the surface charge of both the proteins and the hexa-peptide beads.

The one and two-dimensional electrophoretic analysis of extracts confirmed that the sample treatment with the CPLL technology greatly improved the exploration 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 was 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 [32].

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.

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 Appendix A 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 on the basis of 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 effect of BC in both animals and humans [5,31]; besides immunoglobulins, we found a cluster of 4.02% pro-teins 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 pro-teins 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 which were then clustered on the basis of the 7 biological events leading to wound healing. Hence the wound healing process, which has thus far been linked to the presence of growth factors can also be explained on the basis of above mentioned proteins. 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, a panel of proteins involved in this biological process can be selected for qualitative and quantitative analysis to ensure the quality and efficacy 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 a set of proteins which have been reported to be beneficial for treating epithelial defects in the cornea: lumican, galectin-3 and vitronectin. Lumican is a 40 kDa keratan sulfate proteoglycan which is receiving growing interest both as a promoter of corneal epithelial wound healing and for maintaining ordered collagen fibrils which are vital in keeping the cornea transparent [30,36]. Exogenous galectin-3 was found to enhance re-epithelialization in corneal epithelium in several in vivo animal

experiments by promoting epithelial cell adherence to the major ECMs and onto integrins [33,34]. Finally vitronectin was found to specifically promote ocular surface repair through its migratory activity mediated in part via integrin-ECM engagement and by promoting cell migration [35]. The three proteins were found to be highly homologous in respect to the corresponding human isoforms as confirmed by their marked sequence identity (71% for vitronectin; 79% for galectin; 88% for lumican). High homologies were also found for other bioactive proteins such as transforming growth factor beta-2 (99% of identity), insulin-like growth factor (87%), Xanthine dehydrogenase (89%), fibronectin (94%), lactoperoxidase (83%). Such a high homology between human and bovine proteins further increases the likelihood that BC proteins can contribute to health benefits in humans.

5. Conclusion

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.

Acknowledgments

We gratefully acknowledge financial support from Advances in Medicine (AIM, Bologna, IT). Luca Cannizzaro is a recipient of the Post-Doctoral Scholarship, Chulalongkorn University Rachadaphiseksomphot Fund.

References

- [1] S.I. Kehoe, B.M. Jayarao, A.J. Heinrichs, A survey of bovine colostrum composition and colostrum management practices on Pennsylvania dairy farms, *J. Dairy Sci.* 90 (9) (2007) 4108–4116.
- [2] P. Sacerdote, F. Mussano, S. Franchi, A.E. Panerai, G. Bussolati, S. Carossa, A. Bartorelli, B. Bussolati, Biological components in a standardized derivative of bovine colostrum, *J. Dairy Sci.* 96 (3) (2013) 1745–1754.
- [3] E.J.H. Bechara, F. Dutra, V.E.S. Cardoso, A. Sartori, K.P.K. Olympio, C.A.A. Penatti, A. Adhikari, N.A. Assuncao, The dual face of endogenous alpha-aminoketones: pro-oxidizing metabolic weapons? *Comp. Biochem. Physiol. C—Toxicol. Pharmacol.* 146 (1–2) (2007) 88–110.
- [4] M. Conneely, D.P. Berry, J.P. Murphy, I. Lorenz, M.L. Doherty, E. Kennedy, 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.* 97 (11) (2014) 6991–7000.
- [5] M. Rathe, K. Müller, P.T. Sangild, S. Husby, Clinical applications of bovine colostrum therapy: a systematic review, *Nutr. Rev.* 72 (4) (2014) 237–254.
- [6] D.M. Weaver, J.W. Tyler, D.C. VanMetre, D.E. Hostetler, G.M. Barrington, Passive transfer of colostrum immunoglobulins in calves, *J. Vet. Intern. Med.* 14 (6) (2000) 569–577.
- [7] G.S. Kelly, Bovine colostrums: a review of clinical uses, *Altern. Med. Rev.* 8 (4) (2003) 378–394.
- [8] C. Bühler, H. Hammon, G.L. Rossi, J.W. Blum, 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.* 76 (3) (1998) 758–765.
- [9] Cairangzhuoma, M. Yamamoto, Xijier, M. Inagaki, K. Uchida, K. Yamashita, S. Saito, T. Yabe, Y. Kanamaru, A preparation of cow's late colostrum fraction containing α s1-casein promoted the proliferation of cultured rat intestinal IEC-6 epithelial cells, *Biosci. Biotechnol. Biochem.* 76 (3) (2013) 992–997.
- [10] Cairangzhuoma, M. Yamamoto, H. Muranishi, M. Inagaki, K. Uchida, K. Yamashita, S. Saito, T. Yabe, Y. Kanamaru, Skimmed, sterilized, and concentrated bovine late colostrum promotes both prevention and recovery from intestinal tissue damage in mice, *J. Dairy Sci.* 96 (3) (2013) 1347–1355.
- [11] A.C. Støy, P.M. Heegaard, T. Thymann, M. Bjerre, K. Skovgaard, M. Boye, B. Stoll, M. Schmidt, B.B. Jensen, P.T. Sangild, Bovine colostrum improves intestinal function following formula-induced gut inflammation in preterm pigs, *Clin. Nutr.* 33 (2) (2014) 322–329.
- [12] C.J. Doillon, F. Lehance, L.J. Bordeleau, M.P. Laplante-Campbell, R. Drouin, Modulatory effect of a complex fraction derived from colostrum on fibroblast contractility and consequences on repair tissue, *Int. Wound J.* 8 (3) (2011) 280–290.

