

ARTICLE

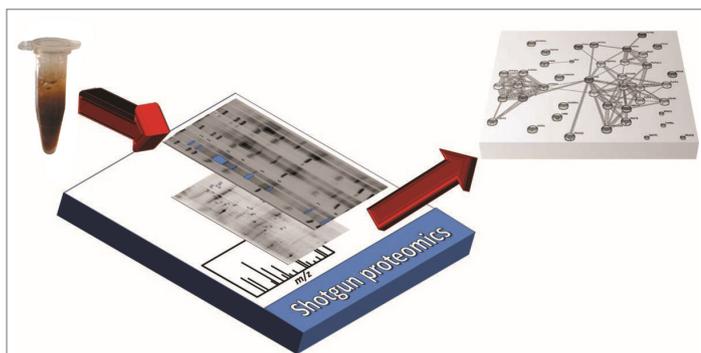
# A Contribution to the Analytical Proteomics of Human Intestinal Mucus: Sampling, Sample Preparation, and Protein Identification Strategies

Paola Andrea Gómez Buitrago<sup>1\*</sup>   Carlos Augusto González Correa<sup>1</sup>  Mario Santacoloma Osorio<sup>2</sup>, Gonzalo Taborda Ocampo<sup>1</sup>  Hector Henrique Ferreira Koolen<sup>3</sup>  Marco Aurélio Zezzi Arruda<sup>3</sup> 

<sup>1</sup>Research Group on Electrical Bio-Impedance, Doctoral Program on Biomedical Sciences, Universidad de Caldas, Calle 65 # 26-10, Edificio de Laboratorios, Office 502, Manizales, Colombia

<sup>2</sup>Clinical Department, Universidad de Caldas, Carrera 25 # 48-57, Manizales, Colombia

<sup>3</sup>Institute of Chemistry, University of Campinas (UNICAMP) and National Institute of Science and Technology in Bioanalytics, PO Box 6154, Campinas, Brazil



The colon is inhabited by circa ten bacteria for each cell of our organism (microbiota) and it is estimated that, altogether, they possess between 100 to 150 times more genes (microbiome) than the human genome. Between these two sets of genomes, acting as a biological interface, there is a mucus bilayer biochemically structured by the Mucin 2 glycoprotein. Several analytical proteomics strategies are presented in this study, which contribute to the proteomic map of human intestinal mucus, enlarging knowledge of this

important sample. Utilizing mucus obtained from sampling via a transanal irrigation procedure, 15 different protocols for running the samples through 2-D PAGE in the pH ranges 3-10 or 4-7 were evaluated. The protocol that showed no streaking in 2-D PAGE and yielded more protein spots (116 and 220 spots were found in the 3 to 10 and 4 to 7 pH ranges, respectively) was considered the best and was the one that was finally used (protocol C). For improving protein identification, SDS-PAGE, 2-D PAGE and shotgun proteomics are carried out through nLC-MS/MS, and a Venn diagram built-up from identified proteins as well as cellular, bacterial and other proteins of clinical interest. Mucin 2, which is the main component of the mucus, is accurately identified, despite its insolubility and high molecular weight, indicating the complementarity and the success of the optimization conditions of all strategies employed. Finally, the interaction between 23 identified proteins is carried out, indicating two major classes of proteins: energy metabolism and energy consumption.

**Keywords:** Intestinal mucus, proteins, sampling method, shotgun proteomics.

**Cite:** Gómez-Buitrago, P-A.; González-Correa, C-A.; Santacoloma-Osorio, M.; Taborda-Ocampo, G.; Koolen, H-H.F.; Arruda, M. A. Z. **A Contribution to the Analytical Proteomics of Human Intestinal Mucus: Sampling, Sample Preparation, and Protein Identification Strategies.** *Braz. J. Anal. Chem.*, 2020, 7 (29), pp 45-63. doi: <http://dx.doi.org/10.30744/brjac.2179-3425.AR-41-2019>

Submitted 25 November 2019, Resubmitted 15 May 2020, Accepted 6 July 2020, Available online 6 August 2020.

## INTRODUCTION

Human intestinal mucus (HIM) is a highly complex sample of incompletely defined chemical structure. However, it is mainly composed of Mucin 2 (MUC2), a glycoprotein containing ca. 5200 amino acids and having a molecular weight of 2.6 MDa when completely O-glycosylated in the proline, threonine and serine (PTS) domains [1]. This macromolecule forms a net-like structure, where, among several other substances, small proteins, bacteria, viruses and Archaea are found [2]

In the colon, the mucus gel comprises a bilayer [3]: the first is a luminal layer with a lower concentration of MUC2, allowing the penetration of bacteria into the mucus and their attachment by means of glycans [1,4]. The second is the mucosal layer, which is in contact with the colonic epithelium. Being impenetrable and acting as a barrier, it prevents passage of bacteria due to the highest concentration of MUC2. The structure of colonic mucus is responsible for two of the main physiological functions necessary for our good health: well-being and survival. Additionally, this mucus possesses a complex variety of microorganisms that inhabit the colon, known as the colonic microbiota [5,6], and acts as an important line of defense against any kind of pathogenic microorganism [7], as well as incoming acids, enzymes and all metabolic waste. In this manner, several authors consider glycans attached to MUC2 to be the body's first line of defense [8,9].

It is important to highlight that the colon has approximately 10 bacteria per single cell in the entire human organism, and the size of the microbiome exceeds the human genome by 100 to 150 times [10]. In the small intestine, the bacterial density is lower when compared to the colon, and the mucus is more permeable to bacteria. However, the mucus from the small intestine has a higher concentration of antibacterial peptides, and proteins are secreted by Paneth cells and enterocytes. These cells seem to contribute to the production of Immunoglobulin A and mucins, and are secreted in the crypt mouth, covering the villi to diminish the risk of its contact with bacteria [11].

To the best of our knowledge, those few studies focusing on intestinal mucus proteomics have been conducted using mouse models. Using agarose-1-D and LC/MS-MS, previous results showed that the main component of the colonic mucus bilayer is MUC2 [1]. Later, these same authors performed a proteomic analysis of both layers of mouse colonic mucus, as well as of samples obtained from human biopsies and reported that cell proteins, serum and bacterial proteins and MUC2 were linked to the Fcgbp protein [12].

Besides being a complex sample, sampling of HIM is not an easy task [13,14]. In contrast to rats, which can be sacrificed, colonoscopy is the most frequently used strategy for sampling HIM [15], accounting for the small amount of mucus obtained with each sampling, which makes its handling and acquisition of possible replicates difficult. This hampers the work with such samples and, more probably, explains the scarce literature regarding HIM proteomics.

Sampling from the recto sigmoid using cytology brushes had been applied by evaluating protein concentrations with Surface Enhanced Laser Desorption Ionization – TOF (SELDI-TOF) through mono- and oligo-saccharides analysis [15]. In order to alleviate sampling problems, some authors use *in vitro* cultured human goblet cells, with the proteome of mucus obtained by ultracentrifugation of such cells. Using LC-MS and MS/MS, a set of membrane proteins were identified, but proteins such as FAM62B and ATP6AP2 were reported as structural components of the colonic mucus [16].

Due to the complexity of HIM and the difficulty to sample and handle it, this study focuses not only on some strategies for analyzing HIM proteins, but also on a sampling strategy for obtaining HIM samples in sufficient quantity and quality, while demonstrating how to isolate and treat them. All these strategies may contribute to enhancing analytical proteomic studies of HIM.

## MATERIALS AND METHODS

### ***Sampling and treatment of HIM samples***

Samples of HIM were collected from four overweight individuals ( $32.4 \pm 6.2$  years old) participating in a weight loss program held at the University of Caldas (UC), Colombia. Exclusion criteria for this program included: uncontrolled high blood pressure, diabetes, pregnancy or breast-feeding, weight

loss  $\geq 5$  kg in the previous six months, major psychiatric conditions (such as nervous bulimia, abuse of psychotropic substances, or under psychiatric treatment), and suspicion of moderate to severe diverticulosis, endoscopically diagnosed. All participants signed an informed consent form approved by the Ethical Committee of the UC. The protocol included 5 days of supplemented fasting, daily consumption of psyllium and montmorillonite and 1 daily transanal colon irrigation using a Colema Board® (California, USA), as suggested by the manufacturer. Throughout the treatment, volunteers were under the medical supervision of Carlos-Augusto González-Correa and Mario Santacoloma-Osorio, a general doctor and a gastroenterologist, respectively. The material eliminated from the 3<sup>rd</sup> to the 5<sup>th</sup> days during the hydrotherapy sessions (from 0.5 mL up to 5.0 mL) was collected using a plastic sieve as, during the first two days, the patients mostly eliminated fecal matter. Visual identifiable mucus associated with this material was separated using tweezers, subsequently suctioned with a Pasteur pipette and collected in 1.5 mL micro centrifuge (Eppendorf) tubes. Mucus collected after the hydrotherapy was then mixed with a cleaning buffer (0.1 M Tris-HCl, in 154 mM NaCl, 0.01% azide, 0.02% Tween-20, pH 7.4) [15] in 1:1 v/v proportion. This material was then centrifuged at 350 g for 5 min, and the supernatant (50  $\mu$ L) discarded to eliminate bacteria and other solid particles. Gram staining confirmed the absence of bacteria in the remaining supernatant. The obtained pellet was then stored at -80 °C for later analysis.

