Università degli Studi di Cagliari

Philosophy Doctor

in

Life, Environmental and Drug Sciences Biomedical Curriculum

XXIX Cycle

Identification of potential disease biomarkers in tissues and saliva by an integrated Top-down and Bottom-up proteomic approach.

BIO/10

Presented by: Doctor Valentina Piras
PhD coordinator: Professor Enzo Tramontano
Tutor: Professor Tiziana Cabras

Final exam academic year 2015 – 2016

Thesis presented in the examination session March to April 2017

“If we knew what it was we were doing, it would not be called research”

Albert Einstein

Contents: Paragraph

- **Abstract**
- **Introduction** **1.0**
  - Proteome and proteomics 1.1
  - Top-down and Bottom-up integrated platforms for proteomic analysis in health and disease 1.2
  - Quantification of proteins and peptides 1.3
  - Proteomics in biomarkers discover. 1.4
- **Objectives of the thesis** **2.0**

- **PART-1**
  Top-Down Proteomics Characterization of the acidic soluble fraction of human saliva revealed significant changes in patients affected by periodic fever syndromes
  - Human Saliva 3.0
  - Top-down proteomics of human saliva 3.1
  - Histatin 3.2
  - Proline Rich-proteins 3.3
  - Statherin and P-B peptide 3.4
  - Salivary ("S"-type) and other cystatins 3.5
  - Proteins of the S100 family 3.6
  - α-defensins 3.7
  - β-thymosins 3.8
  - Antileukoproteinase 3.9
  - Periodic fever syndrome 4.0
  - Familial Mediterranean Fever 4.1
  - Genetics 4.2
  - Pathogenesis 4.3
  - Clinical Manifestations 4.4
  - Abdominal attacks 4.5
  - Arthritis 4.6
  - Myalgia 4.7
  - Skin Manifestations 4.8
  - Isolated febrile attacks 4.9
  - Amyloidosis 4.10
- **Objectives of the study** **5.0**
- **Experimental** **6.0**
  - Materials 6.1
  - Subjects 6.2
  - Clinical data 6.3
  - Salivary sample collection 6.4
  - Top-down proteomics analysis by Low Resolution HPLC-ESI-MS 6.5
  - Quantification. 6.6
  - Statistical analysis 6.7
- **Results** **7.0**
  - Correlation analysis 7.1

- Quantification of protein and peptides by XIC procedure 7.2
- aPRPs 7.3
- Statherin and PB peptide 7.4
- Histatin 7.5
- Cystatins 7.6
- S100A8, S100A9, and antileukoproteinase 7.7
- α-defensins 7.8

- **Discussion** **8.0**
- **Conclusions** **9.0**
- **Bibliography**

- **PART 2**
  TOP-Down and Bottom-up proteomics approaches on the fraction of low molecular weight of protein extracts from human colonic mucosa
- **Introduction**
  - Colorectal cancer 1.1
  - Dukes classification 1.2
  - Biomarker discovery for colo-rectal cancer 1.3
  - Thymosin β-4 and β-10 as potential biomarker in colo-rectal cancer 1.4
- **Objectives of the study** **2.0**
- **Experimental** **3.0**
  - Reagents and apparatus 3.1
  - Samples and Study subjects 3.2
  - Protein extraction 3.3
  - Ultrafiltration of the raw extract 3.3 a)
  - Hydro-organic treatment of the raw extract 3.3 b)
  - Low resolution HPLC-ESI-MS analysis 3.4
  - Quantification and statistical analysis 3.5
  - RP-HPLC/high-resolution-ESI-MS/MS analysis 3.6
  - Trypsin digestion 3.7
- **Results** **4.0**
  - Top-down and Bottom-up characterization of proteins and peptides 4.1
  - Thymosins β 4.2
  - Pro-thymosin α and parathymosin 4.3
  - Ubiquitin 4.4
  - SH3BP-1 protein 4.5
  - FABP1 4.6
  - Carbonic anhydrase 1 4.7
  - 9955 uncharacterized protein 4.8
  - Quantification of peptides and proteins in tumor and healthy intestinal mucosa 4.9
  - Comparison between superficial, deep tumor and healthy mucosa. 4.10
  - Tβ4 and Tβ10 modified proteoforms 4.11
  - Other peptides and proteins detected 4.12

- Correlation with Dukes stadium 4.13
- Thymosins β4 and β10 4.14
- Ubiquitin 4.15
- SH3BP1 4.16
- Carbonic Anhydrase 1 4.17
- FABP1 4.18
- 9955Da protein 4.19

- **Discussion** **5.0**
  - Thymosin β4 and β10 in colorectal cancer 5.1
  - Pro-thymosin α and Parathymosin 5.2
  - Carbonic Anhydrase 5.3
  - SH3 domain Binding Protein 1 5.4
  - Ubiquitin 5.5
  - Fatty acid binding protein 5.6
- **Conclusions** **6.0**
- **Bibliography**
- **Supporting information**
- **Supporting Information Section I**
- **Supporting Information Section II**

Abstract

This thesis has been focused on the proteomic characterization of complex protein mixtures from different biological matrices. The first study has been performed in order to characterize the acid soluble fraction of human saliva in patients affected by Periodic Fever Syndromes. In the second study a top-down and bottom-up proteomic approach has been applied in order to identify the fraction of low molecular weight of protein extracts from Human Colonic Mucosa.

Top-down proteomics has been applied to investigate qualitative and quantitative modifications of the acidic soluble salivary proteome/peptidome in patients affected by auto-inflammatory periodic fever syndromes associated to mutations of pyrin gene. Recurrent episodes of fever are accompanied by abdominal, chest and joint pain, swelling, and aphthous-like oral ulceration; the most severe complication, if disease is untreated, is the development of amyloidosis. 21 adult patients were enrolled and compared with 27 sex/age matched healthy controls, 6 patients with Familial Mediterranean Fever (FMF), and 15 with Unclassified fever syndrome (Uc). Genetic analysis revealed a not correspondence between clinical classification and nonsense or missense mutations in the MEFV gene encoding pyrin, and three patients did not carry mutations. Results highlighted in both FMF and Uc patients significant decreased levels of α-defensins 2, 3, and 4, involved in innate immune-defense, and increased levels of anti-inflammatory proteins, like cystatin C, glutathionylated and cysteinylated cystatin B, antileukoproteinase, and glutathionylated S100A9, with respect to controls. Uc patients showed higher levels of some peptides and proteins involved in the oral cavity protection than both FMF patients and healthy controls. These peptides/proteins were: Histatin 1 (Hst-1), mono- and non-phosphorylated; Hst-3, 5, and 6; di-, mono- and non-phosphorylated proteoforms of statherin and its des1-9 fragment; P-C peptide and the di-, mono- and non-phosphorylated proteoforms of the acidic Proline-Rich Proteins, PRP1 and PRP3. Interestingly, Uc patients exhibited a hypo-phosphorylation of Hst-1, statherin, PRP1 and PRP3 suggesting a lower activity of the Fam20C kinase responsible for their phosphorylation.

In the second study 22 patients submitted to surgical resection of colo-rectal tumors or adenomas have been analysis. Two different regions of the tumor have been explored for each patient: the superficial tumor region (*S*), and the deepest region of the tumor, corresponding to invading tumor cells (*D*). From the same patient, normal mucosa (*H*) was also collected. Structural characterization of peptides/proteins was performed by high-

resolution RP-HPLC-ESI-MS/MS by a top-down and a bottom-up approach. Quantification of peptides and proteins was performed by low-resolution RP-HPLC-ESI-MS with a label-free method based on the area of the extracted ion current (XIC) peaks. Specific multiply-charged ions of each peptide/protein were selected avoiding common m/z values for coeluting species.

The following peptides and proteins belonging to the thymosin family were characterized and quantified: thymosin β4 (Tβ4), thymosin β10 (Tβ10) and derivates. Moreover, we identified two derivatives of Isoform I of pro-thymosin α (proTα1, 111 amino acid residues), corresponding to the N-terminally truncated form at the second residue, and the fragment 2-36, and parathymosin (102 amino acid residues) N-terminally acetylated after removal of the Met residue. The high-resolution MS/MS data allowed us characterizing other components, such as the ubiquitin, the SH3 domain-binding protein 1 (SH3BP-1), the fatty acid-binding protein 1 (FABP1), and its natural variant with the single substitution Thr94>Ala, detected in the form Met1-missing and N-terminally acetylated. Other proteins and peptides detected in the samples are still pending for identification.

Quantitative analysis showed that Tβ4 was more concentrated in the tumor D tissue with respect to the tumor S ($p = 0.0004$), and to the normal tissue H ($p = 0.03$). Tβ4 concentration did not show significant difference in S and H tissues. Also Tβ10 exhibited the same trend: more concentrated in D with respect to S ($p = 0.01$), none difference between S and H. Ubiquitin and proTα1 showed a similar high level in S and D samples, and a low level in H samples (D vs H $p = 0.007$ and D vs H $p = 0.02$ respectively). Tβ4, Tβ10 fragments, and proTα1(2-36) were observed more frequently in the normal mucosa, but not significant differences were obtained for their concentration. FABP1 showed the highest concentration in the H tissues and none changes between S and D tumor tissues (S vs H $p = 0.0002$, D vs H $p = 0.0005$). Similar trends were observed for several peptides and proteins not yet identified. The high concentration of Tβ4 in the invasion front of the tumor is in agreement with an involvement of the peptide in the epithelial-mesenchymal transition of CRC. Moreover, these preliminary results, evidencing a differential expression of peptides/proteins in the deepest region of the tumor, corresponding to invading tumor cells, with respect to the superficial tumor, and the normal intestinal tissue, could suggest different roles for these components in the CRC carcinogenesis.

Corresponding Author:
Dr. Valentina Piras,

Department of Life and Environmental Sciences, University of Cagliari, Monserrato Campus, Monserrato, CA, Italy
dr.ssa.valentina.piras@gmail.com

1.0 Introduction

1.1 *Proteome and proteomics*

The terms “proteome” and “proteomics” were coined in the early 1990s by Marc Wilkins, a student at Australia's Macquarie University, in order to mirror the terms “genomics” and “genome”, which represent the entire collection of genes in an organism.

The proteome is the entire set of proteins expressed by a genome, cell, tissue, or organism at a certain time.

While the genome is considered to be largely static, the proteome exhibits considerable plasticity owing to alternative splicing events, protein modifications, and the amalgamation of proteins into complexes and signaling networks that are regulated both spatially and temporally (Altelaar, Munoz, and Heck 2012).

Proteomics is essential for deciphering how molecules interact as a system and for understanding the functions of cellular systems in healthy and disease states (Patterson and Aebersold 2003); (Yates, Ruse, and Nakorchevsky 2009).

Global protein analysis poses a tough analytical challenge, in part owing to the highly diverse physicochemical properties of amino acids, which are the building blocks of proteins. In fact, the characteristics of the human proteome, which include a high dynamic range of protein expression, degradation, extreme complexity due to a plethora of post-transcriptional modifications (PTMs), and sequence variations, make such analyses challenging.

Proteins manifest physiological as well as pathophysiological processes in a cell or an organism, and proteomics describes the complete protein inventory in dependence on *in vivo* parameters (Kellner 2000). Disease mechanisms or drug effects affect the protein profile of a biological system and reveals information for the understanding of disease and therapy. PTMs modulate protein activity, stability, localization, and function (Mann and Jensen 2003), playing essential roles in many critical cell signaling events in both healthy and disease states (Krueger and Srivastava 2006). Dysregulation of a number of PTMs, such as protein acetylation, glycosylation, hydroxylation, and phosphorylation, have been implicated in a spectrum of human diseases, including, but not limited to, cardiovascular disease, cancer, and neurodegenerative diseases (Krueger and Srivastava 2006); (Karve and Cheema 2011). PTMs are key regulators of protein activity and involve the covalent modification of proteins by chemical groups, lipids or even small proteins. In addition,

proteins can be cleaved by proteases, and the chemical nature of amino acids can be modified.

Taking into account the number of expressed protein-coding human genes (~11,000), the array of PTMs available (more than 200), the number of potentially modified residues, the dynamic nature and the often low stoichiometry of these modifications, one realizes the magnitude of the analytical challenge in order to identify these modifications and to localize their sites. Advancements in MS methodologies, have greatly improved the analysis of PTMs (Altelaar, Munoz, and Heck 2012). Furthermore, proteins often interact with each other in stable or transient multi-protein complexes of distinct composition, with an estimated 130,000 binary interactions in the human interactome, most of which remain to be mapped (Venkatesan et al. 2010). Moreover, proteins can interact with other molecules, such as RNA (Castello et al. 2012) or metabolites (X. Li et al. 2010). Consequently, a comprehensive analysis of all proteoforms is imperative for the understanding, diagnosis, and treatment of human diseases.

At the beginning of the proteomics era, the principal platform consisted of a comparative 2D electrophoresis that allowed the differences in the protein profiles of two (or more) conditions to be detected. Subsequently, the proteomic platforms have been implemented by a variety of instrumental device arrangements, mainly based on the coupling of high-throughput separation methods with different MS equipment and by different biochemical experimental designs conceived to pursue a number of different issues (Messana et al. 2013).

1.2 *Top-down and Bottom-up integrated platforms for proteomic analysis in health and disease*

Several proteomic platforms have been developed to achieve specific goals with the best results.

Proteomic platforms can be classified in quantitative and qualitative (Nikolov, Schmidt, and Urlaub 2012) as well as in top-down and bottom-up on the base of different strategy utilized in the sample treatment (Bogdanov and Smith 2005).

The main point of qualitative platforms is to define the complete set of proteins present in a certain sample, post-translational modifications (PTMs) comprised, the typical set of proteins specifically expressed in cellular sub-compartments, without considering their abundance. However, qualitative proteomics has to face the unequal distribution of the concentration of distinct proteins present in the biological sample, because the highly abundant proteins can prevent the detection of that ones at low concentration (Messana et al. 2013).

Top-down and Bottom-up approaches are two popular approaches that differ in the protocol applied for the sample treatment (Fig. 1).

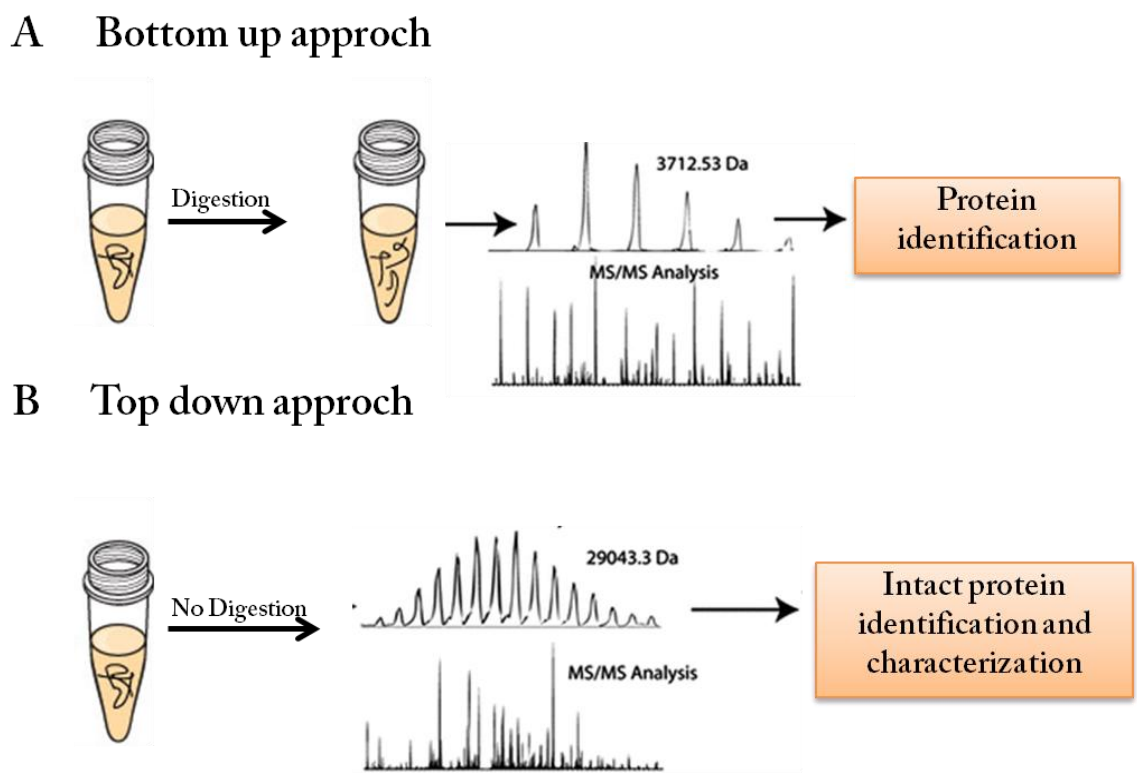
Top-down platforms analyze proteins and peptides in their naturally occurring form, giving particular attention to avoid, as much as possible, any sample alteration (Tipton et al. 2011). Conversely, the bottom-up approach consists in the analysis of the sample digested by specific enzymes, generally trypsin, which cleave proteins in correspondence of defined amino acidic residues. The presence of a protein in the sample is inferred by the detection of one or more of its specific fragments, implying bi-univocal correspondence between the intact protein and the tryptic fragments(Tipton et al. 2011).

Both techniques carried out by tandem mass spectrometry require previous separation steps, in order to reduce the high complexity of the mixture.

The different separation methods can be classified in: gel-based approaches, which can be applied for bottom-up analysis, such as the 2-dimensional gel electrophoresis (2-DE); or gel-free-based approaches, employed for top-down experiments, for example liquid chromatography.

Bottom-up strategies can be further classified in break-then-sort and sort-then break (Han, Aslanian, and Yates 2008).

Fig. 1. Bottom-up and Top-down approaches.



This division takes into account the temporal insertion of the separation step with respect to the fragmentation process. In break-then-sort approaches, the digestion is carried out on the whole set of proteins present in the sample followed by high efficient chromatographic separations (i.e. nano-HPLC, 2D HPLC) coupled to tandem mass spectrometry experiments.

Instead, in sort-then-break strategies, a separation step anticipates the sample digestion such in the case of the determination of the proteome of cellular organelles, the characterization of phospho-proteomes and the analysis of samples submitted to depletion of abundant proteins. In this way, will be digesting only specific proteins of interest further submitted to MS/MS spectrometry analyses.

The majority of proteins are submitted to extensive PTMs, cleavages included, before reaching the mature functional structure and the protein maturation can deeply vary as a function of cellular cycles, tissue, and organ. As a consequence, the minimalistic approach of the bottom-up strategy, when applied to a proteome, can result in the loss of important molecular information. A digestion generating bigger fragments may reduce the problems connected to bottom-up strategies, and this approach is called “middle-down” platform. The major of the fragments can allow the characterization of some PTM codes. Anyway,

middle-down platforms have to be included in the bottom-up platforms because the intact structures of the naturally occurring peptides and proteins are definitely modified by the cleavage. Given the importance of PTMs in the regulation of intracellular signaling and the link between the aberrant or altered PTMs of a number of proteins and human disease, the top down MS approach holds significant promise for the elucidation of proteoform-associated disease mechanisms by providing a powerful method for the identification, characterization, and quantification of proteoforms, which can subsequently be correlated with disease etiology.

On the basis of the above considerations, it could be supposed that top-down platforms should be always preferred to bottom-up. However, top-down platforms have limited performances and they are not prone to extensive and high-throughput management of the datasets.

In fact, top-down tandem mass spectrometry platform technique does not allow characterizing the intact structure of higher molecular weight or glycosylated proteins, due to the complexity of MS/MS spectra that cannot be automatically analyzed by available software (Meyer 2011).

Instead, the bottom-up approach coupled with data banks and bio-informatics tools for automatic analysis of MS/MS data allows to characterize a lot of peptides in a single experiment even if the enzymatic fragmentation preceding the analysis reflects on the inevitable loss of qualitative and quantitative information on the naturally occurring peptidome. In addition, post-translational modification may remain undetected by this approach, and determination of the abundance of different isoforms of the same protein within the sample may be not possible (Tipton et al. 2011), (Massimo Castagnola, Cabras, Iavarone, Vincenzoni, et al. 2012); (Messana et al. 2013); (Cabras et al. 2014).

Therefore, in order to investigate complex proteome of big proteins and glycoproteins, the bottom-up platform is the unique viable strategy.

1.3 *Quantification of proteins and peptides*

Quantitative platforms are employed to determine the amount of each protein component within and among different samples, as levels of proteins and/or their different isoforms can change under different physiological or pathological conditions.

Quantitative approaches can be further divided in relative and absolute.

The relative quantification allows establishing the differences in two (or more) proteomes, (i.e. healthy versus pathological subjects) evidencing statistically significant increases or decreases of protein levels. For large proteomes, the relative quantification is the general approach.

Until now high-throughput analytical platforms for absolute quantization (AQUA) are associated to the development of selected reaction monitoring methods with the application of isotopomers of proteotypic peptides.

Top-down proteomics allows label-free quantification of entire proteins, peptides and their different derivatives and fragments naturally present in the sample by a powerful label-free approach based on the measurement of the eXtracted Ion Current (XIC) peak area.

This approach, avoiding the employment of labeled peptides, consents to perform quantification without any limitation on the number of the species under study (Massimo Castagnola, Cabras, Iavarone, Fanali, et al. 2012) ; (Cabras et al. 2014). The relative percentages of different isoforms of the same protein in a sample can be calculated (Inzitari et al. 2005) (Iavarone et al. 2013) and their diverse abundance, as well as the dissimilar patterns of protein fragmentation, can be compared in different samples and correlated to specific physiological states (Cabras et al. 2009); (Morzel et al. 2012) ; (Hardt et al. 2005) ; (Messana et al. 2015) or pathological conditions (Cabras et al. 2010) ; (Cabras et al. 2013).

1.4 *Proteomics in biomarkers discover.*

One of the most challenging applications of proteomics is the identification of protein biomarkers useful as prognostic or diagnostic clues, as indicator of the disease state or applicable to the monitoring the therapeutic response of the patient.

Biomarkers are defined as "measurable characteristics that reflect physiological, pharmacological, or disease processes" according to the European Medicines Agency (Atkinson A.J. et al. 2001).

Proteomics is a systematic approach to explore the protein compositions of a cell, an organelle or even an entire organism.

Proteomic-based approaches for biomarker investigation can be employed in different aspects of medicine, such as elucidation of pathways affected in disease, identification of individuals who are at a high risk of developing disease for prognosis and prediction of response, identification of individuals who are most likely to respond to specific therapeutic interventions, and prediction of which patients will develop specific side effects.

Recent technological advances have materialized in the design of comprehensive pipelines that integrate discovery and validation phases, enabling plasma biomarkers to be identified for different types of disease (Addona et al. 2011) ; (Whiteaker et al. 2011).

Although several successful biomarkers have been introduced for clinical use, many claimed biomarkers have a limited reliability or remain without proper validation (Poste 2011), leading to skepticism among clinicians. One of the explanations for this is the difficulty in developing effective validation and standardization of procedures to improve reproducibility, sensitivity and specificity (Hanash and Taguchi 2010).

The primary shortcomings of many biomarker studies are a lack of proper controls in the discovery phase, the use of appropriate statistical tools for biomarker definition and the need for independent validation steps in large patient cohorts to certify the legitimacy of the biomarker unambiguously (Liotta and Petricoin 2011); (Mischak et al. 2010).

Such weaknesses lead to claimed biomarkers that are seldom directly related to the disease biology.

There are several different platforms used for this purpose in quantitative clinical proteomics. Identification of a large number of proteins in biological samples has become a routine practice, thanks to advances in instrumentation, computing power and advanced bioinformatics (MP, Wolters, and 3rd 2001).