- [13] D. Kovacs, G. Cardinali, N. Aspite, M. Picardo, Bovine colostrum promotes growth and migration of the human keratinocyte HaCaT cell line, *Growth Factors* 27 (6) (2009) 448–455.
- [14] S. Tucci, R. Mancini, C. De Vitis, A. Noto, E. Marra, A. Lukic, M.R. Giovagnoli, M. Moscarini, Colostrum in menopause effects on vaginal cytology/symptoms, *Clin. Exp. Obstet. Gynecol.* 40 (2) (2013) 219–221.
- [15] A. Galitsopoulou, A.M. Michaelidou, G. Menexes, E. Alichanidis, Polyamine profile in ovine and caprine colostrum and milk, *Food Chem.* 173 (2015) 80–85.
- [16] S. Purup, M. Vestergaard, L.O. Pedersen, K. Sejrsen, Biological activity of bovine milk on proliferation of human intestinal cells, *J. Dairy Res.* 74 (1) (2007) 58–65.
- [17] M. Kramski, R.J. Center, A.K. Wheatley, J.C. Jacobson, M.R. Alexander, G. Rawlin, D.F. Purcell, 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.* 56 (8) (2012) 4310–4319.
- [18] A. Nissen, E. Bendixen, K.L. Ingvarsen, C.M. Røntved, In-depth analysis of low abundant proteins in bovine colostrum using different fractionation techniques, *Proteomics* 12 (18) (2012) 2866–2878.
- [19] P.G. Righetti, E. Boschetti, L. Lomas, A. Citterio, Protein equalizer technology: the quest for a democratic proteome, *Proteomics* 6 (14) (2006) 3980–3992.
- [20] A. Coscia, S. Orrù, P. Di Nicola, F. Giuliani, A. Varalda, C. Peila, C. Fabris, A. Conti, E. Bertino, Detection of cow's milk proteins and minor components in human milk using proteomics techniques, *J. Matern. Fetal Neonatal Med.* 25 (Suppl 4) (2012) 54–56.
- [21] Y. Liao, R. Alvarado, B. Phinney, B. Lönnerdal, Proteomic characterization of human milk whey proteins during a twelve-month lactation period, *J. Proteome Res.* 10 (4) (2011) 1746–1754.
- [22] M. Ha, M. Sabherwal, E. Duncan, S. Stevens, P. Stockwell, M. McConnell, Al El-Din Bekhit, A. Carne, In-depth characterization of sheep (*Ovis aries*) milk whey proteome and comparison with cow (*Bos taurus*), *PLoS One* 10 (10) (2015) e0139774.
- [23] V. Cunsolo, E. Fasoli, R. Saletti, V. Muccilli, S. Gallina, P.G. Righetti, S. Foti, Zeus, Aesculapius, Amalthea and the proteome of goat milk, *J. Proteomics* 128 (2015) 69–82.
- [24] M.M. El-Loly, Identification and quantification of whey immunoglobulins by reversed phase chromatography, *Int. J. Dairy Sci.* 2 (3) (2007) 268–274.
- [26] C. Esteve, A. D'Amato, M.L. Marina, M.C. García, A. Citterio, P.G. Righetti, Identification of olive (*Olea europaea*) seed and pulp proteins by nLC–MS/MS via combinatorial peptide ligand libraries, *J. Proteomics* 75 (8) (2012) 2396–2403.
- [27] M. Blais, Y. Pouliot, S. Gauthier, Y. Boutin, M. Lessard, A gene expression programme induced by bovine colostrum whey promotes growth and wound-healing processes in intestinal epithelial cells, *J. Nutr. Sci.* 3 (2014) e57.
- [28] R.J. Playford, D.N. Floyd, C.E. Macdonald, D.P. Calnan, R.O. Adenekan, W. Johnson, R.A. Goodlad, T. Marchbank, Bovine colostrum is a health food supplement which prevents NSAID induced gut damage, *Gut* 44 (5) (1999) 653–658.
- [29] C. Torre, I. Jeusette, M. Serra, P. Brazis, A. Puigdemont, Bovine colostrum increases proliferation of canine skin fibroblasts, *J. Nutr.* 136 (Suppl. 7) (2006) 2058S–2060S.
- [30] C.Y. Liu, W.W. Kao, Lumican promotes corneal epithelial wound healing, *Methods Mol. Biol.* 836 (2012) 285–290.
- [31] S. Bagwe, L.J. Tharappel, G. Kaur, H.S. Buttar, Bovine colostrum: an emerging nutraceutical, *J. Complement. Integr. Med.* 12 (2015) 175–185.
- [32] U. Restuccia, E. Boschetti, E. Fasoli, F. Fortis, L. Guerrier, A. Bachi, A.V. 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* 72 (6) (2009) 1061–1070.
- [33] A. Fujii, T.R. Shearer, M. Azuma, Galectin-3 enhances extracellular matrix associations and wound healing in monkey corneal epithelium, *Exp. Eye Res.* 137 (2015) 71–78.
- [34] C. Yabuta, F. Yano, A. Fujii, T.R. Shearer, M. Azuma, Galectin-3 enhances epithelial cell adhesion and wound healing in rat cornea, *Ophthalmic Res.* 51 (2) (2014) 96–103.
- [35] S. Chow, N. Di Girolamo, Vitronectin: a migration and wound healing factor for human corneal epithelial cells, *Invest. Ophthalmol. Vis. Sci.* 55 (10) (2014) 6590–6600.
- [36] S. Amjadi, K. Mai, P. McCluskey, D. Wakefield, The role of lumican in ocular disease, *ISRN Ophthalmol.* 2013 (2013) 632302.