### ***Extraction of proteins from HIM samples using protocol C***

For protein extraction, a 300  $\mu$ L sample was treated as explained in the preceding section, adding 300  $\mu$ L of denaturing buffer (8 M urea, Bio-Rad Laboratories Hércules, California, USA), 2% 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate – CHAPS (Bio-Rad Laboratories Hércules, California, USA), 1% Dithiothreitol – DTT (Bio-Rad Laboratories Hércules, California, USA), 50 mM Tris (Bio-Rad Laboratories Hércules, California, USA), pH 7.0 and the addition of 15  $\mu$ L of phenylmethylsulphonyl fluoride – PMSF (Sigma-Aldrich Milwaukee, Wisconsin, USA) to inhibit proteases. Lysis was carried out for 10 min and the proteins were cleaned through precipitation with *ReadyPrep 2-D Clean-up* kit from Bio-Rad™, following the manufacturer's instructions. The obtained precipitate was washed with 100  $\mu$ L cold acetone (Carlo Erba, Paris, France) and water (90:10 mixture) by vortex stirring for 1 min and stored for 10 min at -20 °C. Washing was carried out three times, eliminating the supernatant and adding fresh acetone solution each time. After the third wash and storage at -20 °C, the solution was centrifuged at 350 g for 5 min in a 5430 Eppendorf centrifuge, the supernatant discharged and 100  $\mu$ L of the rehydration buffer (7 M urea, 2 M thiourea, 2% 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate – CHAPS, 50 mM Dithiothreitol – DTT, 1% ampholytes. All reagents were acquired from Bio-Rad Laboratories Hércules, California, USA) [17]. The sample required 1 h to dissolve and it was, intermittently, manually stirred during that time. Protein quantity was determined after protein extraction. Total protein content was measured using the *2-D Quant Kit* from GE Health Care Life Sciences (Uppsala, Sweden) following the manufacturer's instructions. A sample from a single volunteer was included for protein extraction, and the quantification performed three times in the same day. Protein extraction was carried out for all the samples obtained from volunteers that expelled mucus during the third day of treatment.

### ***Sodium dodecyl sulphate polyacrylamide gel electrophoresis***

Samples from three volunteers were run in duplicate, representing two analytical and three biological replicas. The sample protein fraction with protocol C was also evaluated by SDS-PAGE under denaturing conditions. Gels of 17 cm were prepared, and each well was filled with a 1:1 mix of protein fraction and buffer (0.5 M Tris-HCl, 10% (w/v) glycerol (Sigma-Aldrich Milwaukee, Wisconsin, USA), 10% (v/v) SDS, 0.001% (m/v) bromophenol blue and 5% (v/v)  $\beta$ -mercaptoethanol with a pH of 6.8, reagents were acquired from Bio-Rad Laboratories Hércules, California, USA. Gels were run using an Ettan™ Dalt six electrophoresis system (GE Health Care Life Sciences, Uppsala, Sweden), and the processing carried out in two steps: 30 min at 90 V, 15 mA/gel and 5 h at 600 V and 25 mA/gel.

### Two-dimensional gel electrophoresis

For this purpose, 15 different protocols of protein extraction were tested (see Table I) in order to standardize a method for extraction and separation of protein from HIM, carrying out one biological and three analytical replicas for each protocol. Table I lists all the protocols that were used in the search to obtain the best results in 2-D PAGE. After studying the results, protocol C was chosen for protein extraction. The sample from one volunteer was used for 2-D PAGE in order to generate three gels (17 cm-strips) with pH values in the range 3-10 and another three gels with pH values of between 4-7 (13 cm-strips). Two biological replicas were processed, but only the gels from one individual are shown in Figure 3. The sample protein fraction was applied to IPG gels, using 250  $\mu$ L for hydration of the 13 cm-strips and 300  $\mu$ L for hydration of the 17 cm-strips, and then incubated for 12 h.

IEF on the 17 cm gel strips was performed using PROTEAN<sup>®</sup> IEF Cell equipment (Bio-Rad, Hércules, California) at CENICAFE laboratories (Chinchiná, Caldas, Colombia) using a program with four steps: 1) constant 500 V for 500 Vh; 2) a gradient at 1000 V for 800 Vh; 3) a gradient at 1000 V for 16500 Vh, and 4) constant 1000 V for 3700 Vh. IEF on the 13 cm gels was conducted at the GEPAM laboratory (Unicamp, Campinas, Brazil) using Amersham Biosciences equipment (GE Healthcare, model 18113002 EPS601). The program used was the same as that for 17 cm gels, except for step 3, which was performed at 11300 V h. The IEF process was optimized in the laboratory. Once the sample was equilibrated with reducing and alkylating agents, the second dimension was carried out in a 12% (m/v) polyacrylamide; the final concentration of the gels was: 12% Acrilamida/Bis, 0.37 M Tris, 0.1% SDS (Sodium dodecyl sulfate), 0.05% APS (Ammonium Persulfate) and 0.05% TEMED. All reagents were acquired from Bio-Rad Laboratories Hércules, California, USA. The power source was operated in two stages under the following running conditions: 1) 30 min, 15 mA/gel at 90 V, and 2) 4-5 h 25 mA/gel at 600 V. Finally, the gels were stained with either Coomassie colloidal or Sypro Ruby (Bio-Rad Laboratories Hércules, California, USA) and scanned in a Pharos FX Molecular Imager (Bio-Rad, Hércules, California, USA) at CENICAFE laboratories, as well as by an ImageScanner TM II (GE healthcare, Sweden) at the GEPAM laboratory.

**Table I:** Protein extraction protocols evaluated for 2-D PAGE of HIM

Protocols	Brief description
A	Protein precipitation using acetone, 10% v/v trichloroacetic acid and 0.007% DTT, and incubation during 24h at -8 °C.
B	[18]
C	[19] utilizing a denaturing buffer (8 M urea, 2% (m/v) CHAPS, 1% (m/v) DTT, 50 mM Tris pH 7.0), PMSF, cleaning with Protein Clean Up Kit de Bio-Rad (California, USA) and acetone.
D	[18] and [19] utilizing a pool of 3 samples.
E	[19] utilizing the chemical M-PER for extraction and 300 $\mu$ g of sample.
F	[19] utilizing the chemical M-PER for extraction and 100 $\mu$ g of sample.
G	[20] for glucosidases denaturation and separation of small proteins.
H	Protocol C and double cleaning with Protein Clean Up Kit de Bio-Rad (California, USA).
I	[21]

**Table I:** Protein extraction protocols evaluated for 2-D PAGE of HIM (Continuation)

Protocols	Brief description
J	Protein precipitation with ammonium sulfate. Precipitation of 300 $\mu$ L of cleaned mucus sample with ammonium sulfate 0 – 20%. Precipitated incubated overnight. 300 $\mu$ L of sample clean mucus is precipitated with ammonium sulphate ( $(\text{NH}_4)_2\text{SO}_4$ ) in the range of 0 to 20% saturation (0- 20%), then allowed to stand overnight and centrifuged at 14548 g for 5 minutes. The supernatant of 20% saturation was precipitated again to 60% saturation (20-60%), repeating the process; two precipitates (ranges 20-60% 0-20% and saturation) comprising 200 $\mu$ L diluted in hydration buffer containing 7 M urea, 2 M thiourea, 2% CHAPS, 50 mM DTT and 1% ampholytes were then finally obtained.
K	[21] and microdialysis.
L	Protein precipitation with metasulfonic acid (MSA). 300 $\mu$ L of 1 M MSA was added to 300 $\mu$ L of clean mucus sample, and was allowed to precipitate on ice for 40 min. The sample was centrifuged at 14,548 g for 15 min. Then it was cleaned using the Clean Up Kit (Bio-Rad - California, USA), and finally resuspended in 200 $\mu$ L of hydration buffer which contains 7 M urea, 2 M thiourea, 2% CHAPS, 50 mM DTT and 1% ampholytes.
M	Acid hydrolysis and microdialysis. 62 $\mu$ L of 1M HCl is added to 300 $\mu$ L of clean mucus sample, to pH 4.5. Subsequently microdialysis is performed with a membrane pore size 3 kDa for 12 h against deionized water. Finally, the sample was centrifuged at 14,548 g for 5 min, the supernatant is discarded and 100 to 200 $\mu$ L of hydration buffer added which contains 7 M urea, 2 M thiourea, 2% CHAPS, 50 mM DTT and 1% ampholytes.
N	[19] utilizing a denaturing buffer, PMSF and triple cleaning with acetone.
O	[19] utilizing a denaturing buffer, PMSF and triple cleaning with acetone and TCA for precipitating the proteins.