Technological advances in proteomics enable the analysis of hundreds of proteins at a time discovery efforts have generated hundreds to thousands of candidate protein markers in many disease areas (Anderson 2010).

Proteomics provides an attractive approach to study complex diseases including cancer. It has focused on the discovery of diagnostic, prognostic and predictive disease biomarkers, with a particular focus on biomarkers that can be analyzed in easily accessible biological fluids such as saliva, blood or urine (Celis and Moreira 2008). Currently, clinical proteomics studies are increasingly shifting toward the parallel analysis of tissues and biological fluids to address the question of biomarker specificity.

Mass spectrometry is the pillar of proteomics, in fact using mass spectrometry technologies, proteins can be analyzed rapidly, accurately and with high sensitivity at a relatively low cost with high reproducibility (Nilsson et al. 2010).

2.0 Objectives of the thesis

The main point of this thesis was to apply top-down and bottom-up proteomic platforms to study complex protein mixtures from different biological matrices such as saliva and colon mucosa. To this end I have characterized the proteomic profiles in physiological (healthy subjects) and in pathological conditions: autoinflammatory disease and colorectal cancer.
The aim was to highlight the presence of qualitative and quantitative variations in the proteome profile that could be linked to the disease and thus potential biomarkers.

PART 1

TOP-DOWN PROTEOMICS CHARACTERIZATION OF THE ACIDIC SOLUBLE FRACTION OF HUMAN SALIVA REVEALED SIGNIFICANT CHANGES IN PATIENTS AFFECTED BY PERIODIC FEVER SYNDROMES

3.0 *Human Saliva.*

Saliva is a clear body fluid composed by more than 99% of water, and containing significant amounts of proteinaceous material (including enzymes such as amylase, lysozyme, lipase, acid phosphatase, lactoperoxidase, superoxide dismutase, various peptide hormones, and others), glycoproteins (the main constituents of the mucosal secretions), lipid (hormones such as testosterone and progesterone) and inorganic ions such as sodium, chloride, potassium, calcium, magnesium, bicarbonate, phosphate (Cabras et al. 2014).

It is a unique body fluid continually bathing the mucosa of the oral cavity, oropharynx and larynx, the salivary protein content is a complex mixture deriving from the secretion of salivary glands, major (parotid, submandibular and sublingual) and minor (labial, palatine, buccal and lingual), gingival fold and oral mucosa transudate, mucous of the nasal cavity and pharynx, non-adherent oral bacterial, food remainders, desquamated epithelial and blood cells (Humphrey and Williamson 2001).

This fluid is necessary to lubricate mouth tissues, forming a barrier against irritant elements, e.g. hydrolytic enzymes produced by plaque bacteria, and substances derived from smoking. Mucins, complex glycosylated proteins are the main lubricating components for their high viscosity, great elasticity, strong adherence and they participate to the formation of the acquired enamel pellicle that protect tooth and also support speech, mastication and swallowing (Humphrey and Williamson 2001) ; (Messana, Inzitari, et al. 2008); (Amerongen and Veerman 2002).

According to viscosity of their secretions that is dependent from the content of mucins and lipids, salivary glands can be clustered in serous (parotid), mucous (minor glands), and mixed (sublingual and submandibular). About 65% of un-stimulated (resting) saliva originates from the sub-mandibular gland, 25% from the parotid, 4% from the sublingual and 8% from minor salivary glands (Mamta et al. 2013). These percentages vary under stimulation, principally for an increased contribution of parotid saliva. Saliva is responsible for the initial digestion of starch, mainly by the presence of salivary amylase (or ptyalin). This enzyme is considered to be a good indicator of proper functioning of the salivary glands, particularly of the parotid, contributing up to 20-30% of total protein in saliva. The majority of the enzyme (80%) is synthesized in the parotids, and remainder in the submandibular glands.

More than 2000 proteins and peptides have been detected in human saliva (Bandhakavi et al. 2009). More than 90% of these derive from the secretion of three pairs of ‘major’

glands (parotid, submandibular and sublingual glands) and belong to the classes of proline-rich proteins (PRPs) that are divided into acidic, basic and basic glycosylated proteins, α-amylases, mucins, salivary (S-type) cystatins, histatins, statherin and P-B peptide (Siqueira et al. 2008).

All the other components detected in saliva represent the remaining 10% in weight. Some of them are secreted by salivary glands, but the majority probably derive from exfoliated cells, exudates from epithelial tissue, crevicular fluid and from the contributions of the host oral flora. For example, α-defensins and β-thymosins, derive from gingival crevicular fluid (Ngo et al. 2009), whereas albumin, are probably the products of mucosa exudates (Gorr 2009). During the transit in their secretory way the salivary proteins are subjected to a number of changes including removal of the signal peptide and several PTMs such as proteolytic cleavage, glycosylation, phosphorylation, and sulfation (Massimo Castagnola, Cabras, Iavarone, Vincenzoni, et al. 2012).

Further modifications of the proteins and peptides occur in ducts and in the oral cavity after secretion from the cells as a result of a number of proteolytic enzymes of different origin (Schulz, Cooper-White, and Punyadeera 2013) ; (Messana, Inzitari, et al. 2008).

Removal of C-terminal residues by specific carboxypeptidases has been observed in many salivary peptides and it is considered an event common to all secretory processes. Glycosylation and sulfation of salivary peptides follows pathways common to the secretory processes of other tissues too (Messana, Inzitari, et al. 2008).

Saliva plays an important role in the health maintenance of oral surfaces, by means of antibacterial and antiviral activity, in the lubrication and repair of the oral mucosa and in the taste and digestion. Saliva proteins are related to different functions: 21% are associated with immunity, 1.6% are associated with protein replication and reparation, 4.8% are associated with cell mobility and secretion, 2.3% with transcription and ribosomes, 4.2% with cell multiplication and cell cycle, 9.7% with signal transduction, 5.2% with metabolism and 7.1% with the cytoskeleton and endomembrane. Of the remaining, 28.7% are proteins of uncertain and 15.4% of completely unknown function (Wu et al. 2009).

Some salivary elements, i.e. statherin, histatins, cystatins and proline-rich proteins, regulate calcium homeostasis and mobilization, allowing the equilibrium between demineralization and remineralization necessary for the maintenance of the tooth integrity (Messana, Cabras, et al. 2008) ; (Humphrey & Williamson, 2001). Of great importance is saliva antibacterial activity, in which immunologic and nonimmunologic elements are involved:

IgA, secreted by plasma cells; IgG and IgM, α-defensins deriving from gingival crevicular fluid; glycoproteins, statherins, agglutinins, histatins, proline-rich peptides, and enzymes secreted by salivary glands. In addition, mucins give their contribution in the control of bacterial and fungal colonization promoting benign commensal flora growth. On the other hand, glycoproteins, statherin, agglutinins, histatins and salivary proline-rich proteins are involved in a "clumping" process that reduces bacteria ability to adhere and colonize oral tissues (Humphrey & Williamson, 2001).

The great variety of functions in which this body fluid is involved and its role in the health of the oral cavity highlights the importance of a proper salivation and saliva composition. A great number of studies on the salivary proteome have provided an increasingly comprehension of the composition of this biological fluid and the wide spectrum of functions in which salivary proteins are involved has stimulated research (Cabras et al. 2014) (Massimo Castagnola, Cabras, Iavarone, Vincenzoni, et al. 2012); (M Castagnola et al. 2011) to understand their mechanisms of action, their reciprocal interaction and the relations with other components in the oral cavity. Moreover, the easy, rapid and noninvasive collection of saliva samples pushed many researchers to consider the possibility of using this biofluid for diagnostic and prognostic purposes, not only for oral diseases but also for systemic pathologies. Therefore, several proteomic studies have been performed to evidence potential salivary biomarkers, and this is one of the main goal during my PhD.

In fact, saliva represents an increasingly useful auxiliary means of diagnosis.

In the last 10 years, saliva has become the object of various studies (Massimo Castagnola, Cabras, et al. 2011) ; (Hu et al. 2006) ; (Amado et al. 2010) ; (Zhang et al. 2013).

Approximately 20% of total salivary proteins are also seen in plasma and show comparable functional diversity and disease-linkage (Zhang et al. 2013). This fact abets the potential of salivary proteomics not only in the diagnosis and monitoring of oral diseases but also of systemic conditions. Salivary proteomics aims to discriminate between healthy and pathological states through the identification of proteins that are uniquely correlated to a specific state (Ruhl 2012).

The ability to monitor health status, disease onset and progression and treatment outcome through non-invasive means is a highly desirable goal in healthcare management.

There are three prerequisites to this goal: 1) the existence of specific biomarkers associated with a health or disease state; 2) a non-invasive approach to detect and monitor biomarkers; 3) appropriate technologies to discriminate biomarkers.

Alteration in the salivary profile in different physiological states, such as age, diet or circadian variations or in a disease can be potentially helpful for diagnostic purposes.

In fact, changes of saliva composition related to different pathological conditions have been evidenced by various top-down proteomics approaches as reported in Table n°1.

As shown in table n°1, Hu et colleagues found 46 peptides/proteins, from saliva samples of patients affected by oral squamous cell carcinoma, with significantly different levels when compared to controls (Hu and Wong 2007) by MALDI-MS. Studies about oral cancer evidenced increased salivary levels of transferring (Jou et al. 2010) and a truncated form of cystatin SA-I (Shintani et al. 2010).

The use of ClinProt technique associated with MALDI-TOF-MS technology has permitted the identification of zinc finger protein 510 peptide as a novel salivary biomarker for early detection of oral squamous cell carcinoma (Jou et al. 2011).

Modifications of the salivary proteome have been identified by Streckfus et al. and Wu et al. by the use of a surface-enhanced laser desorption/ionization platform (Streckfus, Bigler, and Zwick 2006) and weak cation exchange magnetic beads, MALDI-TOF-MS respectively.

Xiao et colleagues evidenced by 2-D-DIGE combined with MS technique 16 candidate salivary protein biomarkers in lung cancer (Xiao H, et al, 2011). Studies performed on whole saliva from subjects affected by head and neck squamous cell carcinoma have identified several potential tumor markers using SDS-PAGE-MALDI TOF/TOF-MS (Jarai et al. 2012) and 2-D-DIGE analysis followed by MS identification of candidate proteins (Dowling et al. 2008).

Moreover, several proteins involving in inflammation and bone resorption have been characterized by 2-DE coupled to MALDI-TOF/TOF MS as potential biomarkers for the monitoring of orthodontic tooth movement (Ellias et al. 2012). Studies evidenced caries induced modifications of the salivary proteome (Vitorino et al. 2006) and explicated the role of salivary proteins in denture stomatitis (Bencharit et al. 2012).

Alterations in the salivary proteome of subjects affected by generalized aggressive (Wu et al. 2009) and chronic periodontitis with respect to controls (Gonçalves et al. 2010), as well as of individuals with severe periodontitis before and after periodontal treatment (Haigh et al. 2010), were also demonstrated.

Castagnola and colleagues evidenced the hypo-phosphorylation of His-1, statherin and different isoforms of aPRPs in a subset of approximately 60% of subjects affected by autism spectrum disorders with normal to borderline cognitive development. This data

suggesting that analysis of salivary phospho-peptides might help to discriminate a considerable subgroup of autism spectrum disorder patients (Massimo Castagnola et al. 2008).

Analysis of both whole saliva and parotid saliva by top-down proteomics platform has been applied to study the effects of pilocarpine treatment on salivary proteins and peptides in patients with Sjögren's syndrome (Peluso et al. 2007). Giusti and colleagues (Giusti et al. 2007) demonstrated that sclerosis affect the salivary proteome and showed that the chaperon GRP78/BiP increased in saliva of rheumatoid arthritis patients, suggesting its potential role as rheumatoid arthritis biomarker (Giusti et al. 2010). Several studies evidenced significant modifications of the peptide fraction in patients with Type 1 diabetes, probably due to increased activity of oral proteases (Hirtz et al. 2006) ; (A Caseiro et al. 2013).

Analysis on graft versus host disease (Chiusolo et al. 2013) ; (Imanguli et al. 2007) and Down's syndrome (Cabras et al. 2013) shown change in proteome profile when compared to controls, also that only a few studies are present in the literature.

Regarding the non-invasiveness of collection, saliva represents a suitable medium to be explored for health and disease surveillance.

Table 1. Proteomics changes of saliva composition related to different pathological conditions evidenced by top-down proteomics platforms.

PATHOLOGY	BIOFLUID	BIOMARKERS	PLATFORM
Breast cancer	SWS	Proteins in the range 18, 113, 170, 228 and 287 km/z	SELDI-TOF-MS
Gastric cancer	UWS	4 not-characterized peptides (with m/z values of 1472.78 Da, 2936.49 Da, 6556.81 Da and 7081.17 Da) showed a differential expression	MALDI-TOF-MS
Oral squamous cell carcinoma	UWS	46 biomarkers	MALDI-TOF-MS; LC-MS
Oral squamous cell carcinoma (I stage)	UWS	Truncated cystatin SA-I	SELDI-TOF-ProteinChip-MS
Oral squamous cell carcinoma	WS	zinc finger protein 510 peptide	ClinProt technique MALDI-TOF-MS
Inflammatory lung diseases	Sputum	α-defensins 1, 2, 3, C-terminal amidated peptides	MALDI-TOF-MS
Periodontitis in obese patient	WS	α-defensins 1, 2, 3	SELDI-TOF-MS
Root caries	Parotid Saliva	Changes similar to Sjögren's Syndrome, cystatin S and collagen fragments	HPLC-MS/MS
Denture stomatitis	WS	Cystatin SN, statherin, kininogen-1, desmocollin-2, carbonic anhydrase-6, cystatin C, peptidyl-prolyl cis-trans	SELDI-TOF/MS; LC-MALDI-TOF/TOF-MS

		isomerase and Ig fragments	
Primary Sjögren's Syndrome	Stimulated parotid saliva	Increased inflammation proteins; decreased acinar proteins	2DE/MALDI-TOF-MS
Primary Sjögren's Syndrome	WS	Panel of proteins and peptides	2-DE/LC-MS/MS
Primary and Secondary Sjögren's Syndrome	WS	Higher levels of α-defensin 1 in primary Sjögren's Syndrome; restoration of salivary proteins in primary Sjögren's Syndrome after pilocarpine treatment	HPLC-ESI-MS
Sjögren's Syndrome	UWS	7 not-characterized peptides (with m/z values of 1068.1 Da, 1196.2 Da, 1738.4 Da, 3375.3 Da, 3429.3 Da, 3449.7 Da and 3490.6 Da) showed a differential expression	MALDI-TOF-MS
Type 1 diabetes (controlled)	WS	P-B, P-C, Statherin, fragments of PC, histatins, α-defensins 1, 2 and 4, short S100A9	HPLC-ESI-MS
Type 1 diabetes	UWS	Increased percentage of collagen type I peptide fragments (proteases activity)	HPLC-MALDI-TOF-MS
Down Syndrome	UWS	aPRPs, S-cystatins, α-defensins 1, and 2, histatins 3, and 5, S100A7, S100A8, S100A12	HPLC-ESI-MS
Graft versus host disease	WS	S100A8, S100A9, S100A7	HPLC-ESI-MS
Graft versus host disease	Stimulated Submand/ subling saliva	lactoferrin, cystatin-SN, lbumin, α- amylase	2D-DIGE/MALDI-TOF-MS/MS
Autism spectrum disorders	UWS	Hypo-phosphorylation of statherin, aPRPs, histatin 1	HPLC-ESI-MS

UWS: Unstimulated whole saliva; SWS: Stimulated whole saliva; WS: Whole saliva.

3.1 *Top-down proteomics of human saliva*

Top-down mass spectrometry analyses of the human saliva acidic soluble fraction allow the simultaneous detection of all the soluble proteins and peptides in the sample (Massimo Castagnola, Cabras, Iavarone, Fanali, et al. 2012). To obtain accurate structural information an important instrument that can be used is high-resolution MS. High-resolution MS can be used also to identify and characterize different polymorphisms and several PTMs, i.e. phosphorylation, N-terminal acetylation and oxidation (Messana et al. 2004) ; (Inzitari et al. 2005) ; (Inzitari et al. 2006) ; (Messana, Cabras, et al. 2008) ; (Cabras et al. 2010) ; (Cabras, Manconi, et al. 2012) ; (Massimo Castagnola, Cabras, Iavarone, Fanali, et al. 2012) ; (Iavarone et al. 2013) ; (Cabras et al. 2013).

In addition, it is also possible to characterize the naturally occurring peptides generated in the sample by the action of endo- and exo-proteases (Amado et al. 2010); (Thomadaki et al. 2011). Moreover, the top-down approach allowed to establish the specific origin of the proteins (glandular, ductal or oral) and even to clarify when post-translational modifications and proteolytic cleavages occur along the secretory pathway (Messana, Cabras, et al. 2008).

Top-down proteomics allows label-free quantification of entire proteins, peptides and their different derivatives and fragments naturally present in the sample by a powerful label-free approach based on the measurement of the eXtracted Ion Current (XIC) peak area. This approach consents to perform quantification without any limitation on the number of the species under study (Massimo Castagnola, Cabras, Iavarone, Fanali, et al. 2012); (Cabras et al. 2014). The relative percentages of different isoforms of the same protein in a sample can be calculated (Inzitari et al. 2005) ; (Iavarone et al. 2013) and their diverse abundance, as well as the dissimilar patterns of protein fragmentation, can be compared in different samples and correlated to specific physiological states (Cabras et al. 2009); (Morzel et al. 2012) ; (Hardt et al. 2005) ; (Messana et al. 2015) or pathological conditions (Thomadaki et al. 2013); (Cabras et al. 2010) ; (Cabras et al. 2013).

Therefore, to perform a deep characterization of the human salivary proteome and peptidome is necessary to implement both top-down both bottom-up approaches, in order to take advantage of the two strategies and to minimize their limitations (Cabras et al. 2014).

In the following sections, I describe some structural and genetic features of the protein families characterized in saliva by integrated top-down and bottom-up proteomics platforms.

3.2 *Histatins*

Histatins are low molecular weight peptides, deriving their name from the high number of histidine residues on their structure, secreted both by major and minor salivary glands. The name given by the Oppenheim group derives from the high number of histidine residues in their structure (Oppenheim et al. 1988).

It is widely accepted that all the members of this family arise from two parent peptides, histatin 1 and histatin 3, with a very similar sequence and are encoded by two genes (*HIS1* and *HIS2*) located on chromosome 4q13 (Sabatini and Azen 1989). Despite the very high sequence similarity, these two peptides follow different PTM pathways.

Before secretion, histatin 3 is exposed to an extensive proteolytic cleavage, generating at first histatin 6 (His-3 Fr. 1/25), subsequently histatin 5 (His-3 Fr. 1/24) and then other fragments (Massimo Castagnola et al. 2004). Before the proteomic *era* some of these fragments were named histatin 4-12 (Oppenheim et al. 1988). Recently, many other fragments have been detected, and a new nomenclature has been proposed based on the name of the parent peptide (histatin 1 or histatin 3) and the number of the fragment a.a. residues (Massimo Castagnola et al. 2004). The different susceptibility to cleavage of the two histatins derives from the presence in histatin 3 of the RGYR↓ convertase consensus sequence, absent in histatin 1. Histatin 1 is not cleaved and is mostly found phosphorylated on Ser-2 residue, but the non-phosphorylated derivative is always detectable in whole saliva, although at a low percentage. In spite of the presence of a Ser residue at position 2, histatin 3 is not phosphorylated, probably due to the absence of a +2 flanking glutamic acid residue essential for the kinase recognition. Histatin 1 is partly poly-sulfated in submandibular glands on the 4 tyrosines of the C-terminal domain, differently from histatin 3, which lacks a tyrosine equivalent to Tyr-27 of histatin 1, probably essential for the tyrosylprotein sulfotransferase recognition (Cabras et al. 2007).

It has been shown that histatin 5 is active against various microbes (HelmerhorsHMt et al. 2001) and, in particular, it has a powerful antifungal activity against *Candida albicans* species (Oppenheim et al. 1988) and it is involved in the formation of the enamel pellicle

and in the protection of the tooth structure (Humphrey and Williamson 2001);(J. Li et al. 2004); (Yin et al. 2006); (Vitorino et al. 2007); (Vitorino et al. 2008).
Histatin 1 promote in human saliva wound closure by enhancing cell spreading and cell migration, but do not stimulate cell proliferation (Brand, Ligtenberg, and Veerman 2014).

3.3 *Proline-Rich Proteins*

Proline-rich proteins (PRPs) represent the major fraction of salivary proteins, more than 60% in weight of the total salivary proteome, and they can be classified in acidic (aPRPs), basic (bPRPs) and basic glycosylated (gPRPs) (Anders Bennick 2002).
PRPs represent more than 20–30% (w/w) of total proteins in whole human saliva and more than 50–60% (w/w) of proteins of parotid saliva (Manconi, Castagnola, et al. 2016).
Proline is the predominant amino acid in salivary PRP sequences (25–40% of all amino acids), but Gly and Gln are also highly represented, and globally these three amino acids account from 70 to 88% of all the residues (a Bennick 1987).
aPRPs are secreted both by parotid (about 70%) and submandibular/sublingual glands (about 30%). They are the expression products of two loci, PRH1 and PRH2 located on chromosome 12p13, near to the cluster of bPRPs. *PRH1* codes for the PIF-s, Db-s and Pa isoforms, *PRH2* codes for the PRP-1 and PRP-2 isoforms. The acidic properties are due to several glutamic and aspartic acid residues located in the first 30 amino acids. All the isoforms have a pyroglutamic moiety at the N-terminus and are usually di-phosphorylated on Ser-8 and Ser-22, even though minor quantities of mono-, non-phosphorylated and tri-phosphorylated isoforms (on Ser-17) are also detectable (Inzitari et al. 2005). Four of these isoforms (PRP-1, PRP-2, PIF-s and Db-s) can be partially cleaved near to the C-terminus, eventually releasing a common peptide of 44 a.a. residues (P-C peptide) and 4 truncated isoforms called PRP-3, PRP-4, PIF-f and Db-f. The Pa isoform is not cleaved, and it was usually detected in saliva as S-S dimer due to the specific presence of a cysteine residue (Cys-103) in its structure. Minor quantities of other derivatives missing C-terminal residues from almost all isoforms were also detected (Inzitari et al. 2005). Dimerization of aPRPs (under the action of transglutaminases) occur after secretion (Cabras et al. 2006).
bPRPs and *gPRPs*, secreted only by parotid glands, are the expression product of four loci: (*PRB1-PRB4*) located on chromosome 12p13 near aPRP genes.
At least four alleles *S* (small), *M* (medium), *L* (large), VL (very large) are present at *PRB1* and *PRB3* loci, and three S, M, L at *PRB2* and *PRB4* loci in the western population (Lyons,

Stein, and Smithies 1988) ; (Azen et al. 1990). All the bPRPs deriving from bigger pro-proteins and the connection between the most common haplotypes and salivary phenotypes is still waiting for a complete definition. Proteins and peptides deriving from *PRB1* proproteins are: II-2 peptide (from *S*, *M*, *L* alleles), P-E peptides and IB-6 protein (from *S* allele), Ps-1 protein (from *M* allele) and Ps-2 protein (from *L* allele). From *PRB2* proproteins, IB-1, P-J, P-H, P-F peptides and IB-8a protein (from *L* allele) have been characterized while *PRB3* and *PRB4* proproteins give rise to glycosylated proteins and *PRP4* proproteins also to P-D peptide (from *S*, *M*, *L* alleles).