### Shotgun proteomics

A sample from a single volunteer was used for this experiment, giving one biological replica which was processed twice: 150  $\mu$ L of acetonitrile – ACN (JT Baker, Pennsylvania, USA) at 53% (v/v) [22] was added to 100  $\mu$ L of extracted protein, in addition to a denaturing buffer and the absence of any cleaning treatment. This preparation was stirred for one minute and placed in an ultrasonic bath for 10 min. The process was repeated once and the solution was then centrifuged at 12000 g for 10 min. A digestion with trypsin (Promega Corporation Madison, Wisconsin, USA) was carried out in the supernatant and the precipitate was eliminated [23]. Following [22], the precipitate obtained with acetonitrile was eliminated as a depletion method of abundant proteins, especially albumin.

Because proteins did not undergo any cleaning treatment, peptides were desalted using a  $\text{C}_{18}$ -solid phase extraction cartridge (Allcrom 8B-S001-EAK). The cartridges were activated with 3 mL ACN 100%, equilibrated with 1 mL 50/50 ACN/ $\text{H}_2\text{O}$  with 0.1% (v/v) formic acid (JT Baker Center Valley, Pennsylvania, USA) and then loaded in 3 mL of 0.1% (v/v) Trifluoroacetic acid – TFA (Carlo Erba, Paris, France). Desalting was carried out after sample loading. The sample as washed with 3 mL 0.1% (v/v) TFA and equilibrated with 1 mL 0.1% (v/v) formic acid. The sample was eluted in 2 mL 50/50 v/v ACN/ $\text{H}_2\text{O}$  with 0.1% (v/v) formic acid and 1 mL 80/20 v/v ACN/ $\text{H}_2\text{O}$ .

### Mass spectrometry for protein identification

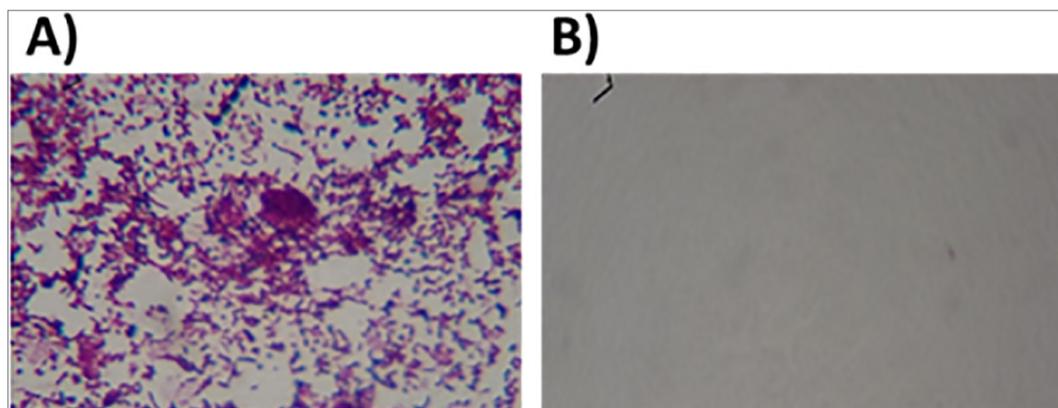
Samples were concentrated to ca. 50  $\mu$ L using a Speed Vac Concentrator (model SPD131DDA, Thermo Scientific). The peptides in the digested samples were analyzed using a UPLC nanoAcquity Waters chromatograph coupled to a Waters SYNAPT HDMS spectrometer equipped with a nanoESI source. A

total of 2 to 5  $\mu\text{L}$  of the aqueous sample was injected using the UPLC autosampler and placed in a guard pre-column (Waters Symmetry  $\text{C}_{18}$ , 20 mm x 180  $\mu\text{m}$ ) for desalination with a 5  $\mu\text{L}/\text{min}$  flow of 97/3 v/v  $\text{H}_2\text{O}/\text{acetonitrile}$  (0.1% v/v formic acid) for 3 minutes. The sample was then transferred to an analytical column (Waters, BEH130  $\text{C}_{18}$ , 100 mm x 100  $\mu\text{m}$  i.d., 1.7  $\mu\text{m}$  particles) and eluted with a 1  $\mu\text{L}/\text{min}$  flow of formic acid. The peptides were detected on line using a mass spectrometer configured to operate in data dependent acquisition (DDA) mode using an MS function (*fullscan* of  $m/z$  200 to 2000), three MS/MS functions, and an external calibration curve function (*lockmass*). The spectrometer operated under the following parameters: 3 kV capillary voltage, 30 V cone voltage, 100  $^\circ\text{C}$  source temperature, 0.5 L/h nanoESI gas flow, 6 and 4 eV collision energy for the Trap and Transfer cells, respectively, and 1700 V detector. The MS (*fullscan*) and MS/MS (spectrum of ions fragmented by collision-induced dissociation) spectra were acquired at a rate of 1 spectrum/s. The instrument was calibrated before the analysis using phosphoric acid oligomers (0.5% v/v  $\text{H}_3\text{PO}_4$  solution in 50/50 v/v ACN/ $\text{H}_2\text{O}$ ) with an  $m/z$  from 90 to 1960. Argon at  $9.7 \cdot 10^{-3}$  mbar was used as the collision gas. The LC-MS runs were processed using ProteinLynx Global Server v.2.2 software (Waters) and all mass spectra were analyzed using a peak list format in the MASCOT program v.2.2 (Matrix Science, London, UK). The selected search parameter was: digestion ignoring trypsin until reaching a cleavage site, oxidation (M) or Methionine oxidation and, then, labelling them because there are oxidant conditions during sample preparation, a variable modification corresponding to those modifications that may or may not be present. For the specific case of M oxidation, MASCOT explains that, when a peptide contains more than 3 methionines, the coincidence with the experimental data has to be evaluated for that peptide that contains 0, 1, 2 and 3 residuals of oxidized Methionine. Precursor and fragment mass tolerance of  $\pm 0.1$  Da was also used as modification variables. For the searches, the SwissProt database (available since 1986) was used. Identification was based on the high specificity of the combination of the 20 essential aminoacids in sequences of peptides with 6 or more aminoacids. SwissProt is characterized because it contains entries that show contrasted, experimental evidence that has been manually reviewed by experts.

## RESULTS AND DISCUSSION

### Sample cleaning

Buffering and centrifugation have been previously implemented [15] for cleaning of intestinal mucus but there is no indication of the effectiveness of the process. In our study, Gram staining was employed as a simple methodology to verify the cleanliness of the sample. For proteomic studies of HIM, besides the quantity of sample, an effective cleaning process is of paramount importance. With this approach, only the mucus forming proteins are extracted, otherwise the sample will be contaminated with proteins from food, faeces, microbiota and external cells. A micrograph of Gram stained bacteria from HIM is shown in Figure 1A, before the sample had undergone the bacterial cleaning process; Figure 1B corresponds to conditions after performing the bacterial cleaning process, thus indicating a sample completely devoid of bacteria.

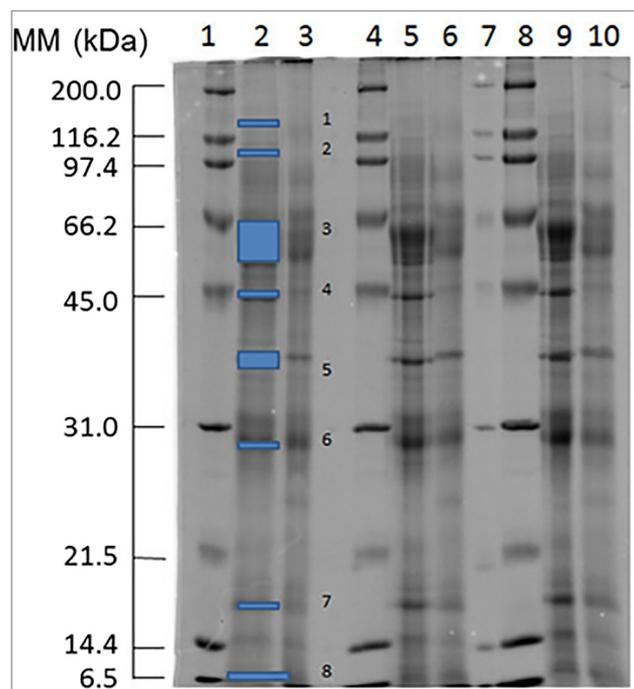


**Figure 1.** Elimination of bacteria from mucus sample. Gram stain of: A) the upper part of the supernatant, rich in bacteria, and B) the lower part of the supernatant, devoid of bacteria.

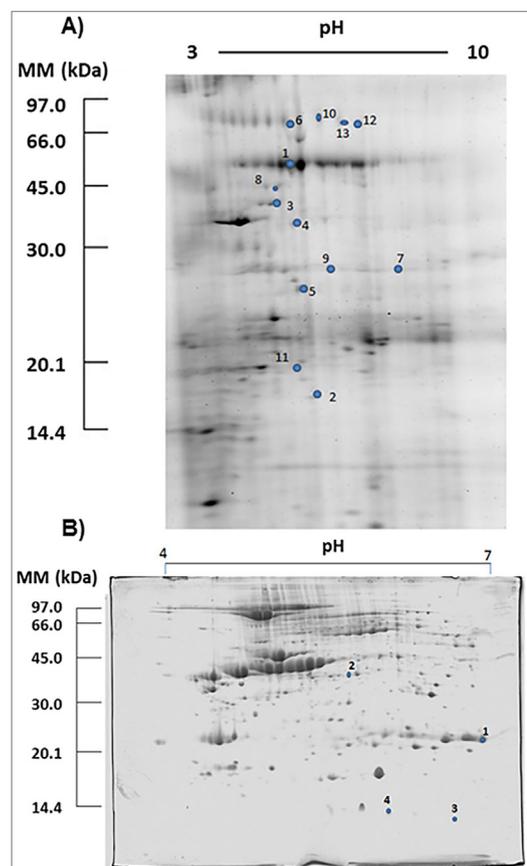
### **Sodium dodecyl sulphate polyacrylamide gel electrophoresis and two-dimensional gel electrophoresis**

According to the results obtained with the patients involved in this study, only 5 to 150  $\mu\text{L}$  of mucus sample can be obtained from a single person during colonoscopy. In fact, this amount of material is not sufficient for evaluation by 2-DE, even using staining reagents like colloidal coomassie, which improves the detection of proteins [24]. The protein quantification showed concentrations ranging from 0.65 to 1.12  $\mu\text{g}/\mu\text{L}$  [25]. On the other hand, a single session of transanal irrigation with one patient can provide from 0.5 to 5 mL of HIM, with protein concentrations ranging from 20.0 to 22.4  $\mu\text{g}/\mu\text{L}$  [25], after quantification. This concentration is enough for analysis by SDS-PAGE (Figure 2) or by 2-D PAGE (Figure 3). In this way, and using the material obtained through this sampling process, Figure 2 shows the SDS-PAGE for protein separation from mucus samples. Although this separation was carried out from three different patients, good similarities in terms of protein distribution are observed (similar protein weight), as well as the good efficiency in the sample cleaning process. After protein separation, all processes for protein identification through MS/MS were conducted (as indicated in section 'Mass spectrometry for protein identification') and supplementary material Table 1 summarizes the list of proteins identified using SDS-PAGE.

As the sampling strategy was efficient in terms of quantity of sample, a proteomic map was also obtained using 2-D PAGE based on samples from the same two patients. Then, gels presenting pH ranges from 3 to 10 and from 4 to 7 were used for protein identification according to procedures presented in section 'Two-dimensional gel electrophoresis'. However, 15 different protocols were tested with the objective of standardizing protein separations from HIM. All the protocols are visualized in Table I. In this way, using protocol A, only poor protein precipitation and partial dilution of the mucus were obtained and the gel presents a bad resolution with only 15 spots observed. On the other hand, protocol B presents a good resolution for protein identification. Even better was protocol C, because it presents the best resolution (ca. 220 spots, see Figure 3), requiring ca. 8 h for the run. Protocol G presents ca. 55 spots but a long time for running the gel. Similar behavior was observed by employing protocol I (only 12 spots), and too much streaking was observed in the gel. No or a minimal amount of proteins were observed employing the other protocols (D – F, H, J – O).



**Figure 2.** Sodium dodecyl sulphate polyacrylamide gel electrophoresis made with the mucus samples obtained from 3 patients. The gel is 17 cm and stained with SYPRO ruby. Lanes 1, 4, 7 and 8 are molecular weight standards (Bio-Rad, Hercules, California, USA). Lanes 2 and 3 are from volunteer 1, lanes 5 and 6 from volunteer 2, and lanes 9 and 10 correspond to volunteer 3. Blue bands indicate the bands analyzed by mass spectrometry.



**Figure 3.** Two-dimensional gel electrophoresis of human intestinal mucus. A) Gel with a pH range from 3 to 10, stained with SYPRO ruby. A total of 116 spots were visually identified with samples from one volunteer. B) Gel presenting a pH range from 4 to 7, stained with Coomassie blue. A total of 220 spots were visually identified with samples from one volunteer. Blue spots indicate the spots analyzed by mass spectrometry.

Protocol C was identified as the best for 2D-PAGE and the image of the best gel is shown in Figure 3. In fact, 116 spots were counted between 3 and 10, while 220 were present between 4 and 7 pH ranges. Additionally, the identified proteins can be visualized in Tables 2 and 3 in supplementary materials for gels from 3-10 and 4-7 pH ranges, respectively.

Finally, it is noteworthy that all steps involved in protein analysis must be conducted without stopping.

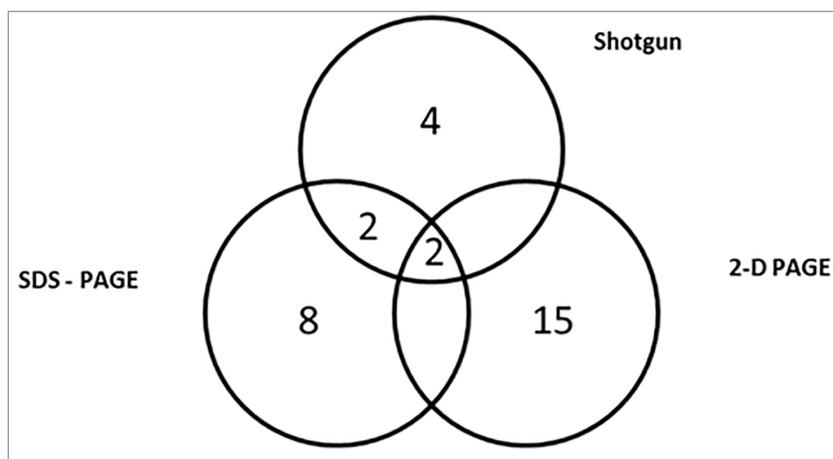
### Shotgun proteomics

Due to the difficulty in solubilizing some proteins, even using chemicals for improving their solubilization [1], as well as the inherent limitation of electrophoretic techniques in separating large molecular weight proteins, shotgun proteomics was selected for examining these problems.

### Identified proteins

#### Some preliminary considerations

Figure 4 presents a Venn diagram showing the overlap between the proteins identified by all the strategies employed in this work (*i.e.* SDS-PAGE; 2-D PAGE and shotgun), indicating the need for such strategies for identification due to their complementarities. It seems that a single technique is not suitable in order to study the largest number of the HIM proteins. It is important to comment that most of the identified proteins present in HIM are cellular proteins. This may be a consequence of the renovation of the cellular epithelium, also observed in the mouse colon mucus, or that such proteins have a specific function in the mucus. However, at present, there is insufficient information to provide a complete explanation of this fact. Besides cellular proteins, plasma proteins were also found. Some authors [17] suggest that the permeability of the intestine allows the passage of the extracellular liquid, or that such proteins are natural components of mucus and possibly have fundamental functions associated with the microbiota. As an example, in the cervical mucus these proteins are abundantly observed [17].



**Figure 4.** This Venn diagram shows the overlap of proteins identified for three separation/identification methods.

From the 31 proteins identified, 27 of them come from *Homo sapiens*, 3 from *Escherichia coli* and 1 from *Saccharomyces cerevisiae*. *E. coli* is a gram-negative intestinal bacterium which participates in the digestion process, although it can produce some intestinal problems. The *S cerevisiae* is an anaerobic unicellular fungus, which could be trapped in the mucus during ingestion of food. It is evident from the literature [12] and the HIM obtained through transanal irrigation, that it contains proteins from luminal (loose) and mucosa (tight) layers, as identified in Table II.

**Table II.** Proteins identified in the loose (L) and tight (T) layers of the colon mucus.

Protein	L	T	Remarks
Calcium activated	x	x	
Protein disulfide-isomerase	x	x	
ATP synthase subunit beta, mitochondrial precursor	x	x	
Actin, cytoplasmic 1		x	Identify from <i>Dreissena polymorpha</i>
IgA heavy 2	x	x	
Malate dehydrogenase mitochondrial	x	x	
Galectin-4			Protein found after the reduction treatment with GuHCl
Annexin 2	x	x	
Ig kappa light chain VLJ region	x		
Zymogen granule membrane protein 16 precursor	x	x	
Mitogen-activated protein kinase scaffold protein 1		x	
Hemoglobin subunit beta A	x	x	
Glutamine amidotransferase [ <i>Escherichia coli</i> ]	x	x	
Ubiquitin	x	x	
Polymeric immunoglobulin receptor	x		
MUC2	x	x	Protein found after the reduction treatment with GuHCl

### **About the identified proteins**

To make this topic more relevant, only those proteins which may have some clinical interest are discussed. Additionally, all the proteins identified were classified according to the biological processes in which they are involved.