Given the number of protein sequences obtained from cDNA or genomic DNA large-scale studies, several other potential bPRP species should be detected in human saliva (Manconi, Castagnola, et al. 2016).

It should be outlined that the deep knowledge on the multiple bPRP species detectable in saliva, including their natural variants, has been possible thanks to the application of top-down proteomics and peptidomics platforms, for their ability to investigate complex protein mixtures in their naturally occurring forms (Messana, Cabras, et al. 2008) ; (Cabras et al. 2009); (Massimo Castagnola, Cabras, Iavarone, Vincenzoni, et al. 2012); (Cabras, Boi, et al. 2012); (Messana et al. 2004). Some protein masses pending for a definitive characterization were tentatively attributed to bPRPs family on the basis of their chromatographic properties and the absence of absorption at 270‑280 nm (Massimo Castagnola, Cabras, Iavarone, Vincenzoni, et al. 2012).

Characterization of glycoprotein species is a difficult task, due to their high heterogeneity deriving from the combination of multiple glycosylation sites and different oligosaccharide structures (Manconi, Castagnola, et al. 2016).

HPLC–ESI-MS-based approaches are the most commonly employed techniques for the analysis of protein glycosylation. MS-based glycoproteomic approach, developed by my research group, has permit the identification for the N- and O-linked profiling of glycosylation occupancy at site-specific level of PRP3M glycoproteins (Manconi, Cabras, et al. 2016).

Acidic PRPs play a role in modulating calcium ions homeostasis (A. Bennick et al. 1981), are absorbed in the hydroxyapatite forming the acquired enamel pellicle (Moreno, Kresak, and Hay 1982); (A Bennick, Kells, and Madapallimattam 1983) and could be involved in the bacterial colonization (Gibbons, Hay, and Schlesinger 1991). On the other hand, basic PRPs bind tannins preventing their absorption and toxic effect on the gastro-intestinal tract (Anders Bennick 2002) and are involved in the perception of the bitter taste (Cabras,

Melis, et al. 2012); (Melis et al. 2015). Glycosylated PRPs not only play lubricating actions (Hatton et al. 1985) but it has been also observed *in vitro* that bacteria can use their glycans as a substrate for their own metabolism and growth (Rudney et al. 2010).

3.4 *Statherin and P-B peptide*

Statherin is an unusual tyrosine-rich phospho-peptide (phosphorylated on Ser-2 and Ser-3) involved in oral cavity calcium ion homeostasis and teeth mineralization (Schlesinger and Hay 1977) ; (Schwartz, Hay, and Schluckebier 1992). Its gene (STATH) is localized on chromosome 4q13.3, near to histatin genes (Sabatini et al. 1987). Secreted by parotid and submandibular glands (Schlesinger and Hay 1977), it is di-phosphorylated on serine 2 and serine 3, but also mono- and non-phosphorylated isoforms of this protein and a cycle-statherin can been observed in low quantities (Cabras et al. 2006); (Messana, Inzitari, et al. 2008). The cyclo-structure derives from an intra-molecular bridge between Lys-6 and Gln-37 generated by the action of oral transglutaminase 2 on statherin. In adult human saliva mono- and non-phosphorylated, as well as N- and C-terminal truncated isoforms are always detectable (Inzitari et al. 2006). Statherin has been demonstrated to play a key role in the oral calcium homeostasis, having high affinity for the hydroxyapatite, in the teeth mineralization and in the formation of the enamel pellicle, especially the cyclized form (Cabras et al. 2006); (Schlesinger and Hay 1977).

P-B peptide is the product of *PROL3* gene, localized on chromosome 4q13.3, very close to the statherin gene. After the determination of its structure, was (erroneously) included in the bPRPs family. Differently from classical bPRPs, P-B peptide is not a fragment of a bigger pro-protein, it is secreted both from parotid and Sm/Sl glands (Messana, Cabras, et al. 2008) and it displays three Tyr residues in the sequence.

Statherin and P-B peptide elute closely in the chromatographic profile suggesting a similar polarity. For these reasons, P-B peptide could be functionally connected to statherin. However, while the statherin role is known, none specific function for P-B peptide has been proposed to date.

3.5 *Cystatins*

"S-type" cystatins comprise cystatin S, SN and SA that belong to family 2 of cystatins, inhibitors of cysteine-proteinases and are mainly secreted by Sm/Sl glands. Recent studies suggested that their secretion is not granule-mediated (Messana, Cabras, et al. 2008).

Cystatin S may be mono-phosphorylated on Ser-3 (cystatin S1; about 65%) or di-phosphorylated on Ser-1 and Ser-3 (cystatin S2; about 25%). Cystatin C was frequently detectable in human saliva, while, until now, no protein mass detected in saliva could be attributed to cystatins D and M. Cystatin A and B, called also stefins, belong to family 1 of cystatins, differing from type 2 cystatins for size and phosphorylation. Cystatin A was detectable in 2 isoforms (acetylated and non-acetylated on its N-terminal) (M Castagnola et al. 2011). Cystatin B was N-terminally acetylated and it was usually not detected as unmodified protein in adult whole saliva, because of the reactivity of Cys-3 residue. Cystatin B, indeed was present in whole saliva as S-glutathionylated (about 55%) S-cysteinylated (about 15%) derivatives or as S-S dimer (about 30%) (Cabras, Manconi, et al. 2012). It is an endogenous cysteine cathepsin inhibitor localized in the cytosol, mitochondria and nucleus where it protects cells from the detrimental release of the lysosomal cysteine cathepsins. Its expression is upregulated upon macrophage activation and cellular stress. A possible role of cystatin B in neuro-inflammation has been proposed, in fact mutations in the gene of this protein are associated with the neurodegenerative disease known as Unverricht-Lundborg disease (EPM1)(Kopitar-Jerala 2015). All of them have been found in other body fluids like urine, tears and seminal plasma while cystatin C has a wider extracellular distribution (Abrahamson et al. 1986) (L. A. Bobek and Levine 1992); (Dickinson 2002). Cystatin C consists of 120 amino acids forming a single polypeptide chainand contains four conserved cysteine residues that can form two disulfide bonds but, unlike other family members, was neither shown to be glycosylated nor phosphorylated (Turk and Bode 1991).

It is a target of proteolysis, and primarily functions as a protease inhibitor. It is degraded by cathepsin D and elastase (Lenarcic et al. 1991). Furthermore, cystatin C plays a very important role in many aspects of human health: in rheumatoid arthritis higher levels of this protein correspond to inflammation and disease (Trabandt et al. 1991). Higher levels are found also in cardiovascular disease (Shlipak et al. 2005), stroke (Zeng et al. 2015), diabetes (Reutens et al. 2013) and in neurodegenerative disease an optimal concentrations of this protein protect neurons against amyloid deposition and degeneration (Nakamura et

al. 1991). Cystatins are inhibitor of the cysteine proteinases and have a stronger inhibitory activity for papain and cathepsin C (Saitoh et al. 1987), thus they protect the oral cavity from the proteolytic action of host, bacterial, viral and parasitic proteinases. Furthermore, they seem to play an antibacterial and antiviral action not related with proteinase inhibitory activity. Salivary cystatin, in particular cystatin C, showed antifungal action and the ability to modulate the immune system (L. a Bobek and Levine 1992); (Gu et al. 1995); (Blankenvoorde et al. 1996); (N. Abe et al. 1998) ; (Dickinson 2002). Salivary cystatins also participate to the mineralization of the tooth and to the formation of the acquired enamel pellicle (L. a Bobek and Levine 1992) ; (Dickinson 2002). Cystatin SN and marginally SA are also able to control lysosomial cathepsins implicated in the destruction of periodontal tissues (L. a Bobek and Levine 1992) ; (Baron, DeCarlo, and Featherstone 1999). Moreover, Cystatin SA has been implicated in the induction of cytokines by human gingival fibroblasts (Kato et al. 2000).

3.6 *Proteins of the S100 family*

The S100 protein family represents the largest subgroup within the Ca^{2+}-binding EF-hand superfamily. Their name has derived from the observation that the first identified S100 proteins were obtained from the soluble bovine brain fraction upon fractionation with saturated (100%) ammonium sulfate (Moore 1965).

Phylogenetically, these proteins appear to be rather young, as they are only present in vertebrates (Shang, Cheng, and Zhou 2008). Most of the S100-coding genes cluster on human chromosome 1q21. This clustered organization gave rise to the systematic nomenclature of S100 proteins: polypeptides encoded by genes located within the cluster on chromosome 1 were assigned as S100A proteins with numbers A1–A16, reflecting the position of the gene in the cluster (Marenholz, Lovering, and Heizmann 2006); (Schäfer et al. 1995).

The remaining S100 genes are located on chromosomes 21q22 (S100B), Xp22 (S100G), 4p16 (S100P), and 5q14 (S100Z). The monomeric forms with molecular weights between 10 and 13 kDa consist of two EF-hand helix–loop–helix structures connected by a flexible linker. The C-terminal EF-hand contains the classical Ca^{2+}-binding motif whereas the N-terminal EF-hand exhibits an extended loop structure which is specific for S100 proteins (“pseudo EF hand”), resulting in reduced Ca^{2+} affinity.

They have no intrinsic catalytic activity but after calcium binding, structural modifications allow them to bind and modulate the action of other proteins. They are constitutively expressed in neutrophils, myeloid cells, platelets, osteoclasts and chondrocytes but can be induced and overexpressed in several cell types (macrophages, monocytes, keratinocytes, fibroblasts) in acute and chronic inflammatory, and oxidative stress conditions (Edgeworth et al. 1991);(Vogl et al. 1999);(Eckert et al. 2004); (Lim et al. 2009);(Goyette and Geczy 2011). It has been demonstrated their involvement in a wide range of intracellular and extracellular functions: regulation of calcium homeostasis, cytoskeletal rearrangement, contraction and motility, cell growth and differentiation, membrane organization, arachidonic acid transport, chemotaxis, apoptosis, promotion of wound repair, protection against microbial proliferation, control of ROS formation, inflammation and protein phosphorylation and secretion (Ravasi et al. 2004); (Santamaria-Kisiel, Rintala-Dempsey, and Shaw 2006); (Lim et al. 2009); (Donato 2003). Their activity can be altered and regulated through formation of homodimers and heterodimers and by numerous PTMs: phosphorylation, methylation, acetylation and oxidation that can change their ability to bind ions or target proteins (Lim et al. 2009); (Andrassy et al. 2006); (Zimmer, Wright Sadosky, and Weber 2003). In particular, S100A8 and S100A9 act as scavengers of ROS, protecting tissues from the excess of oxidant (Lim et al. 2009); (McCormick et al. 2005). Among them, S100A7, S100A8, S100A9, S100A11 and S100A12 were already detected in human saliva (Massimo Castagnola, Inzitari, et al. 2011).

S100A7 (psoriasin) was detected in two isoforms of which the variant E27→D is most abundant. Both S100A7 variants were N-terminal acetylated following the loss of the initial methionine.

Four isoforms of S100A9 (calgranulin B) was detected in in human granulocytes (Strupat et al. 2000), and characterized in human saliva (Massimo Castagnola, Inzitari, et al. 2011). Two isoforms defined as long-types, were found to be acetylated following loss of the N-terminal methionine residue and differed from each other in phosphorylation of the penultimate threonine residue of the sequence (Thr112). The other two isoforms, defined as short-types, were found to be acetylated following the loss of the five N-terminal amino acid residues (MTCKM) and differed in the phosphorylation of the same residue of the long-types (Thr108). S100A11 was found to be acetylated at the N-terminal residue following methionine loss (Massimo Castagnola, Inzitari, et al. 2011).

Furthermore, derivatives of S100A8 and S100A9 with different degree of oxidation are recently characterized by my research group through both a top-down proteomic approach

on the intact proteins and peptides present in the acidic supernatant of whole saliva, and a bottom-up approach on the tryptic digests of salivary enriched fractions. S100A8 oxidation involved methionine 1 and 78 (M1, M78), tryptophan 54 (W54), and cysteine 42 (C42). Three proteoforms of S100A8 showed C42 oxidized to sulfonic acid (S100A8-SO3H). The first showed a further oxidation at W54 (S100A8-SO3H/W54ox), the other two forms were isobaric derivatives of S100A8-SO3H. One form was also oxidized at W54 and M78 (S100A8-SO3H/W54ox/M78ox), the other was dioxidized at W54 (S100A8-SO3H/W54diox). These proteoforms will be named hyper-oxidized S100A8 (Cabras et al. 2015). Was also demonstrated the presence *in vivo* of a glutathionylation of C42 in S100A8 (S100A8-SSG). Cysteine 42 of S100A8 originated also a disulfide bridge with cysteine 3 of S100A9(L) (S100A8/A9-SSdimer) (Cabras et al. 2015).

3.7 *α-defensins*

The α-defensins are 29-35 amino acids long; the three disulfide bridges are between residues 1 and 6, 2 and 4, and 3 and 5, resulting in peptides forming a triple-stranded b-sheet structure with a b-hairpin loop containing cationic charged molecoles. They belong to a family of broad-spectrum antimicrobial peptides, identified originally in human and rabbit leucocytes. In humans, α-defensins are expressed in neutrophils, (1 to 4) whereas α-defensins 5-6 are expressed in epithelial cells of the intestinal and reproductive tracts (Ganz and Lehrer 1994).

All four α-defensins can be found in the azurophilic granules of neutrophil granulocytes. In neutrophils, the α-defensins play a role in the oxygen-independent killing of phagocytosed microorganisms. are involved in the regulation of the cell volume, cytokine production (Chaly et al. 2000); (Lehrer and Lu 2012), chemotaxis and inhibition of natural-killer cells (Goebel et al. 2000).

The α-defensins are the major components detected in the gingival crevicular fluid (GCF): α-defensins 1, 2 and 3 are in major concentration, whereas with minor amounts, α-defensin 4. Therefore, it strongly suggested that GCF is the main source of oral α-defensins (Pisano et al. 2005).

The α-defensin 4, also called corticostatin, exhibits pro-inflammatory effects through its anti-corticotropin property, which inhibits the production of cortisol (Singh et al. 1988).

3.8 *β-thymosins*

β-thymosins are ubiquitous polar peptides, firstly isolated from calf thymus (Klein, Goldstein, and White 1965), which are involved in the prevention of actin filament polymerization, induction of metalloproteinases, chemotaxis, angiogenesis; inhibition of inflammation and bone marrow stem cell proliferation. They have been also associated to cancer and metastasis formation(Huff et al. 2001); (Hannappel 2007); (Hannappel 2010). Thymosin β4, and β4 oxidized (encoded by *TMSB4X* gene clustered on chromosome Xp22.2) and β10 (encoded by *TMSB10* located on chromosome 2p11.2) have been detected in whole saliva; they mainly derive from gingival crevicular fluid (Badamchian et al. 2007); (Inzitari et al. 2009) ; (Massimo Castagnola, Inzitari, et al. 2011).

3.9 *Antileukoproteinase*

Antileukoproteinase, also known as human secretory leukocyte protease inhibitor (SLPI), is an 11.7-kDa cationic protein and a member of the innate immunity-associated proteins. It is a nonglycosylated, highly basic, acid-stable, cysteine-rich, 107-amino acid, single-chain polypeptide (Thompson and Ohlsson 1986). The human SLPI gene is localized on chromosome 20q12- 13.2 (Kikuchi et al. 1998). The SLPI gene consists of four exons and three introns and spans approximately 2.6 kb (Kikuchi et al. 1998). To date, no polymorphism of the SLPI gene and no state of SLPI deficiency have been found (Vogelmeier, Gillissen, and Buhl 1996). SLPI was first isolated from secretions of patients with chronic obstructive pulmonary disease and cystic fibrosis and was thereby considered a major anti-elastase inhibitor (Hochstrasser et al. 1972) ;(Ohlsson and Tegner 1976); (Tegner 1978). This protein was also identified and sequenced by my research in group in a precedent study (Massimo Castagnola, Inzitari, et al. 2011). SLPI is produced by neutrophils, macrophages, beta-cells of pancreatic islets, epithelial cells investing the renal tubules, acinar cells of parotid and submandibular glands, acinar cells of submucosal glands, and epithelial cells lining mucous membranes of respiratory and alimentary tracts (T. Abe et al. 1991) ; (Fahey and Wira 2002) ; (Farquhar et al. 2002) ; (Jin et al. 1997) ; (Nystrom et al. 1999). SLPI was originally isolated from parotid saliva (Thompson and Ohlsson 1986) and has been detected in a variety secretions such as whole saliva, seminal fluid, cervical mucus, synovial fluid, breast milk, tears, and cerebral spinal fluid, as in secretions from the nose and

bronchi, etc. (Farquhar, C., 2002 ; Franken, C., 1989 ; McNeely, T. B., 1997 ; Pillay, K., 2001 ; Shugars, D. C. 1999). The SLPI gene was found to be expressed in lung, breast, oropharyngeal, bladder, endometrial, ovarian, and colorectal carcinomas, and SLPI detection is correlated with poor prognosis (Garver, R. I., 1994 ; Westin, U., 2002). SLPI is also found in neurons and astrocytes in the ischemic brain tissue (Wang, X., 2003). SLPI was found to play a pivotal role in apoptosis and wound healing (Ashcroft, G. S., 2000; Odaka, C., 2003 ; Sorensen, O. E., 2003). The main function of SLPI is to protect local tissue against the detrimental consequences of inflammation. It protects the tissues by inhibiting the proteases, such as cathepsin G, elastase, and trypsin from neutrophils; chymotrypsin and trypsin from pancreatic acinar cells; and chymase and tryptase from mast cells (Gipson, T. S., 1999 ; He, S. H., 2003 ; Jin, F., 1997). It also have a bactericidal and antifungal properties.

4.0 *Periodic fever syndrome*

Autoinflammatory periodic fever syndrome refers to a group of rare hereditary recurrent unprovoked inflammation without high titres of autoantibodies or antigen-specific T lymphocyte in the absence of infection (Samuels and Ozen 2006). It has become clear that autoinflammation is caused by dysregulation of innate immunity, as described by Kastner et al, who proposed that autoinflammatory diseases are "*clinical disorders marked by abnormally increased inflammation, mediated predominantly by cells and molecules of the innate immune system, with a significant host predisposition*" (Kastner, Aksentijevich, and Goldbach-Mansky 2010).

During the last few years, alerts on autoinflammatory diseases have increased, leading the scientific community to become aware of this problem and to focus its efforts on the characterization of these diseases.

These diseases primarily include FMF, TNF receptor-associated periodic fever syndrome (TRAPS), hyperimmunoglobulinaemia D and periodic fever syndrome (HIDS), and the cryopyrin-associated periodic syndrome (CAPS) including familial cold autoinflammatory syndrome (FCAS), Muckle–Wells syndrome (MWS) and neonatal onset multi-system inflammatory disease (NOMID)/chronic infantile neurological cutaneous and articular syndrome (CINCA).

In most cases, diagnosis of these pathologies is difficult because the symptoms that occur are similar to viral or bacterial diseases such as influenza, pharyngitis or intestinal problems. Furthermore, the discrimination between the different pathologies is not easy because the

genetic analysis are not always associated with the characteristic symptoms of the disease and, in the same way, the clinical symptoms are not related to specific mutations. In these pathologies is evident a multifactoriality, which complicates the diagnosis, classification, and, consequently, the therapeutic treatment. When a patient has negative genetics for known mutations involved in autoinflammatory periodic fever syndromes, and the clinical symptoms and other test results do not fit a known condition, researchers will often give a diagnosis of "unclassified autoinflammatory disease."

4.1 *Familial Mediterranean Fever*

Familial Mediterranean fever (FMF) is the most frequent hereditary inflammatory disease characterized by self-limited recurrent attacks of fever, serositis, sterile peritonitis, pleuritis, and arthritis. Other areas less frequently affected are the skin, the pericardium, and the tunica vaginalis. It is transmitted in an autosomal recessive pattern and is most frequently seen in patients from around the Mediterranean, including Turkish, Armenian, Sephardic, and Arabic communities (HELLER, SOHAR, and PRAS 1961); (Pras, Pras, and Kastner 1995). The disease is unusual in other populations, but it has been described in Greeks, Italians, Cubans, and Belgians. FMF affects both sexes in a similar ratio (Tunca et al. 2005), although some studies have reported a male predominance (E Sohar et al. 1967).

Most patients (90%) experience their first attack before 20 years of age (E Sohar et al. 1967). FMF attacks unfold suddenly, persist for only a short time (6–96 h) and subside spontaneously. The high-grade fever, and unendurable and disabling pain leave the patient bedridden during attacks. In between the acute episodes the patients are usually asymptomatic. Emotional stress, fatigue, surgery, menstruation, vigorous exercise and cold exposure may trigger an attack, but no clear precipitant is known (E Sohar et al. 1967). The most frequently used diagnostic criteria are those proposed by Tel-Hashomer and Livneh (Koné-Paut, Hentgen, and Touitou 2011) for adults (Table n° 2) and Yalcinkeya et al. for children (Yalçinkaya et al. 2009).

Table 2. The principal diagnostic criteria of periodic fever syndromes.

Livneh FMF diagnostic criteria	
Major criteria	**Minor criteria (Incomplete attacks affecting one or more sites)**
Generalized peritonitis	Abdomen
Unilateral pleuritis or pericarditis	Chest
Monoarthritis (hip, knee, ankle)	Joints
Isolated fever	Exertion-related leg pain
Favourable response to colchicine	
Supportive criteria	
Family history of FMF	Symptom-free interval
Appropriate ethnic origin	Transient inflammatory response with raised inflammatory markers
Age <20 years at diagnosis	Episodic proteinuria/hematuria
Severe attacks requiring bed rest	Unproductive laparotomy or removal of normal appendix
Spontaneous resolution of attacks	Parental consenguity

4.2 *Genetics*

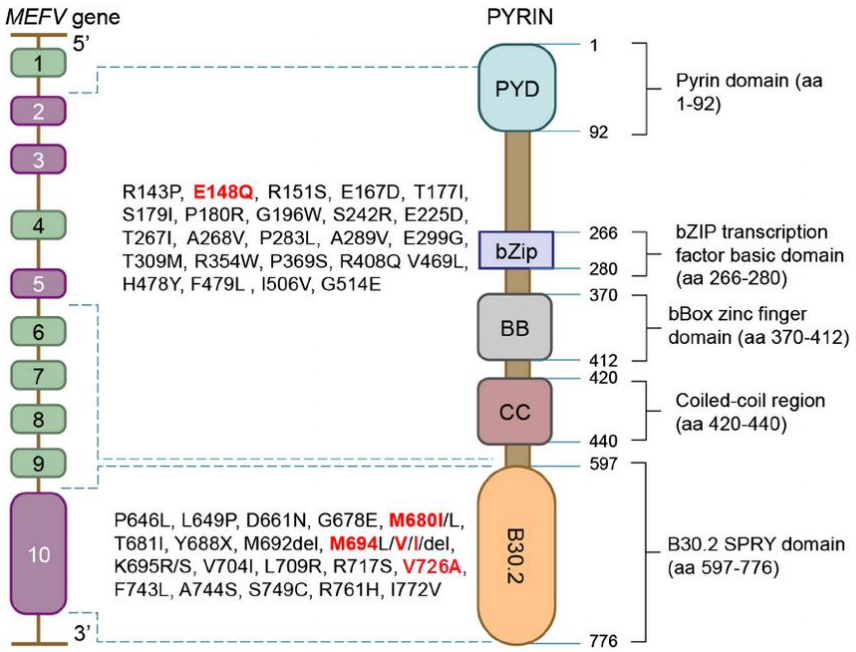
FMF is caused by mutations in MEFV (MEditerranean FeVer) gene, which encodes for a protein called pyrin, also known as marenostin after the Roman name for the Mediterranean Sea—Mare Nostrum.

It is an immunoregulatory molecule made up of 781 amino acids and, while its exact physiological role is unclear, it is thought to have a role in regulating apoptosis, inflammation, cytokine production and modulates IL-1β processing and NF-κB activation. Pyrin is mainly expressed in neutrophils, eosinophils, monocytes, dendritic cells, and fibroblasts (Mankan, Kubarenko, and Hornung 2012); (Seshadri et al. 2007).

A mutated pyrin probably results in uncontrolled inflammation (Onen 2006).

MEFV gene was isolated to chromosome 16 in 1993 (Aksentijevich et al. 1993). Further mapping in the late 1990s refined the 250-kb region associated with FMF to chromosome 16p13.3 (French FMF Consortium 1997). The gene is made up of 10 exons, and 298 variants have been described to date, the majority occurring in exon 10 (Fig. 2) (Lane and Lachmann 2011) ; (Touitou et al. 2004).

Fig. 2. MEFV gene variants.



About 30 mutations associated with FMF were defined. The most prevalent polymorphisms are M680I, M694V, M694I, and V726A on exon 10 (French FMF Consortium 1997); and E148Q on exon 2 (Bernot et al. 1998). The E148Q variant is found in high rates in populations with low disease prevalence, and homozygous E148Q mutations are rarely found in patients with clinical FMF disease (Marek-Yagel et al. 2009); (Naimushin et al. 2011). In a large cohort of Israeli children with FMF, molecular testing for common MEFV mutations identified two mutations in 60% of patients, and no mutation was found in 10% (Padeh et al. 2010). A significant proportion (30%) of patients who had a typical clinical presentation of FMF and a favorable response to colchicine had only one mutation, even after genomic sequencing (Marek-Yagel et al. 2009). Most individuals with only one mutation in MEFV are asymptomatic carriers of the disease; however, classical FMF may clearly occur in carriers of only a single MEFV mutation. Furthermore, carriers of two mutations may exhibit no overt signs of disease. These data, as well as the phenotypic variability of FMF disease, suggest an important role of additional environmental factors, or modifying genes, on the clinical expression of FMF (I Ben-Zvi et al. 2012); (Marek-Yagel et al. 2010). In another study (Ilan Ben-Zvi et al. 2015), patients with mutation negative and genetically heterogeneous MEFV variants present a milder severity phenotype than p.M694 V homozygous patients, leading to the conclusion that the disease in FMF patients without

mutations may be caused by genetic defects upstream or downstream to the MEFV-related metabolic pathway.

FMF is considered to be an autosomal recessive disease. However, an autosomal dominant AID associated with MEFV mutations affecting amino acid 577 in three families has been described. FMF is a clinical diagnosis that can be supported, but not necessarily excluded, by genetic testing. FMF patients with two common mutated alleles, in particular, the M694 V, M680I and M694I on exon 10, are considered at risk of having a more severe disease. Patients homozygous for M694 V mutation are at risk for early onset disease, as was observed in our studies (Padeh et al. 2010). Asymptomatic individuals with risk factors for AA amyloidosis and two common mutated alleles in MEFV gene, especially the M694 V, should be followed closely in order to consider therapy. The E148Q variant is common, of unknown pathogenic significance and when it is the only MEFV variant, does not support the diagnosis of FMF (Giancane et al. 2015).

4.3 *Pathogenesis*

Much work in the field of autoinflammation has focused on the role of the NOD-like receptor family, which is a key player in activation of the innate immune system. One member of this family is the NLRP3 (nucleotide-binding domain and leucine-rich repeat containing family, pyrin domain containing 3) gene, which codes for the cryopyrin protein and is a building block of the inflammasome complex. These inflammasome complexes contain pyrin domains (PYD), and when combined with pyrin they are able to perpetuate inflammation by cleaving the precursor of interleukin-1β (IL-1β) into its mature peptide via a caspase-1-dependent pathway. Activation of the NLRP3 inflammasome is thought to be associated with FMF (Papin et al. 2007) as well as other more common diseases, such as diabetes (Legrand-Poels et al. 2014) and Alzheimer (Tan et al. 2013) disease.

Within this framework, pyrin is thought to have multiple roles:

1) Pyrin is involved in maintenance of the cytoskeleton possibly via its interaction with microtubules (Taskiran et al. 2012) ; (Mansfield et al. 2001).

2) Pyrin increases activation of proinflammatory cytokines via activation of NF-kB via the following pathway (Fig. 3):

(a) Pyrin is cleaved by capsase-1 and its N-terminal fragment interacts with the p65 subunit of NF-kB;

(b) The N-terminal fragment and p65 subunit has increased nuclear permeability when compared to p65 alone;

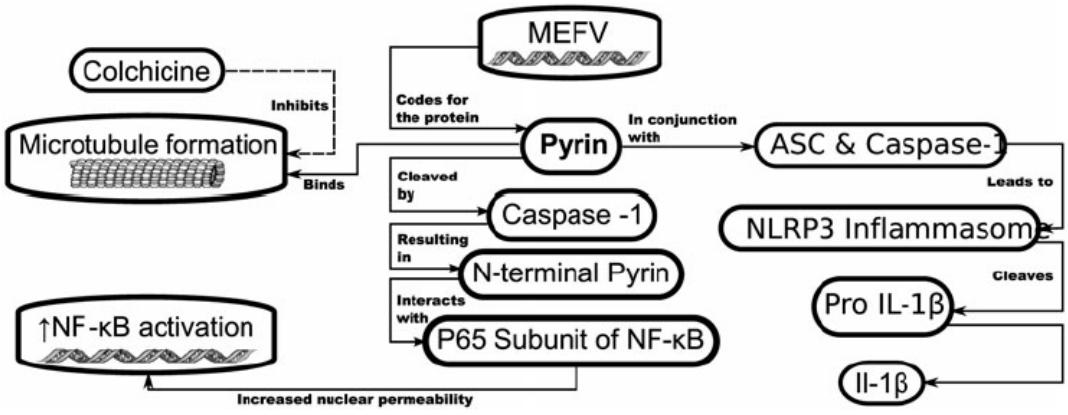
(c) N-terminal pyrin interacts with IkB-α inducing a calpain-mediated degradation of IkB-α, thus potentiating NF-kB activation (Chae et al. 2006) ; (Chae et al. 2008).

3) Pyrin interacts with the inflammasome adaptor protein—ASC (apoptosis-associated speck-like protein with a caspase-recruitment domain (CARD)). Amongst other functions, the inflammasome cleaves (IL-1β) into its mature proinflammatory peptide.

A mutated pyrin probably leads to uncontrolled inflammation by production of IL-1 and inhibition of apoptosis of leukocytes. However, there are some reports that do not support this hypothesis: Gumucio et al.(Gumucio et al. 2002) found no difference between the effect of wildtype pyrin and mutated pyrin with regard to their interaction with the ASC protein and their effect on cell death. Furthermore, Ozen et al. (Ozen et al. 2001) demonstrated that apoptosis was increased in the neutrophils during FMF attacks. Actually, they suggested that the increased neutrophil apoptosis might occur as a response to arrest the inflammation and might explain the spontaneous resolution of the attacks.

The clinical management of FMF is aimed at preventing relapses and prolonging disease-free periods, reducing the severity of attacks, and managing the long-term sequelae of the disease, such as amyloidosis.

Fig. 3. Pro-inflammatory cytokines via activation of NF-kB.



4.4 *Clinical Manifestations*

The hallmarks of FMF are recurrent febrile attacks, accompanied by signs of peritonitis, pleuritis or acute synovitis, lasting 1–3 days, and resolving spontaneously. Attacks occur randomly, from once per week to once in several months, and patients are perfectly well between the attacks.

4.5 *Abdominal attacks*

Abdominal attacks are the most frequent manifestations (90–93% in patients) (Federici et al. 2015); (Demirkaya et al. 2016). Signs of peritonitis, with guarding, rebound tenderness, rigidity and paralytic ileus are often present, mimicking acute appendicitis, and lasting from 24 to 48 h (Ezra Sohar et al. 2016).

4.6 *Arthritis*

Monoarthritis, mostly of the large joints (i.e. the ankle, knee or hip – in that order), though rarely joints in the upper body, is the second most common form of attack occurring in 25–30% of the cases (Federici et al. 2015); (Demirkaya et al. 2016). The joint is hot, tender and often red, resembling septic arthritis precipitated by minor trauma or effort. The synovial fluid is cloudy to purulent and contains large numbers of neutrophils (Garcia-Gonzalez and Weisman 1992).

4.7 *Myalgia*

Muscle pain, mostly of the lower extremities, can develop after physical exertion or prolonged standing. A syndrome of protracted febrile myalgia is characterized by severe debilitating myalgia, prolonged fever and abdominal pain with no peritoneal irritation (Kaplan et al. 2007).

4.8 *Skin Manifestations*

Erysipelas-like erythema are tender, hot, swollen, sharply bordered red lesions usually between the knee and ankle, on the dorsum of the foot or in the ankle region, and are sometimes combined with arthritis (Kolivras, Provost, and Thompson 2013).

4.9 *Isolated Febrile Attacks*

The fever experienced by patients suffering from FMF is typically high over 38 °C and usually resolves after 2–3 days. On occasion it can be the sole presenting sign.

4.10 *Amyloidosis*

Amyloidosis is the most severe complication of FMF and before the introduction of colchicine was the leading cause of mortality in these patients (Herlin, Storm, and Hamborg-Petersen 1985). Type AA amyloid builds up in the kidneys and leads to proteinuria, hematuria, nephrotic syndrome, and a progressively worsening nephropathy, which ultimately leads to renal failure. Patients with certain polymorphisms (M694V) or those with a Sephardic or Turkish background are more likely to develop amyloidosis-related complications (Medlej-Hashim et al. 2004); (Mimouni et al. 2000).

5.0 Objectives of the study

The main goal of this study was to assess whether the auto-inflammation response observed in auto-inflammatory periodic fever syndromes investigated could be associated to qualitative and quantitative variations of the salivary proteins and peptides in the patients with respect to control subjects to have suggestions on potential biomarkers selective for this condition.

6.0 Experimental

6.1 *Materials*

All chemicals and reagents were of analytical grade and were purchased from Sigma Aldrich (St. Louis, MO), and Bio-Rad (Hercules, CA, USA**)**

6.2 *Subjects*

The informed consent process was in agreement with the latest stipulations established by the Declaration of Helsinki. Ethics Committee approval was obtained for the study.
21 adult patients (mean age ± SD: 34.4 ± 10.1; 15 F, 6M) were enrolled by the Unit of Internal Medicine, Allergy and Clinical Immunology of Cagliari University and compared with 27 sex/age matched healthy controls (mean age ± SD: 33.4 ± 9.6; 18 F, 9M). Patients are classified on the base of clinical manifestations as follows: 6 patients with FMF (mean age ± SD: 33 ± 7.9; 5 F, 1M), and 15 with Unclassified periodic fever syndromes (Uc) (mean age ± SD: 35 ± 11.1; 10 F, 5M). Genetic analysis for *MEFV* mutation search were also performed and Table 3 reports the results of genetic analysis on the *MEFV* gene for both the two patient groups.

Table 3. MEFV gene mutations in the Uc and FMF groups.

UNCLASSIFIED PATIENTS	***FMF PATIENTS***
2 WT[a]	1 WT
3 R202 (HO[b])	1 P369S/R408Q (HE)
7 R202Q/WT (HE[c])	1 K695R/WD (HE)
1 E148Q/WT (HE)	1 R202Q/WT (HE)
1 E148Q/R202Q (HE)	1 E148Q/R202Q (HE)
1 I591T/R202Q (HE)	1 A744S (HE)

[a]WT, wild type; [b]HE, heterozygous; [c]HO, homozygous.

6.3 *Clinical data*

The clinical classification of the patients was based on the presence of certain clinical parameters. The main parameter was the presence of febrile episodes and their duration.
In fact, characteristic of these syndromes is a recurrent fever. The age of onset is usually in adulthood except for a few cases where the disease manifests in the first few months after birth. Patients show typical disease's manifestations like oral aphthous ulceration, rash, lymphadenopathies with different localization, arthritis, arthralgia and myalgia as reported in Table n°4.

Table 4. Classification type, onset disease, frequency and duration of febrile episodes and current disease status of each patient at the time of the study.

Pt #	Classification	Onset disease	Febrile episodes frequency	Febrile episodes duration	Disease status
3	Uc	20 years	unknown	1 month	Arthralgia - myalgia
4	Uc	14 years	Every 2-3 months	1 month	Myalgia - oral aphthous ulceration - cervical lymphadenopathy
5	FMF	15 years	Every 1-2 weeks	1 week	Arthritis – arthralgia -myalgia
6	FMF	29 years	Every 1-2 weeks	10 days	Arthralgia - myalgia
7	FMF	3 months	Every 1-2 weeks	1-3 days	Arthritis – arthralgia - erythema nodosum (panniculitis) rash
8	FMF	13 years	Every 2 weeks	2 days	Arthritis – arthralgia – myalgia - papular erythematous rash - cervical, axillary, inguinal lymphadenopathy
9	Uc	16 years	Unknown	Various months	Cervical lymphadenopathy
10	Uc	10 years	Once a month	1 week	Arthritis – arthralgia – papular erythematous, erythema nodosum rash - axillary, retroauricular lymphadenopathy- - oral aphthous ulceration

12	Uc	31 years	Once a year	6 months	Myalgia - lateral cervical lymphadenopathy
13	FMF	28 years	Chronic	Chronic	Arthritis – arthralgia – myalgia - erythematous rash – cervical, mandibular lymphadenopathy
15	Uc	19 years	Unknown	2-4 weeks	erythematous rash (neck and face) - lateral cervical lymphadenopathy
16	Uc	23 years	Unknown	3-4 days	Myalgia – arthralgia - lateral cervical lymphadenopathy
18	Uc	17 years	Every 2-3 months	2-3 days	Myalgia
19	Uc	18 years	Every 2-4 weeks	3 days	Arthritis – arthralgia – myalgia – rash (in the form of blisters) - oral aphthous ulceration - lateral cervical lymphadenopathy
20	Uc	2 years	Every 1-8 weeks	7-10 days	Arthritis – arthralgia – myalgia – rash - lateral cervical lymphadenopathy
21	Uc	22 years	Every 2-8 weeks	3 days	Myalgia
22	Uc	27 years	Every 8-12 weeks	15 days	oral aphthous ulceration - lateral cervical lymphadenopathy
23	Uc	25 years	Every 4-8 weeks	1-2 days	Arthralgia – myalgia - erythema nodosum rash – oral and genital aphthous ulceration
24	Uc	10 years	Unknown	Quotidian fever	Oral aphthous ulceration – cervical, , retroauricular lymphadenopathy
25	Uc	38 years	Unknown	3-7 days	Arthralgia – myalgia – purple rash – cervical lymphadenopathy
26	FMF	23 years	Unknown	Unknown	Unknown

6.4 *Salivary sample collection*

Unstimulated whole saliva was collected, accordingly to a standardized protocol, from patients and healthy controls using a soft plastic aspirator and transferred to a plastic tube in an ice bath. Donors did not eat or drink at least 2 h before the collection, which was established between 10.00 a.m. and 12.00 p.m. Immediately after collection, each salivary sample was diluted in 1:1 v/v ratio with a 0.2% solution of 2,2,2-trifluoroacetic acid (TFA). Then the solution was centrifuged at 20000*g* for 15 min at 4°. The acidic supernatant was separated from the precipitate and either immediately analyzed by HPLC-ESI-MS apparatus (100 µl, corresponding to 50 µl of saliva) or stored at -80°C until low-resolution HPLC-ESI-IT-MS analysis.

6.5 *Top-down proteomics analysis by low resolution HPLC-ESI-MS*

Low-resolution reversed phase (RP)-HPLC-ESI-MS analysis of the acidic soluble fraction of whole saliva was carried out by a Surveyor HPLC system connected to a LCQ Advantage mass spectrometer (ThermoFisher Scientific San Jose, CA). The mass spectrometer was equipped with an electrospray ionization source (ESI) and an ion trap (IT). The chromatographic column was a reversed phase Vydac (Hesperia, CA, USA) C8 column with 5 µm particle diameter (150 x 2.1 mm). The chromatographic separation was carried out using eluent A (0.056% TFA acidic solution) and eluent B (acetonitrile/water 80:20 with 0.05% TFA). The gradient applied for the analysis of saliva was linear from 0 to 55% of B in 40 min, and from 55% to 100% of B in 10 min, at a flow rate of 0.1 ml min^{-1}, entirely addressed to the ESI source. During the first 5 minutes of the analysis the eluate was not directed to the mass spectrometer to avoid that the high salt concentration could damage the instrument.

Mass spectra were collected every 3 ms in the *m/z* range 300-2000 in positive ion mode. The MS spray voltage was 5.0 kV, and the capillary temperature was 260 °C. MS resolution was 6000.

The proteins and peptides analysed in the present study have been already identified in studies previously published by mi research group (Massimo Castagnola, Inzitari, et al. 2011) ; (Cabras et al. 2013) ; (Cabras et al. 2015) thus, their characterization was not here reported. Experimental mass values were obtained by deconvolution of averaged ESI-MS

spectra automatically performed by using MagTran 1.0 software (Zhang, 1998). Experimental masses were compared with theoretical average mass values available at the Swiss-Prot Data Bank (http://us.expasy.org/tools). Moreover, the observed mass values and time elution were compared with those ones determined in previous studies performed with the same experimental conditions. Table 5 reports the Swiss-prot codes, the elution times, the experimental and theoretical average (low-resolution) mass values (Mav), and the multiply-charged ions utilized to selectively extract the ion current peaks used to quantify proteins/peptides and their derivatives.

6.6 *Quantification*

Proteins and peptides reported in Table 5, were searched along the Total Ion Current (TIC) chromatographic profile and quantified by the eXtracted Ion Current (XIC) procedure.

To extract the XIC peaks, selected multiply-charged ions generated by the proteins/peptides at the electrospray ionization source were chosen by excluding values in common with other closely eluting proteins (±0.5 *m/z*).

The area of the XIC peaks is proportional to the protein concentration under constant analytical conditions, allowing to perform relative quantification of the same protein in different samples (Messana, Cabras, et al. 2008).

The estimated percentage error of the XIC procedure was <8%. XIC peaks were considered when the signal to noise ratio was at least 5.

6.7 *Statistical analysis*

GraphPad Prism (version 5.0) was used for statistical analysis. Ranges, medians, means, and standard deviations of the XIC peak areas of the peptides and proteins were calculated in the three groups. A *t*-test was used to compare proteins/peptides XIC peak areas between two different groups. Specifically, a non-parametric *t*-test was chosen if the distribution of the XIC peak areas was not Gaussian for at least one of the two groups (Mann-Whitney *t*-test), while a parametric test was employed if the distribution was Gaussian for both groups (Unpaired *t*-test). Welch's correction was applied if the variance resulted significantly different between the groups. Statistical analysis was considered to be significant when the *p* value was <0.05 (two tailed).

The correlation analysis of clinical parameters was performed comparing the main parameters between FMF and Uc patients. Correlation analysis was considered to be significant when the *p* value was <0.05.

7.0 Results

7.1 *Correlation analysis*

The correlation analysis was performed considering the clinical manifestation indicated in Table 4. No significant difference was found except for the oral aphthous ulceration that seems to be present only in Uc patients with respect to FMF patients. However, a very small number of patients presented this clinical manifestation so it was impossible to obtain a statistically significant data.

7.2 *Quantification of protein and peptides by XIC procedure*

The following proteins and peptides were searched and quantified by XIC procedure, as reported in Table 5: histatins (Hst) 1, 3, 5 (Hst-3 1/24) and 6 (Hst-3 1/25), salivary cystatins S, SN and SA, cystatin A, B, and C, statherin, P-B peptide, acidic proline-rich proteins (aPRPs), α-defensins 1–4, β-thymosins 4 and 10, antileucoproteinase, proteins of S100 family: S100A7 (D27 variant), S100A12, S100A8, S100A9 (short (S) and long (L) species). Moreover, protein species derived from PTMs of these components were searched and quantified (phosphorylated, acetylated, Met-oxidized, truncated forms, disulfide dimers, Cys-gluthathionylated, Cys-cystainylated, Cys-nitrosylated, Cys-sulfonic and sulfinic acid). Basic proline-rich proteins were not evaluated due to their high variability linked to the physiological status (Cabras, Melis, et al. 2012).

In Fig. 4 are shown typical total ion current chromatographic profiles (TIC) obtained by analysis of the acidic soluble fraction of saliva from control (panel A), FMF patient (panel B) and from Uc patient (panel C). The elution ranges of the various classes of proteins and peptides, which are detected in saliva, are indicated on panel A.

The results of the statistical analysis performed by comparing XIC peak areas of the proteins measured in all the Uc and FMF patients and controls are summarized in Table 6. Significant differences between patients and controls and between the two groups of patients were highlighted and reported for each protein family.

None statistically relevant result was highlighted by the correlation analysis of clinical parameters. Despite this, we noticed the presence of oral aphthous ulceration only in UC patients when compared with FMF patients, even if the number of subjects was low to obtain a statistically significant result.

Table 5. Swiss-Prot code, experimental and theoretical average mass values and elution times of proteins and peptides investigated.

Proteins/Peptides Swiss-Prot Code	Elution Time (Minutes)	Exper. (Theor) Av. Mass	Layout m/z (Charge)
Acid Proline-Rich Proteins			
PRP-1 Di-phos. (P02810)	22.9-23.3	15515 ± 2 (15514-15515)	1293.9 (+12), 1194.4 (+13), 1035.3 (+15), 970.7 (+16), 913.6 (+17).
PRP-1 Mono-phos.	23.9-24.3	15435 ± 2 (15434-15435)	1287.2 (+12), 1188.3 (+13), 1030.0 (+15), 965.7 (+16), 908.9 (+17).
PRP-1 Non- phos.	24.2-24.7	15355 ± 2 (15354-15355)	1280.5 (+12), 1182.1 (+13), 1024.6 (+15), 960.7 (+16), 904.2 (+17).
PRP-1 Tri- phos.	22.6-22.9	15595 ± 2 (15594-15595)	1418.7 (+11), 1300.5 (+12), 1200.6 (+13), 1040.6 (+15), 975.7 (+16).
PRP-3 Di- phos. (P02810)	23.3-23.8	11161 ± 1 (11161-11162)	1595.5 (+7), 1396.2 (+8), 1015.7 (+11), 931.1 (+12), 859.6 (+13).
PRP-3 Mono- phos.	23.8-24.2	11081 ± 1 (11081-11082)	1584.1 (+7), 1386.2 (+8), 1008.4 (+11), 924.5 (+12), 853.4 (+13).
PRP-3 Non- phos.	24.8-25.1	11001 ± 1 (11001-11002)	1376.2 (+8), 1101.2 (+10), 917.8 (+12) 786.8 (+14).
PRP-3 Di- phos. Des-Arg$_{106}$	23.5-23.8	11004 ± 1 (11005-11006)	1573.2 (+7), 1223.8 (+9), 1001.5 (+11), 847.6 (+13).
P-C peptide (P02810)	13.6-14.5	4370.9 ± 0.4 (4370.8)	1457.9 (+3), 1093.7 (+4).
Statherin and P-B peptide			
Statherin Di- phos. (P02808)	28.9-29.5	5380.0 ± 0.5 (5379.7)	1794.2 (+3), 1345.9 (+4), 1076.9 (+5).
Statherin Mono- phos.	28.7-29.1	5299.9 ± 0.5 (5299.7)	1767.6 (+3), 1325.9(+4), 1060.9 (+5).
Statherin Non- phos.	28.4-28.8	5220.5 ± 0.5 (5219.7)	1741.2 (+3), 1306.1 (+4), 1045.1 (+5).
SV1 (des-Phe$_{43}$)	27.6-28.0	5232.4 ± 0.5 (5232.5)	1745.1 (+3), 1309.1(+4), 1047.5 (+5).
Statherin des-Thr$_{42}$ Phe$_{43}$	27.7-28.1	5131.2 ± 0.5 (5131.4)	1711.4 (+3), 1283.8 (+4), 1027.2 (+5).
Statherin des-Asp$_1$	28.5-28.9	5264.7 ± 0.5 (5264.6)	1755.9 (+3), 1317.2 (+4), 1053.9 (+5).
Statherin des-$_{1-9}$	27.5-28.8	4127.9 ± 0.4 (4127.6)	1376.9 (+3), 1032.9 (+4).
Statherin des-$_{1-10}$	27.8-28.2	3971.3 ± 0.4 (3971.4)	1986.7 (+2), 1324.8 (+3).
Statherin des-$_{1-13}$	27.8-28.3	3645.2 ± 0.4 (3645.0)	1823.6 (+2), 1216.1 (+3).
P-B peptide (P02814)	29.4-30.5	5792.9 ± 0.5 (5792.7)	1932.0 (+3), 1449.2 (+4), 1159.6 (+5).