#### *Transport and motility*

Proteins identified within this group were: hemoglobine subunit alfa, calcium-activated chloride channel protein 1, Actin cytoplasmic 1, and Zymogen granule membrane protein 16 or Secretory lectin ZG16. The calcium-activated chloride channel protein 1 is one of the proteins most present in the intestinal epithelium [26], and has an important function in the regulation of production and/or secretion of mucus when considering some health problems, such as metaplasia and cystic fibrosis, among others [27], [28]. Additionally, it also acts as a signaling agent between the intestinal environment and some immune perturbations, such as in inflammatory bowel disease. Finally, it is a tumor suppressor and induces the production of the Muc5C [29]. The actin cytoplasmic 1 and actin aortic smooth muscle are proteins involved in a diversity of cellular motility functions; for example, the actin cytoplasmic 1 participates in the secretion of gastric acid. The zymogen granule membrane protein 16 or secretory lectin ZG16 recognizes some pathogens' fungus, such as *Candida* and *Malassezia* species in the digestive system. The iron regulatory protein, Fur, is present in the microbiota, participating as a cofactor in iron (as Fe<sup>2+</sup>) transport.

#### *Immune response*

The Ig A, Ig alpha-2 heavy chain, polymeric immunoglobulin receptor, Ig kappa light chain VLJ, Ig lambda chain C, and Chain A Crystal structure of human calcineurin complexed with cyclosporin A were identified within this group. The chain A, crystal structure of human calcineurin complex with cyclosporine A and human cyclophilin, are calcium receptor enzymes which participate in different cellular regulatory ways. The cyclosporine is an immunosuppressor drug that complexes with cyclophilin. Both can also be associated with the calcineurin for inhibiting its catalytic activity [30]. It is curious that one of the patients remembers that when she was 15 years-old (she was 45 when the mucus sampling was carried out) the doctor prescribed cyclosporine for her. This result is something peculiar, because it highlights the protective importance of the intestinal mucus in trapping some waste substances.

#### *Metabolism*

The proteins identified within this group included: Glutamine-fructose-6-phosphate aminotransferase, and Bifunctional UDP-glucuronic acid decarboxylase/UDP-4-amino-4-deoxy-L-arabinose formyltransferase. The latter, for instance, confers the capacity of accepting antibiotics by mutations to bacteria that hampers the bond between the antibiotic and the protein, or by gene amplification, which codifies the protein. It also participates in lipid metabolism.

#### *DNA damage response*

Polyubiquitin, ubiquitin, and TIP41-like protein were those identified associated with this group. Regarding the TIP41-like protein, it can be an allosteric regulator of serine / threonine protein phosphatase 2A. Isoforms may play a role in regulating the signaling pathway ATM / ATR and control in DNA replication and repair.

#### *Cellular response*

Within this group, mitogen-activated protein kinase scaffold protein 1 or Regulator complex protein LAMTOR3 and UPF0539 protein C7orf59, also known as Regulator complex protein LAMTOR4, were the proteins identified. LAMTOR3 and LAMTOR4 participate in the regulator complex which functions as a factor in guanine nucleotide exchange and activation of small GTPases Rag. In a recent study [31], LAMTOR3 was identified as participating in cellular responses induced by gastrin which is involved, in turn, in proliferation and homeostasis of gastric mucosa.

### *Angiogenesis*

Within this group, the following proteins were identified: ATP synthase subunit beta, mitochondrial precursor, and annexin A2. This last protein is involved in the regulation of epithelial cell migration. When an intestinal lesion happens, these cells migrate to seal the wound. Its action is carried out by controlling the levels of B1 integrin protein, which is responsible for allowing the interaction between the epithelial cells and the matrix [32]. Annexin A2 was evaluated as a biomarker in serum for early diagnosis of hepatocellular carcinoma, which is related to the virus for hepatitis B [33].

### *Enzymatic activity*

Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform 1, and Chain A, crystal structure of the 20s proteasome were the proteins constituting this group. The last one is a proteinase complex whose function is to carry out the proteolysis of those unnecessary or damaged proteins localized in the cytoplasm or in the nucleus.

### *Cellular differentiation and growth*

Within this group the nucleolar protein 8 was found. This protein plays an important role in the apoptosis of gastric cancer cells [34]. It can also be involved in the genic expression regulation at post-transcriptional level or in the biogenesis of ribosomes in cancer cells.

### *Cell adhesion*

Galectin-4, and Mucin 2 (MUC2) were identified in this group, Mucin, the target of this work, was described and commented on in the introduction. Galectin-4 has a similar function to galectin-3, which is related to human colon cancer [35]. A higher abundance of galectin-3 was related to the malignant behavior of colon cancer cells [34]. However, it is still not clear how this protein affects the intracellular signal pathway.

### *Other functions*

Within this group the ferric uptake regulator was identified. This protein [*Escherichia coli* O157: H7EDL933] is grouped with the biological process of transcription and uses Fe<sup>2+</sup> as a cofactor that binds to operator repressed genes and regulates the expression of several proteins of the outer membrane.

### *Unknown functions*

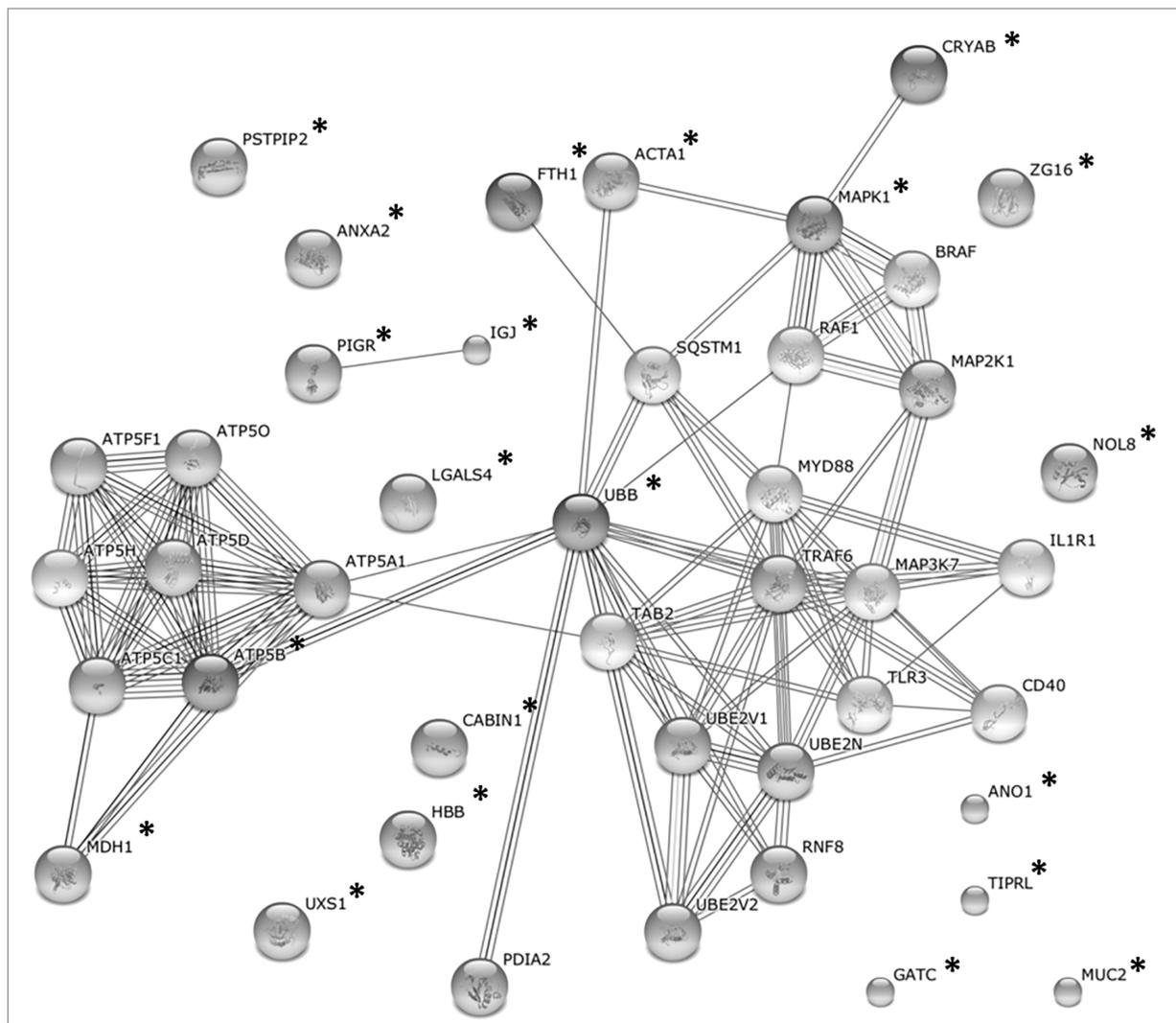
One protein was identified inside this group: the Sccl+ suppressor protein [*Saccharomyces cerevisiae*] is a dominant suppressor of cycloheximide resistance.