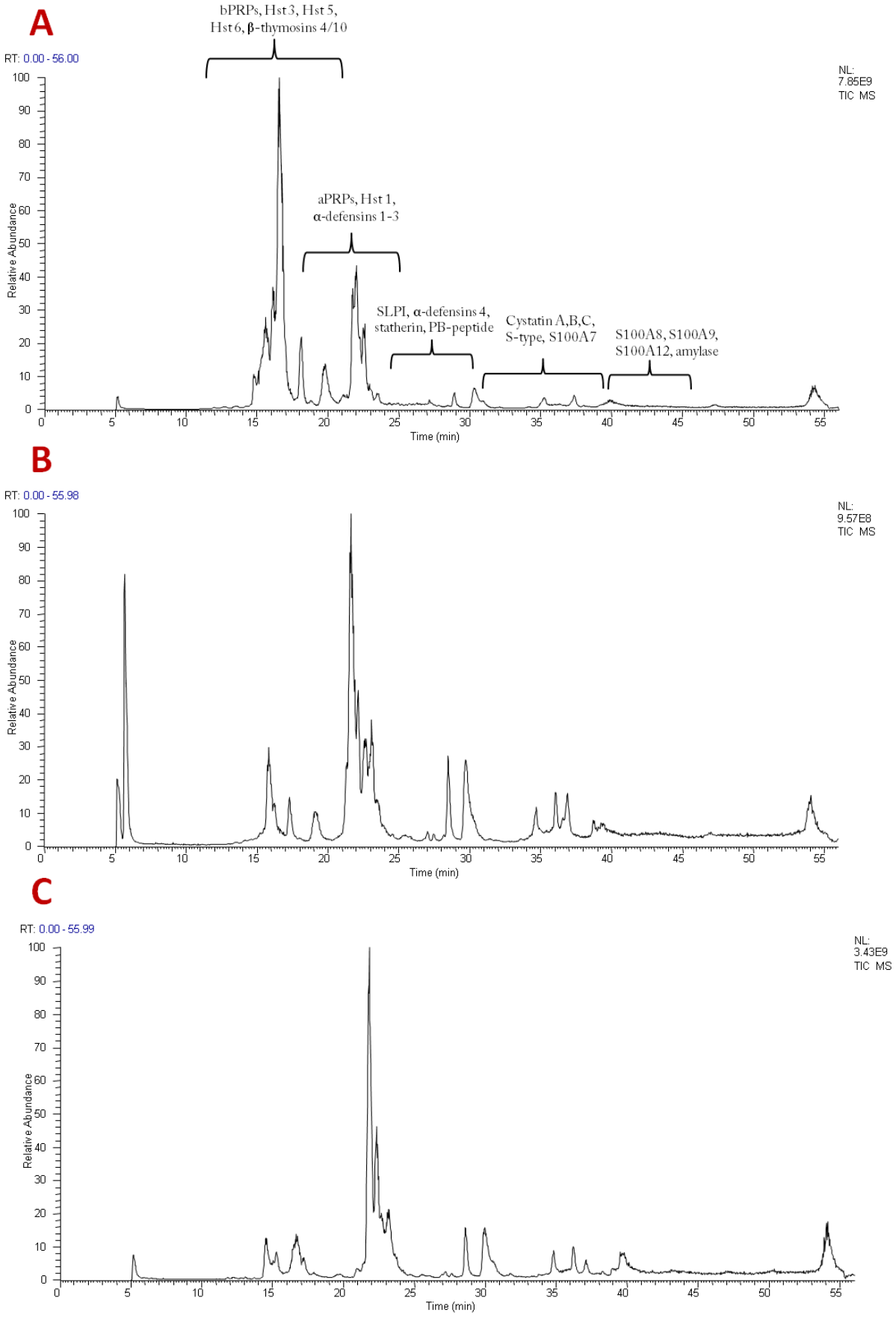
		Histatins	
Hst-1 Di-phos. (P015515)	23.3-23.8	4928.2 ± 0.5 (4928.2)	1644.1 (+3), 1233.5 (+4).
Hst-1 Non- phos.	23.4-23.8	4848.2 ± 0.5 (4848.2)	1617.4 (+3), 1213.5 (+4).
Hst-3 (P15516)	17.6-17.9	4062.2 ± 0.4 (4062.4)	1355.1 (+3), 1016.6 (+4).
Hst-3 1/25 (P15516)	14.0-14.4	3192.4 ± 0.3 (3192.5)	1065.1 (+3), 799.1 (+4).
Hst-3 1/24 (P15516)	14.2-14.7	3036.5 ± 0.3 (3036.3)	1013.2 (+3), 760.1 (+4).
		Cystatins	
S (P01036)	36.5-37.1	14186 ± 2 (14185)	1774.3 (+8), 1577.2 (+9), 1419.6 (+10), 1290.6 (+11), 1183.2 (+12), 1092.2 (+13), 1014.3 (+14).
S1 (S Mono-phos.)	36.6-37.1	14266 ± 2 (14265)	1784.3 (+8), 1586.1 (+9), 1427.6 (+10), 1297.9 (+11), 1189.8 (+12), 1098.4 (+13), 1020.0 (+14).
S2 (S Di- phos.)	36.8-37.2	14346 ± 2 (14345)	1794.3 (+8), 1595.0 (+9), 1435.6 (+10), 1305.2 (+11), 1196.5 (+12), 1104.5 (+13), 1025.7 (+14).
SN (P01037)	34.8-35.2	14312 ± 2 (14313)	1790.0 (+8), 1591.2 (+9), 1432.2 (+10), 1302.1 (+11), 1193.7 (+12), 1101.9 (+13), 1023.3 (+14).
SA (P09228)	38.4-38.9	14347 ± 2 (14346)	1794.4 (+8), 1595.1 (+9), 1435.7 (+10), 1305.3 (+11), 1196.6 (+12), 1104.6 (+13), 1025.8 (+14).
C (P01034)	36.6-37.0	13342 ± 2 (13343.1)	1483.57 (+9) 1335.32 (+10) 1214.02 (+11) 1112.93 (+12) 1027.40 (+13)
C Met-ox	36.9-37.4	13358.5 ± 1 (13358.4)	1485.28 (+9), 1336.85 (+10), 1215.41 (+11), 1114.21 (+12), 1028.58 (+13)
B	31.7-32.2	11179.3 ± 0.5 (11181.6)	1864.6 (+6), 1598.4 (+7), 1398.7 (+8), 1243.4 (9), 1119.2 (+10), 1017.5 (11)
B-SSG (P04080)	33.6-34.4	11485,8 ± 2 (11486.9)	1915.5 (+6), 1642.0 (+7), 1436.9 (+8), 1277.3 (9), 1149.7 (+10), 1045.3 (11)
B-SSC (P04080)	31.4 - 31.8	11299,8 ± 2 (11300.7)	1884.5 (+6), 1615.4 (+7), 1413.6 (+8), 1256.7 (9), 1131.1 (+10), 1028.6 (11)
B- SS dimer (P04080)	32.5-33.1	22358 ± 2 (22361.3)	1862.4 (+12), 1721.1 (+13), 1598.2 (+14), 1491.8 (+15), 1398.6 (+16), 1316.4 (+17), 1243.3 (+18), 1177.9 (+19), 1119.1 (+20), 1065.8 (+21), 1017.4 (+22), 973.2 (+23)
A (P01040)	30.1-31.6	11005.354 ± 2 (11006.5)	1001.59 (+11), 1101.59 (+10), 1223.94 (+9), 1376.81 (+8), 1573.36(+7), 1835.42 (+6)
A des$_{1-2}$ (P01040)	30.2-32	10761.065 ± 2 (10762.1)	979.37(+11), 1077.21(+10), 1196.79(+9), 1346.26(+8), 1538.44(+7), 1794.68(+6)
A N-acet (P01040)	32.4-32.8	11047.43 ± 2 (11048.5)	1005.41(+11), 1105.85(+10), 1228.61(+9), 1382.06(+8), 1579.36(+7), 1842.42(+6)

Defensins			
Def. α 1 (P59665)	22.5 - 23.1	3442.5 ± 2 (3442.1)	1772.03 (+2), 1148.36 (+3), 861.52 (+4)
Def. α 2 (P59665/6)	22.7 – 23.4	3370.4 ± 1 (3370.9)	1686.49 (+2), 1124.66 (+3), 843.75 (+4)
Def. α 3 (P59666)	23.0 – 23.5	3485 ± 2 (3486.1)	1744.03 (+2), 1163.03 (+3), 872.52 (+4)
Def. α 4 (P12838)	27.3 – 27. 7	33708 ± 1 (3709.4)	1855.71 (+2), 1237.48 (+3), 928.36 (+4)
Antileukoproteinase			
Antileukoproteinase (P03973)	25.5 – 26.3	11702.2 ± 1 (11706)	1952.64 (+6) 1673.84 (+7) 1464.73 (+8) 1302.10 (+9)
β-Thymosins			
β-4 (P62328)	19.7 – 20.3	4944.5 ± 1 (4963)	1655.51 (+3) 1241.88 (+4) 993.71 (+5)
β-10 (P62313)	20.4 – 21.0	4934.5 ± 1 (4935)	1646.52 (+3) 1235.14(+4) 988.31 (+5)
β-4 Metox (P62328)	18.0 – 18.5	4977.5 ± 1	1660.84 (+3) 1245.88 (+4) 996.91 (+5)
S100 family			
S100A7 (D27) (P31151[g])	37.4-38.0	11367 ± 2 (11367.8)	1422.0(+8) 1264.1(+9) 1137.8(+10) 1034.4(+11)
S100A12 (P80511)	39.5-40.2	10444 ± 2 (10443.9)	1306.5(+8) 1161.4(+9) 1045.4(+10) 950.4(+11)
S100A8 (P05109)	39.1-39.7	10833 ± 2 (10834.5)	1355.3(+8) 1204.8(+9) 1084.5(+10) 985.9(+11)
S100A8-SO_2H	39.7-40.0	10866 ± 2 (10866.5)	1359.3(+8) 1208.4(+9) 1087.7(+10) 988.9(+11)
S100A8-SO_3H/W54ox	40.2-40.6	10898 ± 2 (10898.6)	1363.3(+8) 1212.0(+9) 1090.9(+10) 991.8(+11)
S100A8 (SO_3H/ Trp_{54}-diox; SO3H/ Trp_{54}-ox - Met-ox)	39.0-39.6	10915 ± 2 (10914.6)	1365.3(+8) 1213.7(+9) 1092.5(+10) 993.2(+11)
S100A8-SSG	38.1-38.4	11140 ± 2 (11139.8)	1393.5(+8) 1238.8(+9) 1115.0(+10) 1013.7(+11)
S100A8-SNO	40.6-40.9	10863 ± 2 (10863.5)	1358.9(+8) 1208.1(+9) 1087.3(+10) 988.6(+11)
S100A8/A9-SS dimer	41.6-41.9	23986 ± 3 (23985)	1600.0(+15) 1500.1(+16) 1411.9(+17) 1333.5(+18) 1263.4(+19) 1200.3(+20) 1143.2(+21) 1091.2(+22) 1043.8(+23) 1000.4(+24) 960.4(+25) 923.5(+26)
S100A9(S) (P06702)	41.3-42.0	12690 ± 2 (12689.2)	1410.9(+9) 1269.9(+10) 1154.6(+11) 1058.4(+12) 977.1(+13)
S100A9(S) Mono- phos.	41.3-42.0	12770 ± 2 (12769.2)	1419.8(+9) 1277.9(+10) 1161.8(+11) 1065.1(+12) 983.3(+13)
S100A9(S)- Met-ox	41.3-42.0	12706 ± 2 (12705.2)	1412.7(+9) 1271.5(+10) 1156.0(+11) 1059.8(+12) 978.3(+13)

S100A9(S)- Met-ox Mono-phos.	41.3-42.0	12786 ± 2 (12785.2)	1421.9(+9) 1279.5(+10) 1163.3(+11) 1066.4(+12) 984.5(+13)
S100A9(L)-SSG	41.1-41.8	13459 ± 2 (13458.1)	1346.8(+10) 1224.5(+11) 1122.5(+12) 1036.3(+13) 962.3(+14)
S100A9(L)-SSG Mono-phos.	41.1-41.8	13539 ± 2 (13538.1)	1354.8(+10) 1231.8(+11) 1129.2(+12) 1042.4(+13) 968.0(+14)
S100A9(L)-SSG/ Met-ox	41.0-41.6	13475 ± 2 (13474.1)	1348.4(+10) 1225.9(+11) 1123.8(+12) 1037.5(+13) 963.4(+14)
S100A9(L)-SSG/ Met-ox, Mono-phos.	41.0-41.6	13555 ± 2 (13554.1)	1356.4(+10) 1233.2(+11) 1130.5(+12) 1043.6(+13) 969.1(+14)
S100A9(L)-SSC	41.1-41.8	13273 ± 2 (13271.9)	1328.2(+10) 1207.6(+11) 1107.0(+12) 1021.9(+13) 949.0(+14)
S100A9(L)- SSC Mono-phos.	41.1-41.8	13353 ± 2 (13351.9)	1336.2(+10) 1214.8(+11) 1113.7(+12) 1028.1(+13) 954.7(+14)
S100A9/A9-SSdimer	41.7-42.5	26306 ± 3 (26304)	1754.6(+15) 1645.0(+16) 1548.3(+17) 1462.3(+18) 1385.4(+19) 1316.2(+20) 1253.5(+21) 1196.6(+22) 1144.6(+23) 1097.0(+24) 1053.1(+25) 1012.7(+26)

SSG= Cys-gluthathionylated, –SSC = Cys-cystainylated, –SNO Cys-nitrosylated

Fig. 4. TIC profiles obtained by analysis of the acidic soluble fraction of saliva from control (panel A), FMF patient (panel B) and from Uc patient (panel C).



7.3 *aPRPs*

As shown in Table 6, Uc patients exhibited a higher level with respect to controls of aPRPs: PRP-1 mono- and di-phosphorylated were very abundant (Fig. 5), and its truncated forms, PRP-3 di-, mono-phosphorylated (Fig. 5) and the non-phosphorylated species, and the P-C peptide.

This difference has not been observed when FMF patients were compared with controls. FMF patients showed basically lower levels of aPRPs with respect to Uc patients, as appreciable form the distribution of XIC peak areas in Fig. 5. PRP-1 di-phosphorylated and PRP-3 mono-phosphorylated were significantly less concentrated in FMF than in Uc patient saliva, even though with a low statistical significance (respectively with p values of 0.047 and 0.04).

Interestingly, when the relative percentage of phosphorylation of aPRPs was consider, Uc patients exhibited a lower % of the di-phosphorylated species of PRP-1 and PRP-3 with respect to controls (Table 6).

Table 6. Results of the statistical analysis performed on proteome of whole saliva. XIC peak areas (mean ± SD ($x10^8$), frequency and p-value obtained by T tests performed.

Proteins/Peptides Swiss-Prot Code	**Exper. (Theor) Av. Mass**	**XIC Peak Areas x 10^8 (Mean ± SD), Frequency**			**p-value**		
		FMF	**Uc**	**Ctrls**	**FMF vs Ctrls**	**Uc vs Ctrls**	**FMF vs Uc**
α-Def. 2	3370.4 ± 0.4	0.16±0.11	0.35±0.61	0.9±1.08	0.002 ↓FMF	0.04↓Uc	NS
P59665/6	(3371.0)	6/6	6/15	19/27			
α-Def. 3	3485.4 ± 0.4	0.07±0.06	0.2±0.4	0.6±1.01	0.01↓FMF	NS	NS
P59666	(3486.1)	5/6	5/15	18/27			
α-Def. 4	3708.5 ± 0.4	0/6	0.03±0.09	0.2±0.3	NA	0.02 ↓Uc	NA
P12838	(3709.5)		2/15	13/27			
Cystatin C	13342 ± 2	0.47±0.53	0.81±0.82	0.28±0.42	NS	0.02 ↑Uc	NS
P01034	(13343.1)	4/6	12/15	13/27			
Cystatin B-SSC	11300 ± 2	0.48±0.38	0.56±0.44	0.24±0.26	NS	0.007 ↑Uc	NS
P04080	(11300.7)	6/6	14/15	22/27			
Cystatin B-SSG	11486 ± 2	2.46±1.52	2.71±2.56	0.88±0.75	0.01 ↑FMF	0.003↑Uc	NS
	(11486.9)	6/6	15/15	27/27			

Proteins/Peptides Swiss-Prot Code	Exper. (Theor) Av. Mass	XIC Peak Areas x 10^8 (Mean ± SD), Frequency			p-value		
		FMF	Uc	Ctrls	FMF vs Ctrls	Uc vs Ctrls	FMF vs Uc
SLPI	11709 ± 2	0.15±0.16	0.12±0.13	0.03±0.09	0.006 ↑FMF	0.005 ↑Uc	NS
P03973	(11709.8)	4/6	8/15	4/27			
S100A9-SSG	13457 ± 2	0.74±0.76	1.48±3.11	0.28±0.69	0.008 ↑FMF	0.03 ↑Uc	NS
P06702	(13458.1)	5/6	8/15	5/27			
Hst-1 monophosph.	4927,1 ± 0.4	3.66±3.33	6.98±4.71	3.14±2.44	NS	0.02↑Uc	NS
P15515	(4928.1)	6/6	15/15	25/27			
Hst-1 nonphosph.	4842,6 ± 0.4	0.92±0.98	1.24±0.84	0.38±0.46	NS	0.002 ↑Uc	NS
	(4848.1)	6/6	14/15	17/27			
% phospho-Hst-1		77±17	79±12	91±8	NS	0.003 ↓Uc	NS
Hst-3	4062.4 ± 0.4	5.10±6.42	15.2±11.7	6.06±5.69	NS	0.01↑Uc	0.02↑Uc
P15516	(4062.4)	5/6	15/15	26/27			
His-3 Fr. 1/24	3036.0 ± 0.4	2.74±3.47	8.02±5.81	3.84±3.76	NS	0.02 ↑Uc	NS
	(3036.3)	6/6	15/15	26/27			
His-3 Fr. 1/25	3192.4 ± 0.4	1.07±1.68	2.84±2.33	1.08±1.14	NS	0.01 ↑Uc	NS
	(3192.5)	5/6	13/15	21/27			
Statherin diphosph.	5378.7 ± 0.4	14.4±11.1	32.8±17.9	14.0±11.1	NS	0.002↑Uc	0.02↑Uc
P02808	(5379.7)	6/6	15/15	27/27			
Statherin	5298.8 ± 0.4	0.23±0.24	0.48±0.46	0.17±0.35	NS	0.001↑Uc	NS
monophosph.	(5299.7)	5/6	15/15	16/27			
Statherin	5218.9 ± 0.4	0.02±0.03	0.04±0.08	0.007±0.03	0.02↑FMF	0.02↑Uc	NS
nonphosph.	(5219.8)	3/6	6/15	2/27			
Statherin Des1-9	4126.6 ± 0.4	1.00±0.83	1.72±1.60	0.60±0.59	NS	0.001↑Uc	NS
	(4127.6)	6/6	15/15	20/27			
PB peptide	5792.9 ± 0.5 (5792.7)	28±26 6/6	38±25 15/15	20±13 27/27	NS	0.004↑Uc	NS
PRP-1 diphosph.	15513 ± 2	74.8±45.9	134±61.9	95.2±58.4	NS	0.02↑Uc	0.047 ↑Uc
P02810	(15514.3)	6/6	15/15	26/27			
PRP-1	15433 ± 2	12.8±8.3	19.3±9.07	8.07±5.33	NS	0.0009↑Uc	NS
monophosph.	(15434.3)	6/6	15/15	26/27			
PRP-3 diphosph.	11160 ± 2	28.1±22.2	51.7±34.3	34.1±23.9	NS	0.02↑Uc	NS
P02810	(11161.5)	6/6	15/15	26/27			

Proteins/Peptides Swiss-Prot Code	Exper. (Theor) Av. Mass	XIC Peak Areas x 10^8 (Mean ± SD), Frequency			p-value		
		FMF	Uc	Ctrls	FMF vs Ctrls	Uc vs Ctrls	FMF vs Uc
PRP-3 monophosph.	11080 ± 2 (11081.6)	4.21±3.06 6/6	9.19±5.40 15/15	4.34±3.02 26/27	NS	0.007↑Uc	0.04↑Uc
PRP-3 nonphosph.	11000 ± 2 (11001.6)	0.31±0.35 6/6	0.54±0.29 15/15	0.20±0.19 20/27	NS	0.0006↑Uc	NS
P-C peptide *P02810*	4370.1 ± 0.4 (4370.8)	18.6±21.3 6/6	34.3±15.7 15/15	19±14.2 20/27	NS	0.002↑Uc	NS
% phosphorylation of PRP-1		81±11	86±4	91±4	NS	0.0008↓Uc	NS
% phosphorylation of PRP-3		80±13	84±5	87±5	NS	0.02↓Uc	NS

7.4 *Statherin and P-B peptide.*

The levels of statherin mono- di- and non-phosphorylated and its fragment Des1-9 were significant more abundant in Uc patients with respect to controls, di-phosphorylated statherin was more abundant in Uc patients also with respect the FMF group (Fig. 6, Table 6).

As shown in Fig. 6 and in Table 6, P-B peptide was significant more abundant in Uc patients with respect to controls (p values of 0.004). FMF patients showed levels of staherin species and P-B peptide comparable to those ones of controls. None difference was observed in FMF patients when compared with Uc group, even if, similarly to aPRPs, statherin species and P-B peptide showed a tendency to have lower levels in FMF than in Uc subjects.

7.5 *Histatins*

Hst-1 mono- and non-phosphorylated, Hst-3 and its fragments Hst-3 1/24 and Hst-3 1/25 were more concentrated in Uc patients with respect to controls (Fig. 8, Table 6).

None differences were observed in the level of histatins between FMF patients and controls, whereas FMF subjects exhibited lower levels of histains with respect to Uc group, if the difference was significant only for Hst-3 (Table 6).

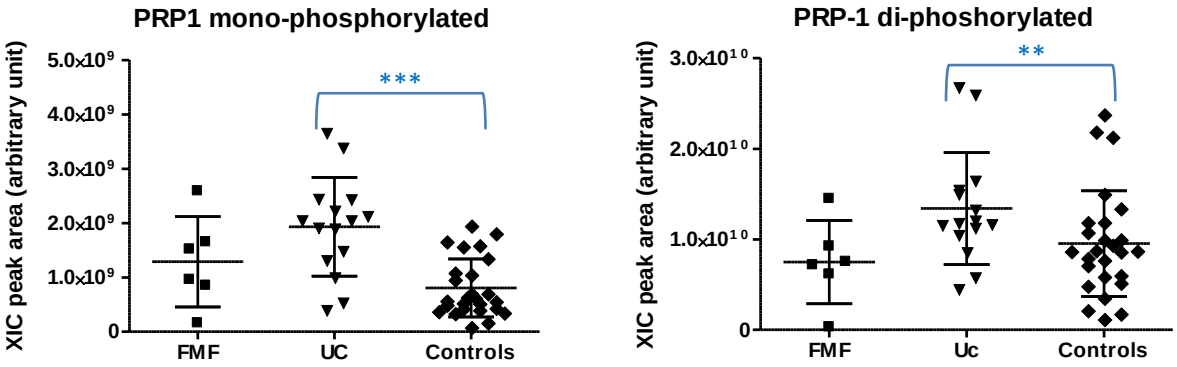
Analogously with aPRPs, the % of phosphorylated Hst-1 was lower in Uc patients than in controls (Table 4).

7.6 *Cystatins*

Among the cystatins analyzed, only cystatins B and C showed significant differences between patients and controls, none differences were observed in the leves of salivary cystatins, and cystatin A.

Cystatin B (both -SSG and -SSC forms), and cystatin C were more abundant in Uc patients with respect to controls (Table 6). Cystatin C, like the previously discussed proteins and peptides of glandular origin, was present in FMF patients with levels comparable to controls (Table 6, Fig. 9). Differently to cystatin C, the cystatin B species were more abundant in FMF patients with respect to controls, and this difference resulted significant for the glutathionylated species (Table 6, Fig. 9).

Fig. 5. Distribution of XIC peak areas among the three groups analyzed (FMF, Uc patients and controls) of PRP-1 mono- and di-phosphorylated, and PRP-3 mono- and di-phosphorylated.



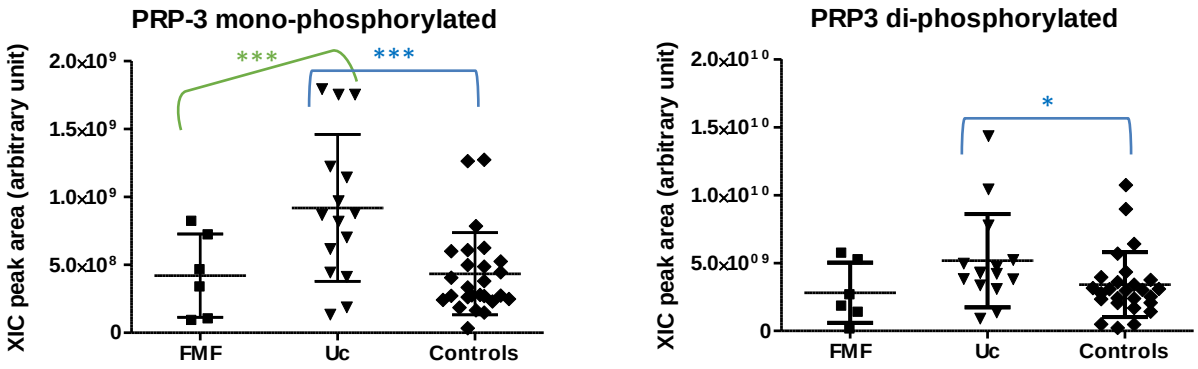


Fig. 6. Distribution of XIC peak areas among the three groups analyzed (FMF, Uc patients and controls) of statherin di-phosphorylated and P-B peptide.

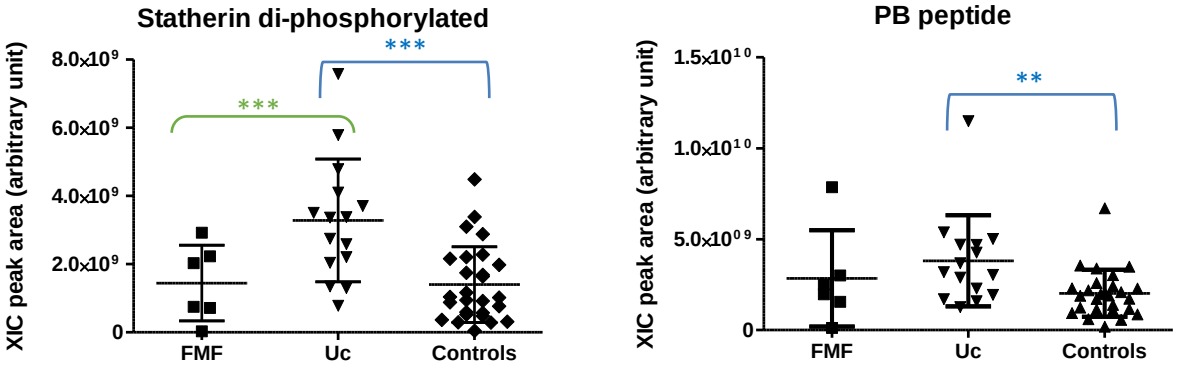
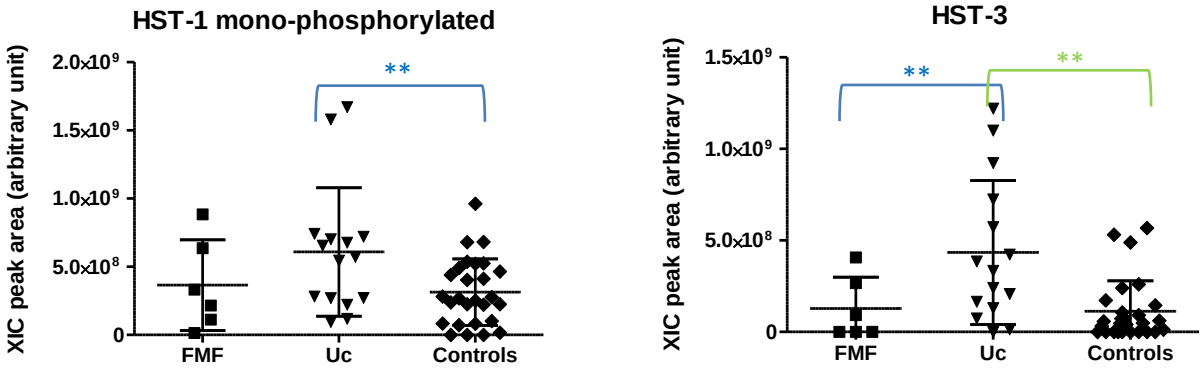


Fig. 8. Distribution of XIC peak areas among the three groups analyzed (FMF, Uc patients and controls) of HST-1 mono-phosphorylated, HST-3, and HST-3 1/24.



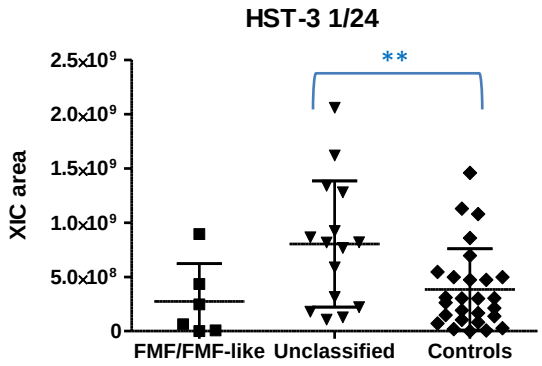
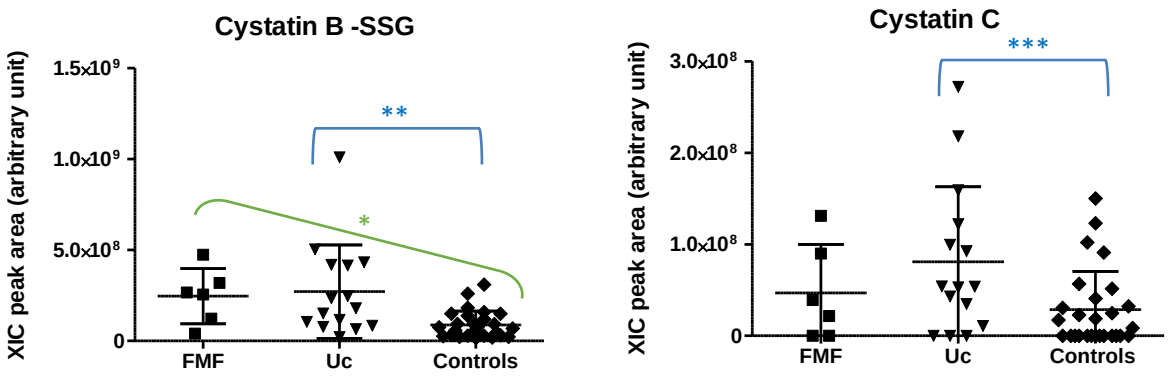


Fig. 9.

Distribution of XIC peak areas in the three groups (FMF, Uc patients and controls) of cystatin B-SSG and cystatin C.



7.7 *S100A8, S100A9, and antileukoproteinase.*

Results obtained from the analysis of S100A proteins showed significant variations only in the levels of glutathionylated S100A9(L). Both FMF and Uc patients exhibited increased levels of S100A9(L)-SSG with respect to controls (Fig. 10, Table 6).

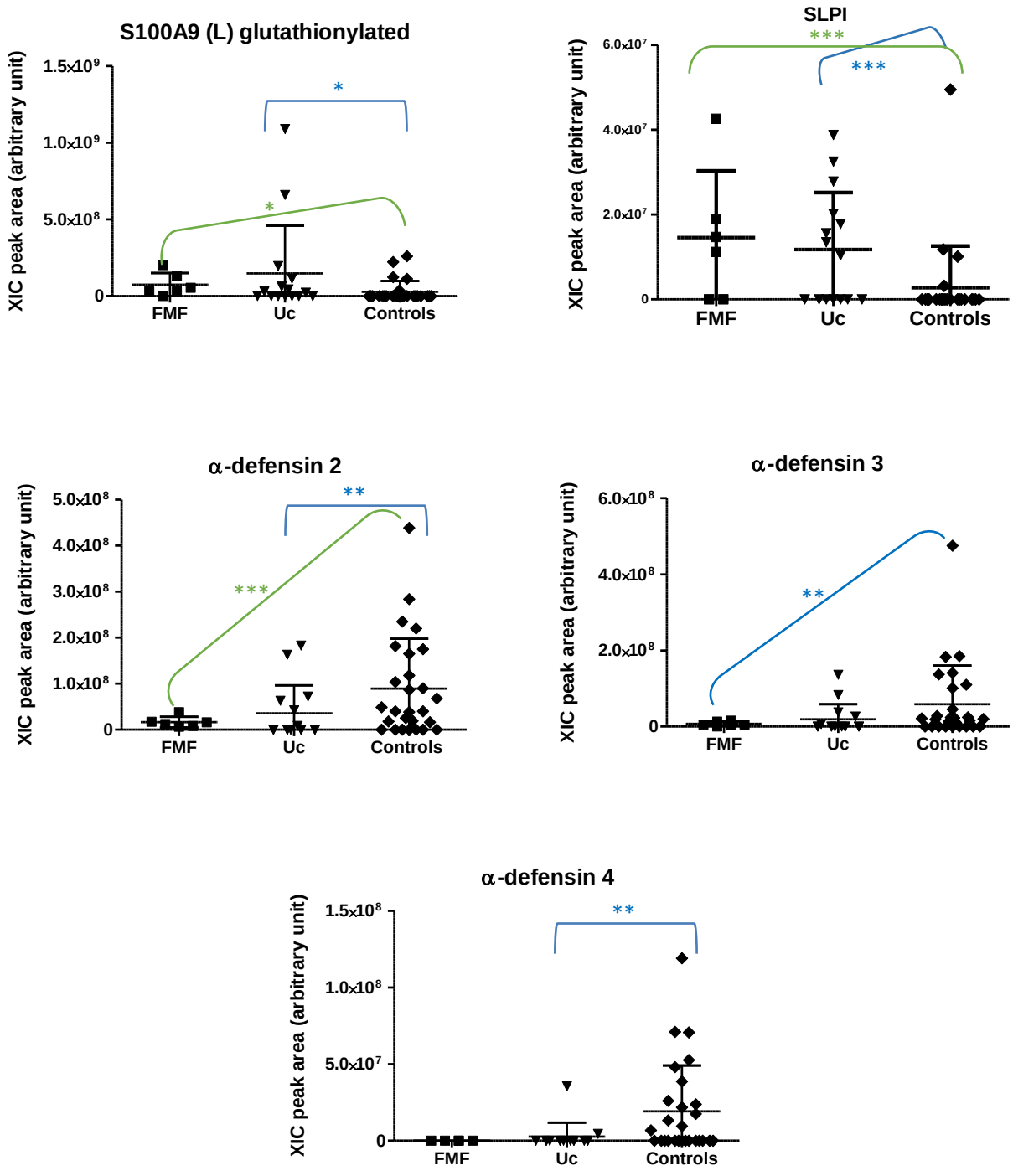
In the same way, in both groups was possible to observe a significant increase of concentration of antileukoproteinase (SLPI) (Fig. 10, Table 6).

7.8 *α-Defensins*

Differently to the other classes of proteins, α-defensin 2, 3 and 4 showed lower levels in both group of patients when compared to controls, and α-defensin 4 have never detected in

FMF patients (Table 6, Fig. 10). Whereas, α-defensin 1 did not exhibit changes in its concentration among the three groups.

Fig. 10. Distribution of XIC peak areas in the three groups (FMF, Uc patients and controls) of S100A9(L)-SSG, SLPI, α-defensin 2, 3 and 4.



8.0 *Discussion*

The study on salivary proteome of periodic fever syndrome patients shows important quantitative differences between patients and controls and between the two groups of patients.

These syndromes are characterized by a wide spectrum of clinical manifestations, involving different organs and apparatus, which can be extremely dangerous and lead to the amyloidosis, the most severe complication (Lane et al. 2013). The absence of a confident method of diagnosis and the overlap of some clinical symptoms with other similar diseases, implicate difficulty in diagnosis, and in the clinical classification; thus, the characterization and validation of new disease-specific biomarkers urges. As it has been demonstrated by this study and previous studies, salivary protein composition is affected by the pathology and may reflect the features of the disease (Cabras et al. 2013) ; (Cabras et al. 2015) ; (M Castagnola et al. 2011) ; (Armando Caseiro et al. 2012); (Wittkowski et al. 2008).

Patients affected by periodic fever syndromes show recurrent episodes of high fever, pharyngitis, cervical adenitis, and aphthous stomatitis. Many possible causative factors have been explored so far, including infectious agents, immunologic mechanisms and genetic predisposition (Theodoropoulou, Vanoni, and Hofer 2016), but the exact etiology remains unclear. Recent studies suggest a potential genetic origin for this entity due to the increased frequency of the condition in members of the same family (Akelma et al. 2013) ; (Mizuno et al. 2012), this fact strongly suggests that periodic fever syndromes have a potential genetic background. Due the clinical similarities with other monogenic periodic fever syndromes, several studies have investigated a possible genetic ethiology for this disease, and the principal gene responsible for this disease (Güncan et al. 2016); (Beheshtian et al. 2016) ; (Kilic et al. 2015) appear to be, the *Mediterranean Fever* (*MEFV*) gene. However, the data published on the subject are often contrastant. Federici et al. found mutations in the *MEFV* gene significantly correlated with the clinical manifestations in a portion of their patients (Federici et al. 2012); (Kubota et al. 2014) ; Salehzadeh et al. has reported statistically insignificant results (Salehzadeh et al. 2014), and in the study of Chandrakasan et al. no relevant gene mutations were found in their patients whatsoever (Chandrakasan et al. 2014). These findings might suggest a polygenic background in periodic fever syndrome with involvement of the inflammasome-related genes. Another hypothesis might be that periodic fever syndrome does not represent a homogenous entity and may involve some yet-uncharacterized genetic disease or attenuated forms of monogenic autoinflammatory

diseases (Gattorno and Caorsi 2009). These conclusions, agree with the heterogeneous set of *MEFV* mutations found in the patients enrolled in this study, and with the absence of correlation between the MEFV mutations and the clinical manifestations.

The results of this study highlighted that salivary protein profiles of the two groups of patients were different, relative to one another and with respect to healthy controls.

The levels of proteins and peptides secreted by salivary glands, except for salivary cystatins, were significantly increased in Uc patients with respect to both controls and FMF group.

Levels of proteins and peptides of not-glandular origin were significantly changed in both groups of patients with respect to controls.

Uc patients were characterized by high salivary levels of aPRPs and statherin. It is recognized that these proteins play an important role in the creation of a protective environment for the teeth, in the modulation of the bacteria adhesion to the oral surfaces. Moreover, they are involved in the formation of the protein pellicles covering the the oral surfaces (A. Bennick et al. 1983). Therefore, an increased secretion of these proteins might be useful in the oral cavity of patients to protect the oral mucosa from damages caused by the inflammatory rush. Another class of salivary proteins, playing a protective role in the mouth, is represented by histatins. Histatins play a key role in the innate defense system of the oral cavity having a powerful antibacterial and antifungal activity (White et al. 2009). In fact, Hst-3 1/25 is the most efficient salivary antimicrobial peptide in killing Candida albicans at physiological concentration (15–30 μM)(Raj, Edgerton, and Levine 1990). Histatin 1 is active in the wound healing processes occurring on the oral mucosa (Oudhoff et al. 2008).

In addition, an increase of Hst 3 and its fragments (Hst-3 1/24 and 1/25) in Uc patients, could be a response of the organism against local inflammation in the oral cavity, since that these peptides play a role in the down-regulation of pro-inflammatory cytokine, interleukin 8 and 18 (IL-8, IL-18), and tumor necrosis factor (TNF) (Imatani et al. 2000) ; (Imamura and Wang 2014).

In fact, in periodic fever syndromes, an increased serum levels of pro-inflammatory cytokines (IL-18, IL-6) and chemokines for activated T lymphocytes (CXCL10, CXCL9) was observed (Savic et al. 2012), and suggested the hypothesis of an inflammasome-mediated innate immune system activation. This hypothesis could explain the increased secretion not only of Hst-3, but also of the other salivary peptides and proteins involved in the innate immune defence of the oral cavity, such as aPRPs and statherin.

It is interesting to note that FMF patients did not exhibit oral aphthous ulceration whereas six out of 15 Uc patients show this clinical manifestation. According to these data, aPRP, statherin and histatins that are all implicated in oral defenses are more abundant in Uc patients, probably due to a major need of defense of oral mucosa.

Oddly, salivary cystatins did not show any variations in their levels in patients and controls, even if they are considered as proteins involved in the innate immune response of the oral cavity (Fábián et al. 2012). It is noteworthy that cystatins, differently to the other proteins and peptides secreted by salivary glands, are leader less secretory proteins that follow a not *trans*-Golgi secretory pathway (Lie et al. 2001). This may suggest that in our Uc patients was probably activated only the specific secretory processes involving the *trans*-Golgi network.

Interestingly, comparing the levels of phosphorylation of Hst-1, PRP-1, PRP-3, and statherin in the three groups, a significant hypo-phosphorylation of Hst-1, PRP-1 and PRP-3 in Uc patients when compared to controls, was observed, but not of statherin. This result suggests a decreased activity of the Fam20C kinase, a pleiotropic enzyme responsible for the phosphorylation of the proteins/peptides secreted by salivary glands on serine residues in the sequon SXE/S(Phos) or $S(X)_{3/4}(E/D/S(Phos))_3$ (Messana, Inzitari, et al. 2008) such as for the phosphorylation of secretory proteins/peptides in different tissues (Cozza et al. 2015).

Similar hypo-phosphorylation of Hst-1 and aPRPs was observed previously in subjects suffering of autistic disorders (Massimo Castagnola et al. 2008).

Cystatins are inhibitor of the cysteine proteinases and can suppress some viral infections (Ruzindana-Umunyana and Weber 2001); (Gu et al. 1995). Cystatin C promotes maturation of dendritic cells (DC) allowing them to take up antigens through exposure of major histocompatibility complex (MHC) molecules and during the regulation of inflammatory processes promotes maturation of macrophages and monocytes (Kopitar-Jerala 2006). Also cystatin B plays a role in inflammation, indeed, this protein is present in high levels during inflammation processes and as a result of bacterial infections (Zavasnik-Bergant 2008).

Having a pro-inflammatory activity (Maher et al. 2014), an increase salivary levels of cystatin C and cystatin B species in both group of patients might be due to a stimulation of inflammatory reaction caused by infections, and may be considered a clue of an inflammatory process occurring not only in the oral cavity but also in other organs and tissues. Interestingly, Uc patients were characterized by high levels of both cystatins, instead, FMF patients only by high level of cystatin B species, suggesting a different pro-inflammatory mechanism in the two patient groups.

The high levels of SLPI, in both patient groups, like for cystatins B and C, may be due to a response of organism at the inflammation reaction occurring in our patients.

SLPI plays a pivotal role in inflammation regulating neutrophils function through their inhibition (Stetler, Brewer, and Thompson 1986). This protein plays important roles in several areas: (i) as an antimicrobial agent, it can provide a first line of defense against infection (Reviglio et al. 2007), (ii) it controls the processing of inflammatory mediators and protects the host from excessive tissue damage by proteolytic enzymes released during inflammation (McKiernan, McElvaney, and Greene 2011), (iii) it suppresses inflammatory responses by controlling the activity of the transcription factor NFκB (Wen et al. 2011), (iv) it regulates the production and pro-immunogenic function of neutrophil extracellular traps (Koizumi et al. 2008), (v) it fosters repair and is a component of the molecular machinery that controls cell growth, differentiation and apoptosis (Majchrzak-Gorecka et al. 2016).

The result of SLPI actions is to counteract excessive inflammatory responses and to initiate healing processes, although SLPI can also potentially participate in the pathological outcome of inflammatory diseases (Song et al. 1999).

Both FMF patients and Uc patients showed low levels of α-defensins. These peptides can stimulate the phagocytosis of neutrophils (Ericksen et al. 2005) and exert anti-bacterial actions, recent studies suggest that these antimicrobial peptides contribute to host defense and homeostasis of tissues and biological fluids by recruiting immune cells in the site of infection (Suarez-Carmona et al. 2015) ; (Ganz 2003) ; (Bevins 2013).

So, their low concentration, in saliva of the patients enrolled in this studies, may be correlated with a lower phagocytic activity by neutrophils or with their dysregulation, and, thus, to a major susceptibility of the patients to infections and to a dysregulated inflammatory rush.

S100A9 plays a prominent role in the regulation of inflammatory processes and immune response. It can induce neutrophil chemotaxis, adhesion, can increase the bactericidal activity of neutrophils by promoting phagocytosis via activation of SYK, PI3K/AKT, and ERK1/2 and can induce degranulation of neutrophils by a MAPK-dependent mechanism. S100A9 it has been shown to have a proinflammatory activity. Its proinflammatory activity includes recruitment of leukocytes, promotion of cytokine and chemokine production, and regulation of leukocyte adhesion and migration (Ryckman et al. 2003).

The glutathionylated S100A9(L) is the predominant form in saliva. An increase level of this protein is due to systemic inflammation that afflicts patients.