### ***Interactions between some proteins identified***

When considering protein-protein interactions (Figure 5) using the STRING program, from the total amount of proteins identified, 23 are correlated. From these proteins, a good correlation between their functions is then noted, and at least two major clusters observed. Firstly, in the center of Figure 5 is Ubiquitin, a protein that mediates transcriptional activation of target genes and plays an important role in controlling the progress of cell cycle and differentiation, in the error-free DNA repair pathway, and in the survival of cells after DNA damage. Secondly, on the left side of this figure, both ATP synthase (ATP5B) and malate dehydrogenase (MDH1) were correlated, belonging to a cluster in which a diversity of ATP synthases is present. In fact, the ATP synthases are involved in the production of ATP from ADP in the presence of a proton gradient across the membrane, which is generated by electron transport complexes in the respiratory chain. As this enzyme is involved in energy production, its correlation with the synthesis of glucose through MDH is then correct.

Some other proteins identified (see Tables 1-4 supplementary material), such as annexin A2 (ANXA2), Galactin 4 (LGALS4), calcium binding protein 1 (CABIN1) and others, are not connected to each other, but

play important roles regarding biological processes, as described in section about the identified proteins. Looking at the right side of Figure 5, another cluster is noted in which mitogen-activate protein kinase 1 (MAPK1), crystalline alpha B (CRYAB), ferritin (FTH1) and actin (ACTA1) are present. The presence of such proteins denotes the link between the clusters: while the cluster on the left side is involved in energy storage, the cluster on the right hand side is involved in energy consumption due to the work developed by the biochemical machinery in relation to cell growth, adhesion, survival (MAPK1), motility (ACTA1) and iron uptake (FTH1).



**Figure 5.** Protein interaction network generated and visualized with STRING 9.1 program for proteins identified in our study. Only those proteins identified in this study are named and marked with an asterisk in this Figure. The others are inserted into the program in order to compose the network. Center: UBB - Ubiquitin B; Left side: ATP5B – ATP synthase; MDH1- Malate dehydrogenase; UXS1 – UDP glucuronic acid decarboxylase; HBB – hemoglobin; CABIN1 – calcinerium binding protein 1; LGALS4 – Lectin; PIGR – Polymeric immunoglobulin receptor; IGJ – Immunoglobulin J peptide; ANXA2 – Annexin A2; PSTPIP2 – Proline-serine-threonine phosphatase; Right side: GATC – glutamyl-tRNA amidittransferase; MUC2 – Mucin 2; TIPRL – TIP41-like protein; ANO1 – calcium activated chloride channel; NOL8 – Nucleolar protein 8; ZG16 – zymogen granule protein 16; CYAB – Crystallin alpha B; MAPK1 – mitogen-activated protein kinase; ACTA 1 – actin; FTH1 – ferritin.

Most probably this “hard work” developed by the organism may give rise to some stress conditions, thus explaining the presence of CRYAB in this cluster, as one of its functions is just to prevent aggregation of various proteins under a wide range of stress conditions. Some other identified proteins, such as Mucin 2 (MUC2), zymogen granule protein 16 (ZG16), calcium-activated chloride channel (ANO1) and others, are also present, but apparently with no correlation to that cluster. However, some of them are proteins which also participate in the process of energy consumption, such as gastrointestinal smooth muscle contraction (ANO1), protein trafficking (ZG16), and exclusion of bacteria from the inner mucus layer (MUC2), among others.

Finally, according to the analysis of the identified proteins and their correlations, some of the most important biological processes are then identified (as indicated in the section about the identified proteins) and correlated with each other.

## CONCLUSIONS

The main objectives of this study were successfully achieved. By using a sampling method based on transanal irrigation and some strategies involved in analytical proteomics, the first, and most detailed, proteomic human intestinal mucus map was obtained. It is necessary to note the extreme difficulty in working with such samples, especially when attaining accurate identification of those proteins belonging exclusively to HIM. In this way, all analytical steps for treating the samples were efficiently performed, producing a selective proteomic map, free of those proteins from food, faeces and microbiota. With all the strategies applied here, a good number of proteins was then identified using SDS-PAGE, 2-D PAGE and shotgun analysis, and 9 of them related with other studies, playing an important role in some diseases, particularly considering those related to the digestive system. The majority are secreted proteins, making their purification possible for future studies. This study contributes to a better understanding of the HIM proteome through a methodology for obtaining, treating, separating and identifying its proteins. Additionally, it provides a methodology for analysis of the HIM proteins, without protein separation. This includes their extraction, desalination, depletion and digestion for MUC2 identification, the main protein of HIM, generally separated by SDS-agarose or ultracentrifugation [36]. In summary, the content of this work should facilitate further studies of complex bodily fluids, such as HIM, as well as the identification of the differential proteins in a variety of pathological conditions that could be used as therapeutic agents or biomarkers.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgments

The authors thank: a) In Colombia: The University of Caldas, Cenicafé and COLCIENCIAS, for their partial financial support and permission to use their equipment and facilities; b) In Brazil: “Fundação de Amparo a Pesquisa do Estado de São Paulo” (FAPESP, São Paulo, Brazil), the “Conselho Nacional de Desenvolvimento Científico e Tecnológico” (CNPq, Brasília, Brazil), and the “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior” (CAPES, Brasília, Brazil) for financial support and fellowships.

## REFERENCES

1. Johansson, M. E.; Phillipson, M.; Petersson, J.; Velcich, A.; Holm, L.; Hansson, G. C. *Pro. Natl. Acad. Sci.*, **2008**, *105* (39), pp 15064–15069 (<https://doi.org/10.1073/pnas.0803124105>).
2. Gum, J. R.; Hicks, J. W.; Toribara, N. W.; Siddiki, B.; Kim, Y. S. *J Biol Chem.*, **1994**, *269* (4), pp 2440–2446.
3. Matsuo, K.; Ota, H.; Akamatsu, T.; Sugiyama, A.; Katsuyama, T. *Gut.*, **1997**, *40* (6), pp 782–789.
4. Sommer, F.; Adam, N.; Johansson, M. E. V.; Xia, L.; Hansson, G. C.; Bäckhed, F. *PLoS One*, **2014**, *9* (1), e85254 (<https://doi.org/10.1371/journal.pone.0085254>).
5. Blaut, M.; Collins, M. D.; Welling, G. W.; Doré, J.; Van Loo, J.; de Vos, W. M. *Br. J. Nutr.*, **2002**, *87* (S2), pp S203–S211 (<https://doi.org/10.1079/BJN/2002539>).