The different protein profile observed in the acid soluble fraction of saliva of the two groups of patients may reflect pathogenic mechanisms distinctive for patients affected by FMF and for those ones affected by not classifiable periodic fever syndromes.
In these last, it appears to be a great involvement of proteins and peptides originated from salivary glands, and probably the triggering of protection tools of the oral cavity in response to injuries caused by the pathology, like the aphthous stomatitis.
Furthermore, the alteration of other proteins not strictly salivary, reflects the presence of systemic inflammation typical of this disease.

9.0 *Conclusions*

Due to the small number of subjects available, the statistical power of this study suffers from some limitations; though, the rarity of this disease reduces the chance of recruiting a larger group of patients.
Therefore, the validation of our observations requires a larger study including additional information as well as longitudinal sampling and analysis.
Despite these limitations, the present study bears some novel points. In fact, this is the first study that investigates the salivary proteome and peptidome in periodic fever syndromes and some interesting results have been obtained.
Uc patients that present aphthous stomatitis have the concentration of proteins implicated in oral defenses is more abundant with respect to FMF patients which in fact do not exhibit this manifestation. S100A9(L)-SSG protein, cystatin B and C, SLPI showed altered levels in both groups reflecting the effects of systemic inflammation in these patients.
In contrast α-defensins, those have an anti-inflammatory function, show lower levels in patients when compared to controls due, probably, to a dysregulation of innate immune responses.
One of the major difficulties of these inflammatory diseases is due to the variability of the clinical manifestations of patients that impedes the creation of homogeneous group on which perform a powerful statistical analysis.
Despite this, we have been able to highlight that proteomic and peptidomic modifications observed in patients with respect to controls were distinctive for FMF and Uc diseases to indicating that saliva reflects the typical features of disorders.
Further studies to confirm the variations of the salivary proteome/peptidome observed and their disease-specificity as well as validation of the results by orthogonal methods will allow

in the future establishing the actual applicability and the diagnostic power of a salivary test for these pathologies.

Bibliography

Abe, N, T Kadowaki, K Okamoto, K Nakayama, M Ohishi, and K Yamamoto. 1998. "Biochemical and Functional Properties of Lysine-Specific Cysteine Proteinase (Lys-Gingipain) as a Virulence Factor of Porphyromonas Gingivalis in Periodontal Disease." *J. Biochem. (Tokyo)* 123 (2): 305–12.

Abe, T, N Kobayashi, K Yoshimura, B C Trapnell, H Kim, R C Hubbard, M T Brewer, R C Thompson, and R G Crystal. 1991. "Expression of the Secretory Leukoprotease Inhibitor Gene in Epithelial Cells." *The Journal of Clinical Investigation* 87 (6). United States: 2207–15. doi:10.1172/JCI115255.

Abrahamson, M, A J Barrett, G Salvesen, and A Grubb. 1986. "Isolation of Six Cysteine Proteinase Inhibitors from Human Urine. Their Physicochemical and Enzyme Kinetic Properties and Concentrations in Biological Fluids." *The Journal of Biological Chemistry* 261 (24): 11282–89. doi:10.1002/elps.1150181516.

Addona, Terri A, Xu Shi, Hasmik Keshishian, D R Mani, Michael Burgess, Michael A Gillette, Karl R Clauser, et al. 2011. "A Pipeline That Integrates the Discovery and Verification of Plasma Protein Biomarkers Reveals Candidate Markers for Cardiovascular Disease." *Nat Biotech* 29 (7). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 635–43.

Akelma, Ahmet Zulfikar, Mehmet Nevzat Cizmeci, Mehmet Kenan Kanburoglu, Emin Mete, Davut Bozkaya, Naile Tufan, and Ferhat Catal. 2013. "Is PFAPA Syndrome Really a Sporadic Disorder or Is It Genetic?" *Medical Hypotheses* 81 (2). Elsevier Ltd: 279–81. doi:10.1016/j.mehy.2013.04.030.

Aksentijevich, Ivona, Elon Pras, Luis Gruberg, Yang Shen, Katherine Holman, Sharon Helling, Leandrea Prosen, et al. 1993. "Refined Mapping of the Gene Causing Familial Mediterranean Fever, by Linkage and Homozygosity Studies." *Am. J. Hum. Genet* 531: 451–46.

Altelaar, A F Maarten, Javier Munoz, and Albert J R Heck. 2012. "Next-Generation Proteomics: Towards an Integrative View of Proteome Dynamics." *Nature Reviews Genetics* 14 (1). Nature Publishing Group: 35–48. doi:10.1038/nrg3356.

Amado, Francisco, Maria João Calheiros Lobo, Pedro Domingues, José Alberto Duarte, and Rui Vitorino. 2010. "Salivary Peptidomics." *Expert Review of Proteomics* 7 (5). Taylor & Francis: 709–21. doi:10.1586/epr.10.48.

Amerongen, a V Nieuw, and E C I Veerman. 2002. "Saliva--the Defender of the Oral Cavity." *Oral Diseases* 8 (1): 12–22. doi:10.1034/j.1601-0825.2002.1o816.x.

Anderson, N Leigh. 2010. "Libraries of Specific Assays Covering Whole Proteomes: From Yeast to Man." *Clinical Chemistry* 56 (10): 1521 LP-1522.

Andrassy, Martin, John Igwe, Frank Autschbach, Christian Volz, Andrew Remppis, Markus F Neurath, Erwin Schleicher, et al. 2006. "Posttranslationally Modified Proteins as Mediators of Sustained Intestinal Inflammation." *The American Journal of Pathology* 169 (4): 1223–37. doi:10.2353/ajpath.2006.050713.

Atkinson A.J., Jr, W. A. Colburn, V. G. DeGruttola, D. L. DeMets, G. J. Downing, D. F. Hoth, J. A. Oates, et al. 2001. "Biomarkers and Surrogate Endpoints: Preferred Definitions and Conceptual Framework." *Clinical Pharmacology and Therapeutics* 69 (3): 89–95. doi:10.1067/mcp.2001.113989.

Azen, E a, K Minaguchi, P Latreille, and H S Kim. 1990. "Alleles at the PRB3 Locus Coding for a Disulfide-Bonded Human Salivary Proline-Rich Glycoprotein (Gl 8) and a Null in an Ashkenazi Jew." *American Journal of Human Genetics* 47 (4): 686–97.

Badamchian, Mahnaz, Ali A Damavandy, Hadi Damavandy, Sonal D Wadhwa, Barrett Katz, and Allan L Goldstein. 2007. "Identification and Quantification of Thymosin beta4 in Human Saliva and Tears." *Annals of the New York Academy of Sciences* 1112: 458–65. doi:10.1196/annals.1415.046.

Bandhakavi, Sricharan, Matthew D Stone, Getiria Onsongo, Susan K Van Riper, and Timothy J Griffin. 2009. "A Dynamic Range Compression and Three-Dimensional Peptide Fractionation Analysis Platform Expands Proteome Coverage and the Diagnostic Potential of Whole Saliva." *Journal of Proteome Research* 8 (12). American Chemical Society: 5590–5600. doi:10.1021/pr900675w.

Baron, A C, A A DeCarlo, and J D B Featherstone. 1999. "Functional Aspects of the Human Salivary Cystatins in the Oral Environment." *Oral Diseases* 5: 234–40.

Beheshtian, Maryam, Nasim Izadi, Gernot Kriegshauser, Kimia Kahrizi, Elham Parsi Mehr, Maryam Rostami, Masoumeh Hosseini, et al. 2016. "Prevalence of Common MEFV Mutations and Carrier Frequencies in a Large Cohort of Iranian Populations." *Journal of Genetics* 95 (3): 667–74. doi:10.1007/s12041-016-0682-6.

Ben-Zvi, I, B Brandt, Y Berkun, M Lidar, and A Livneh. 2012. "The Relative Contribution of Environmental and Genetic Factors to Phenotypic Variation in Familial Mediterranean Fever (FMF)." *Gene* 491 (2): 260–63. doi:10.1016/j.gene.2011.10.005.

Ben-Zvi, Ilan, Corinne Herskovizh, Olga Kukuy, Yonatan Kassel, Chagai Grossman, and Avi Livneh. 2015. "Familial Mediterranean Fever without MEFV Mutations: A Case-Control Study." *Orphanet Journal of Rare Diseases* 10 (1): 34. doi:10.1186/s13023-015-0252-7.

Bencharit, Sompop, Sandra K Altarawneh, Sarah Schwartz Baxter, Jim Carlson, Gary F Ross, Michael B Border, C Russell Mack, et al. 2012. "Elucidating Role of Salivary Proteins in Denture Stomatitis Using a Proteomic Approach." *Molecular bioSystems* 8 (12): 3216–23. doi:10.1039/c2mb25283j.

Bennick, a. 1987. “Structural and Genetic Aspects of Proline-Rich Proteins.” *Journal of Dental Research* 66 (2): 457–61. doi:10.1177/00220345870660021201.

Bennick, A., G. Chau, R. Goodlin, S. Abrams, D. Tustian, and G. Madapallimattam. 1983. “The Role of Human Salivary Acidic Proline-Rich Proteins in the Formation of Acquired Dental Pellicle in Vivo and Their Fate after Adsorption to the Human Enamel Surface.” *Archives of Oral Biology* 28 (1): 19–27. doi:10.1016/0003-9969(83)90022-5.

Bennick, A., A. C. McLaughlin, A. A. Grey, and G. Madapallimattam. 1981. “The Location and Nature of Calcium-Binding Sites in Salivary Acidic Proline-Rich Phosphoproteins.” *Journal of Biological Chemistry* 256 (10): 4741–46.

Bennick, A, D Kells, and G Madapallimattam. 1983. “Interaction of Calcium Ions and Salivary Acidic Proline-Rich Proteins with Hydroxyapatite. A Possible Aspect of Inhibition of Hydroxyapatite Formation.” *The Biochemical Journal* 213 (1): 11–20.

Bennick, Anders. 2002. “Interaction of Plant Polyphenols with Salivary Proteins.” *Critical Reviews in Oral Biology and Medicine* 13 (2): 184–96. doi:10.1177/154411130201300208.

Bernot, a, C da Silva, J L Petit, C Cruaud, C Caloustian, V Castet, M Ahmed-Arab, et al. 1998. “Non-Founder Mutations in the MEFV Gene Establish This Gene as the Cause of Familial Mediterranean Fever (FMF).” *Human Molecular Genetics* 7 (8): 1317–25. doi:9668175.

Bevins, Charles. 2013. “Innate Immune Functions of α-Defensins in the Small Intestine.” *Digestive Diseases (Basel, Switzerland)* 31 (3–4): 299–304. doi:10.1159/000354681.

Blankenvoorde, M F, Y M Henskens, W van’t Hof, E C Veerman, and A V Nieuw Amerongen. 1996. “Inhibition of the Growth and Cysteine Proteinase Activity of Porphyromonas Gingivalis by Human Salivary Cystatin S and Chicken Cystatin.” *Biol Chem* 377 (12): 847–50.

Bobek, L a, and M J Levine. 1992. “Cystatins--Inhibitors of Cysteine Proteinases.” *Critical Reviews in Oral Biology and Medicine* 3 (4): 307–32. doi:10.1177/10454411920030040101.

Bobek, Libuse A, and Michael J Levine. 1992. “Cystatins - Inhibitors of Cysteine Proteinases.” *Critical Reviews in Oral Biology and Medicine* 3 (4): 307–32. doi:10.1177/10454411920030040101.

Bogdanov, Bogdan, and Richard D. Smith. 2005. “Proteomics by Fticr Mass Spectrometry: TOP down and Bottom up.” *Mass Spectrometry Reviews* 24 (2): 168–200. doi:10.1002/mas.20015.

Brand, Henk S., Antoon J M Ligtenberg, and Enno C I Veerman. 2014. “Saliva and Wound Healing.” In *Saliva: Secretion and Functions*, 24:52–60. doi:10.1159/000358784.

Cabras, Tiziana, Roberto Boi, Elisabetta Pisano, Federica Iavarone, Chiara Fanali, Sonia Nemolato, Gavino Faa, Massimo Castagnola, and Irene Messana. 2012. "HPLC-ESI-MS and MS/MS Structural Characterization of Multifucosylated N-Glycoforms of the Basic Proline-Rich Protein IB-8a CON1+ in Human Saliva." *Journal of Separation Science* 35 (9): 1079–86. doi:10.1002/jssc.201101066.

Cabras, Tiziana, Chiara Fanali, Joana A. Monteiro, Francisco Amado, Rosanna Inzitari, Claudia Desiderio, Emanuele Scarano, Bruno Giardina, Massimo Castagnola, and Irene Messana. 2007. "Tyrosine Polysulfation of Human Salivary Histatin 1. A Post-Translational Modification Specific of the Submandibular Gland." *Journal of Proteome Research* 6 (7): 2472–80. doi:10.1021/pr0700706.

Cabras, Tiziana, Federica Iavarone, Barbara Manconi, Alessandra Olianas, Maria Teresa Sanna, Massimo Castagnola, and Irene Messana. 2014. "Top-down Analytical Platforms for the Characterization of the Human Salivary Proteome." *Bioanalysis* 6 (4): 563–81. doi:10.4155/bio.13.349.

Cabras, Tiziana, Rosanna Inzitari, Chiara Fanali, Emanuele Scarano, Maria Patamia, Maria T. Sanna, Elisabetta Pisano, Bruno Giardina, Massimo Castagnola, and Irene Messana. 2006. "HPLC-MS Characterization of Cyclo-Statherin Q-37, a Specific Cyclization Product of Human Salivary Statherin Generated by Transglutaminase 2." *Journal of Separation Science* 29 (17): 2600–2608. doi:10.1002/jssc.200600244.

Cabras, Tiziana, Barbara Manconi, Federica Iavarone, Chiara Fanali, Sonia Nemolato, Antonella Fiorita, Emanuele Scarano, et al. 2012. "RP-HPLC-ESI-MS Evidenced That Salivary Cystatin B Is Detectable in Adult Human Whole Saliva Mostly as S-Modified Derivatives: S-Glutathionyl, S-Cysteinyl and S-S 2-Mer." *Journal of Proteomics* 75 (3): 908–13. doi:10.1016/j.jprot.2011.10.006.

Cabras, Tiziana, Melania Melis, Massimo Castagnola, Alessandra Padiglia, Beverly J. Tepper, Irene Messana, and Iole Tomassini Barbarossa. 2012. "Responsiveness to 6-N-Propylthiouracil (PROP) Is Associated with Salivary Levels of Two Specific Basic Proline-Rich Proteins in Humans." *PLoS ONE* 7 (2). doi:10.1371/journal.pone.0030962.

Cabras, Tiziana, Elisabetta Pisano, Roberto Boi, Alessandra Olianas, Barbara Manconi, Rosanna Inzitari, Chiara Fanali, Bruno Giardina, Massimo Castagnola, and Irene Messana. 2009. "Age-Dependent Modifications of the Human Salivary Secretory Protein Complex Research Articles." *Journal of Proteome Research* 8: 4126–34.

Cabras, Tiziana, Elisabetta Pisano, Andrea Mastinu, Gloria Denotti, Pietro Paolo Pusceddu, Rosanna Inzitari, Chiara Fanali, Sonia Nemolato, Massimo Castagnola, and Irene Messana. 2010. "Alterations of the Salivary Secretory Peptidome Profile in Children Affected by Type 1 Diabetes." *Molecular & Cellular Proteomics : MCP* 9 (10): 2099–2108. doi:10.1074/mcp.M110.001057.

Cabras, Tiziana, Elisabetta Pisano, Caterina Montaldo, Maria Rita Giuca, Federica Iavarone, Giuseppe Zampino, Massimo Castagnola, and Irene Messana. 2013. "Significant Modifications of the Salivary Proteome Potentially

Associated with Complications of Down Syndrome Revealed by Top-down Proteomics." *Molecular & Cellular Proteomics : MCP* 12 (7): 1844–52. doi:10.1074/mcp.M112.026708.

Cabras, Tiziana, Monica Sanna, Barbara Manconi, Daniela Fanni, Luigi Demelia, Orazio Sorbello, Federica Iavarone, Massimo Castagnola, Gavino Faa, and Irene Messana. 2015. "Proteomic Investigation of Whole Saliva in Wilson â€™S Disease." *Journal of Proteomics* 128. Elsevier B.V.: 154–63. doi:10.1016/j.jprot.2015.07.033.

Caseiro, A, R Ferreira, A Padrao, C Quintaneiro, A Pereira, R Marinheiro, R Vitorino, and F Amado. 2013. "Salivary Proteome and Peptidome Profiling in Type 1 Diabetes Mellitus Using a Quantitative Approach." *J Proteome Res* 12 (4): 1700–1709. doi:10.1021/pr3010343.

Caseiro, Armando, Rui Vitorino, António S Barros, Rita Ferreira, Maria João Calheiros-Lobo, Davide Carvalho, José Alberto Duarte, and Francisco Amado. 2012. "Salivary Peptidome in Type 1 Diabetes Mellitus." *Biomedical Chromatography : BMC* 26 (5): 571–82. doi:10.1002/bmc.1677.

Castagnola, M, P M Picciotti, I Messana, C Fanali, A Fiorita, T Cabras, L Calò, et al. 2011. "Potential Applications of Human Saliva as Diagnostic Fluid." *Acta Otorhinolaryngologica Italica : Organo Ufficiale Della Società Italiana Di Otorinolaringologia E Chirurgia Cervico-Facciale* 31 (6): 347–57.

Castagnola, Massimo, Tiziana Cabras, Federica Iavarone, Chiara Fanali, Sonia Nemolato, Giusy Peluso, Silvia Laura Bosello, Gavino Faa, Gianfranco Ferraccioli, and Irene Messana. 2012. "The Human Salivary Proteome: A Critical Overview of the Results Obtained by Different Proteomic Platforms." *Expert Review of Proteomics* 9 (1): 33–46. doi:10.1586/epr.11.77.

Castagnola, Massimo, Tiziana Cabras, Federica Iavarone, Federica Vincenzoni, Alberto Vitali, Elisabetta Pisano, Sonia Nemolato, et al. 2012. "Top-down Platform for Deciphering the Human Salivary Proteome." *The Journal of Maternal-Fetal & Neonatal MedicineOnline) Journal The Journal of Maternal-Fetal & Neonatal Medicine The Journal of Maternal-Fetal and Neonatal Medicine* 25 (25S5): 1476–7058. doi:10.3109/14767058.2012.714647.

Castagnola, Massimo, Tiziana Cabras, Alberto Vitali, Maria Teresa Sanna, and Irene Messana. 2011. "Biotechnological Implications of the Salivary Proteome." *Trends in Biotechnology*. Elsevier Ltd. doi:10.1016/j.tibtech.2011.04.002.

Castagnola, Massimo, Rosanna Inzitari, Chiara Fanali, Federica Iavarone, Alberto Vitali, Claudia Desiderio, Giovanni Vento, et al. 2011. "The Surprising Composition of the Salivary Proteome of Preterm Human Newborn." *Molecular & Cellular Proteomics : MCP* 10 (1): M110.003467. doi:10.1074/mcp.M110.003467.

Castagnola, Massimo, Rosanna Inzitari, Diana Valeria Rossetti, Chiara Olmi, Tiziana Cabras, Vincenzo Piras, Paola Nicolussi, et al. 2004. "A Cascade of 24 Histatins (Histatin 3 Fragments) in Human Saliva: Suggestions for a Pre-Secretory Sequential Cleavage Pathway." *Journal of Biological Chemistry* 279 (40): 41436–43.

doi:10.1074/jbc.M404322200.

Castagnola, Massimo, Irene Messana, Rosanna Inzitari, Chiara Fanali, Tiziana Cabras, Alessandra Morelli, Anna Maria Pecoraro, Giovanni Neri, Maria Giulia Torrioli, and Fiorella Gurrieri. 2008. "Hypo-Phosphorylation of Salivary Peptidome as a Clue to the Molecular Pathogenesis of Autism Spectrum Disorders." *Journal of Proteome Research* 7 (12): 5327–32. doi:10.1021/pr8004088.

Castello, Alfredo, Bernd Fischer, Katrin Eichelbaum, Rastislav Horos, Benedikt M. Beckmann, Claudia Strein, Norman E. Davey, et al. 2012. "Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins." *Cell* 149 (6). Elsevier Inc.: 1393–1406. doi:10.1016/j.cell.2012.04.031.

Celis, Julio E., and Jose M. A. Moreira. 2008. "Clinical Proteomics." *Mol. Cell. Proteomics* 7 (10): 1779.

Chae, Jae Jin, Geryl Wood, Seth L Masters, Katharina Richard, Grace Park, Brian J Smith, and Daniel L Kastner. 2006. "The B30.2 Domain of Pyrin, the Familial Mediterranean Fever Protein, Interacts Directly with Caspase-1 to Modulate IL-1beta Production." *Proceedings of the National Academy of Sciences of the United States of America* 103 (26). United States: 9982–87. doi:10.1073/pnas.0602081103.

Chae, Jae Jin, Geryl Wood, Katharina Richard, Howard Jaffe, Nona T Colburn, Seth L Masters, Deborah L Gumucio, Nitza G Shoham, and Daniel L Kastner. 2008. "The Familial Mediterranean Fever Protein, Pyrin, Is Cleaved by Caspase-1 and Activates NF-kappaB through Its N-Terminal Fragment." *Blood* 112 (5). United States: 1794–1803. doi:10.1182/blood-2008-01-134932.

Chaly, Y V, E M Paleolog, T S Kolesnikova, I I Tikhonov, E V Petratchenko, and N N Voitenok. 2000. "Neutrophil Alpha-Defensin Human Neutrophil Peptide Modulates Cytokine Production in Human Monocytes and Adhesion Molecule Expression in Endothelial Cells." *European Cytokine Network* 11 (2): 257–66.

Chandrakasan, Shanmuganathan, Saurabh Chiwane, Matthew Adams, and Basil M. Fathalla. 2014. "Clinical and Genetic Profile of Children with Periodic Fever Syndromes from a Single Medical Center in South East Michigan." *Journal of Clinical Immunology* 34 (1): 104–13. doi:10.1007/s10875-013-9960-8.

Chiusolo, Patrizia, Sabrina Giammarco, Chiara Fanali, Silvia Bellesi, Elisabetta Metafuni, Simona Sica, Federica Iavarone, et al. 2013. "Salivary Proteomic Analysis and Acute Graft-versus-Host Disease after Allogeneic Hematopoietic Stem Cell Transplantation." *Biology of Blood and Marrow Transplantation* 19 (6): 888–92. doi:10.1016/j.bbmt.2013.03.011.

Cozza, Giorgio, Mauro Salvi, Sourav Banerjee, Elena Tibaldi, Vincent S. Tagliabracci, Jack E. Dixon, and Lorenzo A. Pinna. 2015. "A New Role for Sphingosine: Up-Regulation of Fam20C, the Genuine Casein Kinase That Phosphorylates Secreted Proteins." *Biochimica et Biophysica Acta - Proteins and Proteomics* 1854 (10): 1718–26. doi:10.1016/j.bbapap.2015.04.023.

Demirkaya, Erkan, Burak Erer, Seza Ozen, and Eldad Ben-Chetrit. 2016. “Efficacy and Safety of Treatments in Familial Mediterranean Fever: A Systematic Review.” *Rheumatology International*. doi:10.1007/s00296-015-3408-9.

Dickinson, D. P. 2002. “Salivary (SD-Type) Cystatins: Over One Billion Years in the Making--but to What Purpose?” *Critical Reviews in Oral Biology and Medicine : An Official Publication of the American Association of Oral Biologists* 13 (6): 485–508. doi:10.1177/154411130201300606.

Donato, Rosario. 2003. “Intracellular and Extracellular Roles of S100 Proteins.” *Microscopy Research and Technique* 60 (6): 540–51. doi:10.1002/jemt.10296.