6. Tennyson, C. A; Friedman, G. *Curr Opin Endocrinol Diabetes Obes.*, **2008**, *15* (5), pp 422–427.
7. Tamboli, C. P; Neut, C.; Desreumaux, P.; Colombel, J. F. *Gut.*, **2004**, *53* (1), pp 1–4.
8. Johansson, M. E. V.; Sjövall, H.; Hansson, G. C. *Nat Rev Gastroenterol Hepatol.*, **2013**, *10* (6), pp 352–361 (<https://doi.org/10.1038/nrgastro.2013.35>).
9. Ouwerkerk, J. P.; De Vos, W. M.; Belzer, C. *Best Pract Res Clin Gastroenterol.*, **2013**, *27* (1), pp 25–38.
10. Bäckhed, F.; Ley, R. E.; Sonnenburg, J. L.; Peterson, D.; Gordon, J. I. *Science*, **2005**, *307* (5717), pp 1915–1920 (<https://doi.org/10.1126/science.1104816>).
11. Hansson, G. C. *Curr Opin Microbiol.*, **2012**, *15* (1), pp 57–62 (<https://doi.org/10.1016/j.mib.2011.11.002>).
12. Johansson, M. E. V.; Thomsson, K. A.; Hansson, G. C. *J Proteome Res.*, **2009**, *8* (7), pp 3549–3557.
13. Du, L. Y.; Qian, D. W.; Shang, E. X.; Jiang, S.; Liu, P.; Guo, J. M.; Su, S. L.; Duan, J.; Xu, J.; Zhao, M. *Anal Methods.*, **2015**, *7* (13), pp 5574–5582 (<https://doi.org/10.1039/C5AY00931F>).
14. Licea-Perez, H.; Junnotula, V.; Zohrabian, S.; Karlinsey, M. *Anal Methods*, **2016**, *8* (15), pp 3023–3033.
15. Hamer, H. M.; Jonkers, D.; Loof, A.; Vanhoutvin, S. A. L. W.; Troost, F. J.; Venema, K.; Kodde, A.; Koek, G.; Schipper, R.; van Heerde, W.; et al. *Dig Liver Dis.*, **2009**, *41* (8), pp 559–564.
16. Rodríguez-Piñeiro, A. M.; Post S, V. D.; Johansson, M. E. V.; Thomsson, K. A.; Nesvizhskii, A. I.; Hansson, G. C. *J Proteome Res.*, **2012**, *11* (3), pp 1879–1890 (<https://doi.org/10.1021/pr2010988>).
17. Panicker, G.; Ye, Y.; Wang, D.; Unger, E. R. *Clin Proteomics.*, **2010**, *6* (1-2), pp 18–28.
18. Dellios, N. L.; Lappas, M.; Young, I. R.; Palliser, H. K.; Hirst, J. J.; Oliva, K.; Ayhan, M.; Rice, G. *Eur J Obs. Gynecol Reprod Biol.*, **2010**, *153* (1) pp 16–22 (<https://doi.org/10.1016/j.ejogrb.2010.06.014>).
19. Panicker, D. R.; Lee, E. R.; Unger, E. *J Proteomics.*, **2009**, *71* (6) pp 637–646.
20. Bovee-Oudenhoven, I.; Termont, D.; Dekker, R.; Van der Meer, R. *Gut.*, **1996**, *38* (1) pp 59–65.
21. Nice, E. C., Fabri, L.; Whitehead, R. H.; James, R.; Simpson, R. J.; Burgess, A. W. *J Biol Chem.*, **1991**, *266* (22) pp 14425–14430.
22. Kay, R.; Barton, C. *Rapid Commun Mass Spectrom.*, **2008**, pp 3255–3260.
23. Ribeiro, D.; Cota, J.; Alvarez, T. M.; Brüchli, F.; Bragato, J.; Pereira, B.; Pauletti, B.; Jackson, G.; Pimenta, M.; Murakami, M.; et al. *PLoS One.*, **2012**, *7* (12), e50571 (<https://doi.org/10.1371/journal.pone.0050571>).
24. Dyballa, N.; Metzger, S. *J. Vis. Exp.*, **2009**, (30), e1431 (<https://doi.org/10.3791/1431>).
25. Gómez-Buitrago, P-A.; González-Correa, C-A.; Santacoloma-Osorio, M.; Taborda-Ocampo, G.; Zezzi-Arruda, M. A. *Rev Colomb Quim.*, **2014**, *43* (3), pp 5–10 (<http://dx.doi.org/10.15446/rev.colomb.quim.v43n3.53586>).
26. Nyström, E.; Arike, L.; Ehrencrona, E.; Hansson, G. C.; Malin, E. V. *J. Biol. Chem.*, **2019**, *1*, pp 1–25.
27. Hoshino, M.; Morita, S.; Iwashita, H.; Sagiya, Y.; Nagi, T.; Nakanishi, A.; Ashida, Y.; Nishimura, O.; Fujisawa, Y.; Fujino, M. *Am J Respir Crit Care Med.*, **2002**, *165* (8), pp 1132–1136.
28. Qiu, J.; Heller, J.; Guo, X.; Chen, Z.; Fish, K.; Fu, Y.; Zhou, L. *Immunity.*, **2012**, *36* (1), pp 92–104.
29. Bustin, S.; Li, S.; Dorudi, S. *DNA Cell Biol.*, **2001**, *20* (6), pp 331–338.
30. Jin, L.; Harrison, S. *Proc Natl Acad Sci.*, **2002**, *99* (21) , pp 13522–13526.
31. Steigedal, T.; Prestvik, W.; Selvik, L.; Fjeldbo, C.; Bruland, T.; Laegreid, A.; Thommesen, L. *In Vitro Cell. Dev. Biol. - Animal*, **2013**, *49* (3), pp 162–169 (<https://doi.org/10.1007/s11626-013-9588-2>).
32. Rankin, C.; Hilgarth, R. S.; Leoni, G.; Kwon, M.; Den Beste, K.; Parkos, C.; Nusrat, A. *J Biol Chem.*, **2013**, *288* (21), pp 15229–15239 (<https://doi.org/10.1074/jbc.M112.440909>).
33. Sun, Y.; Gao, G.; Cai, J.; Wang, Y.; Qu, X.; He, L.; Liu, F.; Zhang, Y.; Lin, K.; Ma, S.; et al. *Carcinogenesis*, **2013**, *34* (3), pp 595–604.
34. Jinawath, N.; Furukawa, Y.; Nakamura, Y. *Cancer Sci.*, **2004**, *95* (5), pp 430–435.
35. Kim, D. W.; Kim, K. H.; Yoo, B. C.; Hong, S. H.; Lim, Y. C.; Shin, Y. K.; Park, J. *Cancer Sci.*, **2008**, *99* (10), pp 1884–1891.
36. Holmén, J. M.; Karlsson, H.; Sjövall, H.; Hansson, G. C. *Glycobiology*, **2009**, *19* (7), pp 756-766 (<https://doi.org/10.1093/glycob/cwp048>).

## Supplementary Material

Table 1. Proteins identified of HIM through SDS-PAGE

Line	Protein name	Access	Theoretical MM(Da)	Experimental MM(Da)	Score	% Coverage	Peptides	# Peptides
4	Actin, cytoplasmic 1; AltName: Full=Beta actin.	gij113271	42169	42200	868	55	K.AGFAGDDAPR.A, R.AVFPISIVGRPR.H, R.HQGVMVGMGQK.D, K.DSYVGDEAQS.K.R, K.DSYVGDEAQS.K.R.G,K.IWHHTFYNELR.V, R.VAPEEHPVLLTEAPLNPK.A, R.TTGIVMDSGDGVTHTVPIYEGYALPHAILR.L, R.LDLAGRDLTDYLMK.I, R.DLTDYLMK.I,K.ILTERGYSFTTTAER.E, R.GYSFTTTAER.E,R.GYSFTTTAEREIVR.D, K.SYELPDGQVITIGNER.F, K.DLYANTVLSGGTTMYPGIADR.M, R.MQKEITALAPSTMK.I, K.EITALAPSTMK.I,K.IIAPPER.K, K.IIAPPERK.Y,K.QEYDESGPSIVHR.K, K.QEYDESGPSIVHRK.C	21
5	Annexin A2.	gij18645167	38784	38500	288	24	K.AYTNFDAER.D,R.DALNIETAIK.T, K.GVDEVTIVNILTNR.S, K.SALSGHLETVILGLLK.T, K.TPAQYDASELK.A,R.TNQELQEINR.V, K.SYSPYDMLESIR.K	7
3	ATP synthase subunit beta, mitochondrial precursor.	gij32189394	56525	56000	1254	59	K.VLDSGAPIK.I, K.TVLIMELINNVAK.A, K.IGLFGGAGVGK.T,K.IPVGPETLGR.I, K.VVDLLAPYAK.G,R.TIAMDGTEGLVR.G, R.IMNVIGEPIDER.G,K.AHGGYSVFAGVGER.T, R.FTQAGSEVSALLGR.I,R.VALTGLTVAEYFR.D, K.GFQQILAGEYDHLPEQAFYMGPIEEAVAK.A	13
2	Calcium-activated chloride channel protein 1.	gij4585469	100894	103000	212	10	K.LAEYGPQGR.A, R.IHLTPDFIAGK.K, K.LAEYGPQGR.A, R.WGVFDEYNNDEK.F, K.TTPMTTQPPNPTFSLQLIQGR.I, K.MAYLQIPGIAK.V, K.TVTLELLDNGAGADATK.D	9
5	Galectin-4.	gij5453712	36034	38500	514	30	R.FFVNFVVGQDPGSDVAFHFNPR.F, R.FDGWDK.V,K.VVFNLTQGGK.W, K.VVNGNPFYEGYHR.L,R.LQGGLTAR.R, K.SFAINFK.V,K.VGSSGDIALHINPR.M, K.KITHNPFPGQFFDLSIR.C, K.IITHNPFPGQFFDLSIR.C	10

**Table 1.** Proteins identified of HIM through SDS-PAGE (Continuation)

Line	Protein name	Access	Theoretical MM(Da)	Experimental MM(Da)	Score	% Coverage	Peptides	# Peptides
5	Ig alpha-2 heavy chain constant region.	gij184761	37226	38500	587	20	R.DASGATFTWTPSSGK.S, R.GFSPKDLVLR.W,R.WLQGSQELPR.E, R.WLQGSQELPR.E, R.WLQGSQELPREK.Y, R.WLQGSQELPREK.Y,R.EKYLWASR.Q, K.YLTWASR.Q, R.VAAEDWK.K, K.YLTWASRQEPSQGTTFVAVTSILR.V, R.QEPSQGTTFVAVTSILR.V	11
6	Ig kappa light chain VLJ region.	gij21669449	28825	29000	65	18	R.TVAAPSVFIFPPSDEQLK.S, R.TVAAPSVFIFPPSDEQLK.S, K.DSTYLSSTLTLSK.A, K.VYACQVTHQGLSSPVTK.S	4

**Table 2.** Proteins identified of HIM from 2-D PAGE, pH 3-10

Spot	Protein name	Access	Theoretical pI/MM(Da)	Experimental pI	Experimental MM(Da)	Score	% Coverage	Peptides	# Peptides
10	Bifunctional UDP-glucuronic acid decarboxylase [Escherichia coli]	gij15802804	6.34/74859	6.40	74000	110	5	R.APLNWVLVNGETETGVTLHR.M, R.VLILGVNGFIGNHLTER.L	3
3	Chain A, Crystal structure of human calcineurin complexed with cyclosporin A and Human cyclophilin	gij24987750	5.26/43392	5.30	40000	101	8	R.IITEGASILR.Q, R.YLFLGDYVDR.G	2
5	Chain A, Crystal structure of the 20s proteasome From Yeast At 2.4 Angstroms Resolution	gij24987750	5.90/27440	5.90	28000	74	10	R.TIGMVVNGPIPDAR.N, R.MANLSQIYTQR.A	2
11	Ferric uptake regulator [Escherichia coli O157:H7 EDL933]	gij15800386	5.68/17016	5.60	18000	93	16	R.VLNQFDDAGIVTR.H, K.VIEFSDDSIAR.Q	2
6	Glutamine amidotransferase [Escherichia coli]	gij290577	5.52/67065	5.52	68000	69	4	R.DVAEILLEGLR.R, R.FIFLEEGDIAEITRR.S	2

**Table 2.** Proteins identified of HIM from 2-D PAGE, pH 3-10 (Continuation)

Spot	Protein name	Access	Theoretical pI/MM(Da)	Experimental pI	Experimental MM(Da)	Score	% Coverage	Peptides	# Peptides
2	Hemoglobin subunit beta A	gi 122539	6.36/16012	6.35	16200	331	55	K.VHVDEVGGEALGR.L, R.FFESFGDLSTADAVMNNPK.V, K.VLDSFSDGMK.H, K.HLDDLK.G, K.LHVDPENFK.L, K.LLGNVLVVVLAR.N, K.EFTPVLQADFQK.V	7
4	Ig A-2 heavy chain	gi 184761	5.71/37226	5.60	37000	132	17	R.DASGATFTWTPSSGK.S, K.SAVQGPPER.D, R.WLQGSQELPR.E, K.YLTWASR.Q, R.QEPSQGTTTFAVTSILR.V	6
7	Ig kappa light chain VLJ region	gi 21669449	8.15/28885	8.11	29000	76	13	K.VDNALQSGNSQESVTEQDSK.D, K.VYACQVTHQGLSSPVTK.S	2
13	Polyubiquitin	gi 2627129	7.13/68448	7.10	68000	193	6	MQIFVK.T, R.TLSDYNIQK.E , K.ESTLHLVLR.L,K. TITIEVEPSDTIENVK.A	4
9	Sc11+ suppressor protein [Saccharomyces cerevisiae]	gi 172546	6.7/30552	6.70	29000	74	9	R.TIGMVVNGPIPDAR.N, R.MANLSQIYTQR.A	2
1	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform 1	gi 6715568	5.58/59346	5.48	59000	313	19	K.EVFDNDGKPR.V , R.LEESVALR.I, R.IITEGASILR.Q, K.LFEVGGSPANTR.Y, R.YLFLGDYVDR.G , K.TQEHFTHNTVR.G, K.TQEHFTHNTVR.G, R.AHEAQDAGYR.M, R.DAMPSDANLNSINK. AR.EEESVLTLLK.G, K.YENNVMNIR.Q	11
8	Unnamed protein product [Homo sapiens]	gi 34039	5.04/44079	5.00	49000	269	20	R.ALEAANGELEVK.I,R. DYSHYYTTIQDLR.D, K.ILGATIENSR.I,R.IVLQIDNAR.L, K.FETEQLR.M ,R.MSVEADINGLR.R, R.VLDELTLAR.T, R.TDLEMQIEGLK.E	8

**Table 2.** Proteins identified of HIM from 2-D PAGE, pH 3-10 (Continuation)

Spot	Protein name	Access	Theoretical pI/MM(Da)	Experimental pI	Experimental MM(Da)	Score	% Coverage	Peptides	# Peptides
12	Unnamed protein product [Homo sapiens]	gi 189054178	7.62/66154	7.39	67000	341	12	R.SLVNLGGSK.S,R. SLVNLGGSKSISIVAR.G, K.SLNNQFASFIDK.V,R. FLEQQNQVLQTK.W, R.TNAENEFVTIK.K,R. SLDLDSIIAEVK.A, K.AQYEDIAQK.S,K.YEELQITAGR.H	9

**Table 3.** Proteins identified of HIM from 2-D PAGE, pH 4-7

Spot	Protein name	Access	Theoretical pI/MM(Da)	Experimental pI/MM (Da)	Score	% Coverage	Peptides	# Peptides
1	Alpha-crystallin B chain	CRYAB_HUMAN	6.76/20146	6.80/20150	104	35	R.RPFFPFHSPSR.L, R.FSVNLQVH, R.QDEHGFISR.E, K.YRIPADVDPDLTITSSLDGVLTVNGPR.K, R.IPADVDPDLTITSSLDGVLTVNGPR.K, R.TIPITR.E, R.IPADVDPDLTITSSLDGVLTVNGPR.K, R.IPADVDPDLTITSSLDGVLTVNGPR.K	11
2	TIP41-like protein	TIPRL_HUMAN	5.6/31652	5.75/34000	470	43	K.LADELHMPSLPEMMFGDNVLR.I, R.IQHSGSGFIEFNATDALR.C, K.GTLLGESLK.L, R.VMPSSFLLLR.F, R.IDGVLIR.M, R.LYHEADK.T, R.LYHEADK.T, R.LYHEADKTYMLR.E, K.TYMLR.E, K.ISSLMHVPPSLFTEPNEISQYLPK.E, K.LIFPER.I,R.IDPNPADSQK.S, R.IDPNPADSQK.S	12
3	Ubiquitin	UBIQ_HUMAN	6.56/8560	6.55/10100	19	32	K.TITLEVEPSDTIENVK.A, K.EGIPPDQQR.L, K.EGIPPDQQR.L, K.EGIPPDQQR.L	4
4	UPF0539 protein C7orf59	CG059_HUMAN	6.07/10792	5.91/13400	97	26	R.GMNVPFK.R, R.GMNVPFKR.L, R.LSVVFGHEHTLLVTVSGQR.V	4

**Table 4.** Proteins identified of HIM from shotgun analyses

Protein name	Access	Theoretical pI/ MM(Da)	Score	% Coverage	Peptides	# Peptides
Calcium-activated chloride channel regulator	CLCA1_HUMAN	5.97/100847	63	8	K.TKADYVRPK.L, K.NADVLVAESTPPGNDEPYTEQMGNCGEKGER.I, R.IHLTPDFIAGKK.L, R.AFVHEWAHLR.W, R.WGVFDEYNNDEKFYLSNGR.I	6
Galectin-4	LEG4_HUMAN	9.21/36032	136	5	R.FFVNFVVGQDPGSDVAFHFNPR.F, R.FDGWDKVVFNLTQGGK.W, K.VGSSGDIALHINPR.M, K.VGSSGDIALHINPR.M	5
IgA-1 chain C region	IGHA1_HUMAN	6.08/38486	148	23	R.DASGVTFTWTPSSGK.S, K.TFTCTAAYPEK.T, R.WLQGSQELPR.E, R.EKYLTWASR.Q, R.QEPSQGTTFFAVTSILR.V, K.KGDTFSCMVGHEALPLAFTQK.T	6
IgA-2 chain C region	IGHA2_HUMAN	5.71/37301	148	38	K.HYTNPSQDVTVPCPVPPPPCCHPR.L, R.DASGATFTWTPSSGK.S, R.DLCGCYSVSSVLPGCAQPWNHGETFTCTAAHPELK.T, R.WLQGSQELPR.E, R.EKYLTWASR.Q, R.QEPSQGTTFFAVTSILR.V, K.KGDTFSCMVGHEALPLAFTQK.T	9
Ig kappa chain C region	IGKC_HUMAN	5.58/11773	229	53	K.VQWKVDNALQSGNSQESVTEQDSK.D, K.VDNALQSGNSQESVTEQDSKDYSLSTLTLK.A, K.DSTYLSSTLTLK.A, K.VYACEVTHQGLSSPVTK.SK.HKVYACEVTHQGLSSPVTK.S, K.HKVYACEVTHQGLSSPVTK.S	8
Ig lambda chain C regions	LAC_HUMAN	6.92/11401	36	49	K.ADSSPVKAGVETTTSPK.Q, K.AGVETTTSPK.Q, K.QSNKYAASSYLSLTPEQWK.S, K.YAASSYLSLTPEQWK.S, R.SYSCQVTHEGSTVEK.T	6
Muc-2	MUC2_HUMAN	5.49/552218	28	<1	D RPIYEEDLKK, K.TVHMMPMQVQVQVNR.Q, K.GVCVHGNAEYQPGSPVYSSK.C	4
Polymeric immunoglobulin receptor	PIGR_HUMAN	5.58/84429	95	15	R.GGCITLISSEGYVSSK.Y,R.TVTINCPFKTENAQK.R, K.NADLQVLKPEPELVYEDLRGSVTFHCALGPEVANVAK.F, K.RAPAFEGR.I, R.ILLNPQDKDGSFVITGLR.K, K.QGHFYGETAAVYVAVEERK.AK.QGHFYGETAAVYVAVEERK.A,	8