Dowling, Paul, Robert Wormald, Paula Meleady, Michael Henry, Aongus Curran, and Martin Clynes. 2008. “Analysis of the Saliva Proteome from Patients with Head and Neck Squamous Cell Carcinoma Reveals Differences in Abundance Levels of Proteins Associated with Tumour Progression and Metastasis.” *Journal of Proteomics* 71 (2): 168–75. doi:10.1016/j.jprot.2008.04.004.

Eckert, Richard L., Ann Marie Broome, Monica Ruse, Nancy Robinson, David Ryan, and Kathleen Lee. 2004. “S100 Proteins in the Epidermis.” *Journal of Investigative Dermatology*. doi:10.1111/j.0022-202X.2004.22719.x.

Edgeworth, Jonathan, Michael Gorman, Robert Bennett, Paul Freemont, and Nancy Hogg. 1991. “Identification of p8,14 as a Highly Abundant Heterodimeric Calcium Binding Protein Complex of Myeloid Cells.” *Journal of Biological Chemistry* 266 (12): 7706–13.

Ellias, Mohd Faiz, Shahrul Hisham Zainal Ariffin, Saiful Anuar Karsani, Mariati Abdul Rahman, Shahidan Senafi, and Rohaya Megat Abdul Wahab. 2012. “Proteomic Analysis of Saliva Identifies Potential Biomarkers for Orthodontic Tooth Movement.” *TheScientificWorldJournal* 2012: 647240. doi:10.1100/2012/647240.

Ericksen, Bryan, Zhibin Wu, Wuyuan Lu, and Robert I Lehrer. 2005. “Antibacterial Activity and Specificity of the Antibacterial Activity and Specificity of the Six.” *Antimicrobial Agents and Chemotherapy* 49 (1): 8–15. doi:10.1128/AAC.49.1.269.

Fábián, Tibor Károly, Péter Hermann, Anita Beck, Pál Fejérdy, and Gábor Fábián. 2012. “Salivary Defense Proteins: Their Network and Role in Innate and Acquired Oral Immunity.” *International Journal of Molecular Sciences*. doi:10.3390/ijms13044295.

Fahey, John V, and Charles R Wira. 2002. “Effect of Menstrual Status on Antibacterial Activity and Secretory Leukocyte Protease Inhibitor Production by Human Uterine Epithelial Cells in Culture.” *The Journal of Infectious Diseases* 185: 1606–13. doi:10.1086/340512.

Farquhar, Carey, Thomas C VanCott, Dorothy A Mbori-Ngacha, Lena Horani, Rose K Bosire, Joan K Kreiss,

Barbra A Richardson, and Grace C John-Stewart. 2002. “Salivary Secretory Leukocyte Protease Inhibitor Is Associated with Reduced Transmission of Human Immunodeficiency Virus Type 1 through Breast Milk.” *The Journal of Infectious Diseases* 186 (8): 1173–76. doi:10.1086/343805.

Federici, Silvia, Giuseppina Calcagno, Martina Finetti, Romina Gallizzi, Antonella Meini, Agata Vitale, Francesco Caroli, et al. 2012. “Clinical Impact of MEFV Mutations in Children with Periodic Fever in a Prevalent Western European Caucasian Population.” *Annals of the Rheumatic Diseases* 71 (12): 1961–65. doi:10.1136/annrheumdis-2011-200977.

Federici, Silvia, Maria Pia Sormani, Seza Ozen, Helen J Lachmann, Gayane Amaryan, Patricia Woo, Isabelle Koné-Paut, et al. 2015. “Evidence-Based Provisional Clinical Classification Criteria for Autoinflammatory Periodic Fevers.” *Annals of the Rheumatic Diseases* 74 (5): 799–805. doi:10.1136/annrheumdis-2014-206580.

French FMF Consortium. 1997. “A Candidate Gene for Familial Mediterranean Fever.” *Nature Genetics* 17 (1): 25–31. doi:10.1038/ng0997-25.

Ganz, Tomas. 2003. “Defensins: Antimicrobial Peptides of Innate Immunity.” *Nature Reviews. Immunology* 3 (9): 710–20. doi:10.1038/nri1180.

Ganz, Tomas, and Robert Lehrer. 1994. “Defensins.” *Current Opinion in Immunology* 6 (4): 584–89.

Garcia-Gonzalez, A, and M H Weisman. 1992. “The Arthritis of Familial Mediterranean Fever.” *Seminars in Arthritis and Rheumatism* 22 (3): 139–50.

Gattorno, Authors Marco, and Roberta Caorsi. 2009. “Differentiating PFAPA Syndrome From Monogenic Periodic Fevers” 124 (4). doi:10.1542/peds.2009-0088.

Giancane, Gabriella, Nienke M Ter Haar, Nico Wulffraat, Sebastiaan J Vastert, Karyl Barron, Veronique Hentgen, Tilmann Kallinich, et al. 2015. “Evidence-Based Recommendations for Genetic Diagnosis of Familial Mediterranean Fever.” *Annals of the Rheumatic Diseases* 74 (4): 635–41. doi:10.1136/annrheumdis-2014-206844.

Gibbons, R. J., D. I. Hay, and D. H. Schlesinger. 1991. “Delineation of a Segment of Adsorbed Salivary Acidic Proline-Rich Proteins Which Promotes Adhesion of Streptococcus Gordonii to Apatitic Surfaces.” *Infection and Immunity* 59 (9): 2948–54.

Giusti, Laura, Chiara Baldini, Federica Ciregia, Gino Giannaccini, Camillo Giacomelli, Francesca de Feo, Andrea Delle Sedie, et al. 2010. “Is GRP78/BiP a Potential Salivary Biomarker in Patients with Rheumatoid Arthritis?” *Proteomics - Clinical Applications* 4 (3): 315–24. doi:10.1002/prca.200900082.

Giusti, Laura, Laura Bazzichi, Chiara Baldini, Federica Ciregia, Giovanni Mascia, Gino Giannaccini, Mario Del

Rosso, Stefano Bombardieri, and Antonio Lucacchini. 2007. “Specific Proteins Identified in Whole Saliva from Patients with Diffuse Systemic Sclerosis.” *Journal of Rheumatology* 34 (10): 2063–69. doi:07/13/0817.

Goebel, C., L. G. Mackay, E. R. Vickers, and L. E. Mather. 2000. “Determination of Defensin HNP-1, HNP-2, and HNP-3 in Human Saliva by Using LC/MS.” *Peptides* 21 (6): 757–65. doi:10.1016/S0196-9781(00)00205-9.

Gonçalves, Lorena Da Rós, Márcia Regina Soares, Fábio C S Nogueira, Carlos Garcia, Danielle Resende Camisasca, Gilberto Domont, Alfredo C R Feitosa, Denise de Abreu Pereira, Russolina B Zingali, and Gilda Alves. 2010. “Comparative Proteomic Analysis of Whole Saliva from Chronic Periodontitis Patients.” *Journal of Proteomics* 73 (7). Elsevier B.V.: 1334–41. doi:10.1016/j.jprot.2010.02.018.

Gorr, Sven-Ulrik. 2009. “Antimicrobial Peptides of the Oral Cavity.” *Periodontology 2000* 51 (57): 152–80. doi:10.1111/j.1600-0757.2009.00310.x.

Goyette, Jesse, and Carolyn L. Geczy. 2011. “Inflammation-Associated S100 Proteins: New Mechanisms That Regulate Function.” *Amino Acids*. doi:10.1007/s00726-010-0528-0.

Gu, M, G G Haraszthy, A R Collins, and E J Bergey. 1995. “Identification of Salivary Proteins Inhibiting Herpes Simplex Virus 1 Replication.” *Oral Microbiology and Immunology* 10 (1): 54–59.

Gumucio, Deborah L., A. Diaz, P. Schaner, N. Richards, C. Babcock, M. Schaller, and T. Cesena. 2002. “Fire and ICE: The Role of Pyrin Domain-Containing Proteins in Inflammation and Apoptosis.” *Clinical and Experimental Rheumatology*.

Güncan, Sabri, N Şule Y Bilge, Döndü Üsküdar Cansu, Timuçin Kaşifoğlu, and Cengiz Korkmaz. 2016. “The Role of MEFV Mutations in the Concurrent Disorders Observed in Patients with Familial Mediterranean Fever.” *European Journal of Rheumatology* 3 (3): 118–21. doi:10.5152/eurjrheum.2016.16012.

Haigh, Brendan J., Kevin W. Stewart, John R K Whelan, Matthew P G Barnett, Grant A. Smolenski, and Thomas T Wheeler. 2010. “Alterations in the Salivary Proteome Associated with Periodontitis.” *Journal of Clinical Periodontology* 37 (3): 241–47. doi:10.1111/j.1600-051X.2009.01525.x.

Han, Xuemei, Aaron Aslanian, and John Yates. 2008. “NIH Public Access.” *Curr Opin Chem Biol.* 12 (5): 483–90. doi:10.1016/j.cbpa.2008.07.024.Mass.

Hanash, S, and A Taguchi. 2010. “The Grand Challenge to Decipher the Cancer Proteome.” *Nat Rev Cancer* 10 (9). Nature Publishing Group: 652–60. doi:10.1038/nrc2918.

Hannappel, E. 2007. “Beta-Thymosins.” *Annals of the New York Academy of Sciences* 1112: 21–37. doi:10.1196/annals.1415.018.

Hannappel, E. 2010. “Thymosin beta4 and Its Posttranslational Modifications.” *Annals of the New York Academy of Sciences* 1194: 27–35. doi:10.1111/j.1749-6632.2010.05485.x.

Hardt, M, H E Witkowska, S Webb, L R Thomas, S E Dixon, S C Hall, and S J Fisher. 2005. “Assessing the Effects of Diurnal Variation on the Composition of Human Parotid Saliva: Quantitative Analysis of Native Peptides Using iTRAQ Reagents.” *Anal Chem* 77 (15): 4947–54. doi:10.1021/ac050161r.

Hatton, M N, R E Loomis, M J Levine, and L A Tabak. 1985. “Masticatory Lubrication. The Role of Carbohydrate in the Lubricating Property of a Salivary Glycoprotein-Albumin Complex.” *The Biochemical Journal* 230 (3): 817–20.

HELLER, H, E SOHAR, and M PRAS. 1961. “Ethnic Distribution and Amyloidosis in Familial Mediterranean Fever (FMF).” *Pathologia et Microbiologia* 24. Switzerland: 718–23.

HelmerhorsHMt, Eva J., Wim Van’T Hof, Pieter Breeuwer, Enno C I Veerman, Tjakko Abee, Robert F. Troxler, Arie V Nieuw Amerongen, and Frank G. Oppenheim. 2001. “Characterization of Histatin 5 with Respect to Amphipathicity, Hydrophobicity, and Effects on Cell and Mitochondrial Membrane Integrity Excludes a Candidacidal Mechanism of Pore Formation.” *Journal of Biological Chemistry* 276 (8): 5643–49. doi:10.1074/jbc.M008229200.

Herlin, T, K Storm, and B Hamborg-Petersen. 1985. “Remission of Progressive Renal Failure in Familial Mediterranean Fever during Colchicine Treatment.” *Archives of Disease in Childhood* 60 (5): 477–79. doi:10.1136/adc.60.5.477.

Hirtz, Christophe, François Chevalier, Nicolas Sommerer, and Isabelle Raingeard. 2006. “Salivary Protein Profiling in Type 1 Diabetes Using Two-Dimensional Electrophoresis and Mass Spectrometry.” *Clinical Proteomics* 0: 117–27. doi:10.1385/CP:2:1:117.

Hochstrasser, K, R Reichert, S Schwarz, and E Werle. 1972. “[Isolation and characterisation of a protease inhibitor from human bronchial secretion].” *Hoppe-Seyler’s Zeitschrift fur physiologische Chemie* 353 (2). Germany: 221–26.

Hu, S, Y Li, J Wang, Y Xie, K Tjon, L Wolinsky, R R O Loo, J A Loo, and D T Wong. 2006. “Human Saliva Proteome and Transcriptome.” *Journal of Dental Research* 85 (12): 1129–33. doi:10.1177/154405910608501212.

Hu, S, and D T Wong. 2007. “Oral Cancer Proteomics.” *Current Opinion in Molecular Therapeutics* 9 (5): 467–76. isi:000250093500006.

Huff, Thomas, Christian S G Müller, Angela M. Otto, Roland Netzker, and Ewald Hannappel. 2001. “β-Thymosins, Small Acidic Peptides with Multiple Functions.” *International Journal of Biochemistry and Cell*

Biology. doi:10.1016/S1357-2725(00)00087-X.

Humphrey, Sue P., and Russell T. Williamson. 2001. "A Review of Saliva: Normal Composition, Flow, and Function." *Journal of Prosthetic Dentistry* 85 (2): 162–69. doi:10.1067/mpr.2001.113778.

Iavarone, Federica, Tiziana Cabras, Elisabetta Pisano, Maria Teresa Sanna, Sonia Nemolato, Giovanni Vento, Chiara Tirone, et al. 2013. "Top-down HPLC-ESI-MS Detection of S-Glutathionylated and S-Cysteinylated Derivatives of Cystatin B and Its 1-53 and 54-98 Fragments in Whole Saliva of Human Preterm Newborns." *Journal of Proteome Research* 12 (2): 917–26. doi:10.1021/pr300960f.

Imamura, Yasuhiro, and Pao-Li Wang. 2014. "Salivary Histatin 3 Inhibits Heat Shock Cognate Protein 70-Mediated Inflammatory Cytokine Production through Toll-like Receptors in Human Gingival Fibroblasts." *Journal of Inflammation (London, England)* 11 (1): 4. doi:10.1186/1476-9255-11-4.

Imanguli, Matin M., Jane C. Atkinson, Kristen E. Harvey, Gerard T. Hoehn, Ok Hee Ryu, Tianxia Wu, Albert Kingman, et al. 2007. "Changes in Salivary Proteome Following Allogeneic Hematopoietic Stem Cell Transplantation." *Experimental Hematology* 35 (2): 184–92. doi:10.1016/j.exphem.2006.10.009.

Imatani, T, T Kato, K Minaguchi, and K Okuda. 2000. "Histatin 5 Inhibits Inflammatory Cytokine Induction from Human Gingival Fibroblasts by Porphyromonas Gingivalis." *Oral Microbiology and Immunology* 15 (6): 378–82. doi:10.1034/j.1399-302X.2000.150607.x.

Inzitari, Rosanna, Tiziana Cabras, Giuseppina Onnis, Chiara Olmi, Andrea Mastinu, Maria Teresa Sanna, Maria Guiseppina Pellegrini, Massimo Castagnola, and Irene Messana. 2005. "Different Isoforms and Post-Translational Modifications of Human Salivary Acidic Proline-Rich Proteins." *Proteomics* 5 (3): 805–15. doi:10.1002/pmic.200401156.

Inzitari, Rosanna, Tiziana Cabras, Elisabetta Pisano, Chiara Fanali, Barbara Manconi, Emanuele Scarano, Antonella Fiorita, et al. 2009. "HPLC-ESI-MS Analysis of Oral Human Fluids Reveals That Gingival Crevicular Fluid Is the Main Source of Oral Thymosins β4 and β10." *Journal of Separation Science* 32 (1): 57–63. doi:10.1002/jssc.200800496.

Inzitari, Rosanna, Tiziana Cabras, Diana Valeria Rossetti, Chiara Fanali, Alberto Vitali, Mariagiuseppina Pellegrini, Gaetano Paludetti, et al. 2006. "Detection in Human Saliva of Different Statherin and P-B Fragments and Derivatives." *Proteomics* 6 (23): 6370–79. doi:10.1002/pmic.200600395.

Jarai, Tamas, Gabor Maasz, Andras Burian, Agnes Bona, Eva Jambor, Imre Gerlinger, and Laszlo Mark. 2012. "Mass Spectrometry-Based Salivary Proteomics for the Discovery of Head and Neck Squamous Cell Carcinoma." *Pathology Oncology Research : POR* 18 (3): 623–28. doi:10.1007/s12253-011-9486-4.

Jin, Fen Yu, Carl Nathan, Danuta Radzioch, and Aihao Ding. 1997. "Secretory Leukocyte Protease Inhibitor: A

Macrophage Product Induced by and Antagonistic to Bacterial Lipopolysaccharide." *Cell* 88 (3): 417–26. doi:10.1016/S0092-8674(00)81880-2.

Jou, Yu Jen, Chia Der Lin, Chih Ho Lai, Chao Hsien Chen, Jung Yie Kao, Shih Yin Chen, Ming Hsui Tsai, Su Hua Huang, and Cheng Wen Lin. 2010. "Proteomic Identification of Salivary Transferrin as a Biomarker for Early Detection of Oral Cancer." *Analytica Chimica Acta* 681 (1–2). Elsevier B.V.: 41–48. doi:10.1016/j.aca.2010.09.030.

Jou, Yu Jen, Chia Der Lin, Chih Ho Lai, Chih Hsin Tang, Su Hua Huang, Ming Hsui Tsai, Shih Yin Chen, Jung Yie Kao, and Cheng Wen Lin. 2011. "Salivary Zinc Finger Protein 510 Peptide as a Novel Biomarker for Detection of Oral Squamous Cell Carcinoma in Early Stages." *Clinica Chimica Acta* 412 (15–16). Elsevier B.V.: 1357–65. doi:10.1016/j.cca.2011.04.004.

Kaplan, Eytan, Masza Mukamel, Judith Barash, Riva Brik, Shai Padeh, Yaakov Berkun, Yosef Uziel, Tsivia Tauber, Jacob Amir, and Liora Harel. 2007. "Protracted Febrile Myalgia in Children and Young Adults with Familial Mediterranean Fever: Analysis of 15 Patients and Suggested Criteria for Working Diagnosis." *Clinical and Experimental Rheumatology* 25 (4 SUPPL. 45).

Karve, Tejaswita M., and Amrita K. Cheema. 2011. "Small Changes Huge Impact: The Role of Protein Posttranslational Modifications in Cellular Homeostasis and Disease." *Journal of Amino Acids* 2011: 1–13. doi:10.4061/2011/207691.

Kastner, Daniel L., Ivona Aksentijevich, and Raphaela Goldbach-Mansky. 2010. "Autoinflammatory Disease Reloaded: A Clinical Perspective." *Cell*. doi:10.1016/j.cell.2010.03.002.

Kato, T., T. Imatani, T. Miura, K. Minaguchi, E. Saitoh, and K. Okuda. 2000. "Cytokine-Inducing Activity of Family 2 Cystatins." *Biological Chemistry* 381 (11): 1143–47. doi:10.1515/BC.2000.141.

Kellner, Roland. 2000. "Proteomics. Concepts and Perspectives." *Fresenius' Journal of Analytical Chemistry* 366 (6–7): 517–24. doi:10.1007/s002160051547.

Kikuchi, Toshiaki, Tatsuya Abe, Sachiko Hoshi, Nobumichi Matsubara, Yasuyuki Tominaga, Ken Satoh, and Toshihiro Nukiwa. 1998. "Structure of the Murine Secretory Leukoprotease Inhibitor (Slpi) Gene and Chromosomal Localization of the Human and Murine SLPI Genes." *American Journal of Respiratory Cell and Molecular Biology* 19 (6): 875–80.

Kilic, Ayse, Muhammet Ali Varkal, Mehmet Sait Durmus, Ismail Yildiz, Zeynep Nagihan Yürük Yıldırım, Gorkem Turunc, Fatma Oguz, et al. 2015. "Relationship between Clinical Findings and Genetic Mutations in Patients with Familial Mediterranean Fever." *Pediatric Rheumatology Online Journal* 13: 59. doi:10.1186/s12969-015-0057-1.

Klein, Jerome J., Allan L. Goldstein, and Abraham White. 1965. “Enhancement of in Vivo Incorporation of Labeled Precursors into DNA and Total Protein of Mouse Lymph Nodes after Administration of Thymic Extracts.” *Proceedings of the National Academy of Sciences of the United States of America* 53 (4): 812–17.

Koizumi, Masahiro, Aiko Fujino, Kay Fukushima, Takashi Kamimura, and Midori Takimoto-Kamimura. 2008. “Complex of Human Neutrophil Elastase with 1/2SLPI.” *Journal of Synchrotron Radiation* 15 (3): 308–11. doi:10.1107/S0909049507060670.

Kolivras, Athanassios, Philippe Provost, and Curtis T. Thompson. 2013. “Erysipelas-like Erythema of Familial Mediterranean Fever Syndrome: A Case Report with Emphasis on Histopathologic Diagnostic Clues.” *Journal of Cutaneous Pathology* 40 (6): 585–90. doi:10.1111/cup.12132.

Koné-Paut, Isabelle, Véronique Hentgen, and Isabelle Touitou. 2011. “Current Data on Familial Mediterranean Fever.” *Joint Bone Spine Revue Du Rhumatisme*. doi:10.1016/j.jbspin.2010.09.021.

Kopitar-Jerala, Nataša. 2006. “The Role of Cystatins in Cells of the Immune System.” *FEBS Letters*. doi:10.1016/j.febslet.2006.10.055.

Kopitar-Jerala, N. 2015. “The Role of Stefin B in Neuro-Inflammation.” *Frontiers in Cellular Neuroscience* 9: 458. doi:10.3389/fncel.2015.00458.

Krueger, Karl E, and Sudhir Srivastava. 2006. “Posttranslational Protein Modifications: Current Implications for Cancer Detection, Prevention, and Therapeutics.” *Molecular & Cellular Proteomics* 5 (10): 1799–1810. doi:10.1074/mcp.R600009-MCP200.

Kubota, Kazuo, Hidenori Ohnishi, Takahide Teramoto, Norio Kawamoto, Kimiko Kasahara, Osamu Ohara, and Naomi Kondo. 2014. “Clinical and Genetic Characterization of Japanese Sporadic Cases of Periodic Fever, Aphthous Stomatitis, Pharyngitis and Adenitis Syndrome from a Single Medical Center in Japan.” *Journal of Clinical Immunology* 34 (5): 584–93. doi:10.1007/s10875-014-0043-2.

Lane, Thirusha, and Helen J Lachmann. 2011. “The Emerging Role of Interleukin-1beta in Autoinflammatory Diseases.” *Current Allergy and Asthma Reports* 11 (5). United States: 361–68. doi:10.1007/s11882-011-0207-6.

Lane, Thirusha, Jutta M. Loeffler, Dorota M. Rowczenio, Janet A. Gilbertson, Alison Bybee, Tonia L. Russell, Julian D. Gillmore, Ashutosh D. Wechalekar, Philip N. Hawkins, and Helen J. Lachmann. 2013. “AA Amyloidosis Complicating the Hereditary Periodic Fever Syndromes.” *Arthritis and Rheumatism* 65 (4): 1116–21. doi:10.1002/art.37827.

Legrand-Poels, Sylvie, Nathalie Esser, Laurent L’homme, Andre Scheen, Nicolas Paquot, and Jacques Piette. 2014. “Free Fatty Acids as Modulators of the NLRP3 Inflammasome in Obesity/type 2 Diabetes.” *Biochemical Pharmacology* 92 (1). England: 131–41. doi:10.1016/j.bcp.2014.08.013.

Lehrer, Robert I., and Wuyuan Lu. 2012. "??-Defensins in Human Innate Immunity." *Immunological Reviews*. doi:10.1111/j.1600-065X.2011.01082.x.

Lenarcic, B., M. Krasovec, A. Ritonja, I. Olafsson, and V. Turk. 1991. "Inactivation of Human Cystatin C and Kininogen by Human Cathepsin D." *FEBS Letters* 280 (2): 211–15. doi:10.1016/0014-5793(91)80295-E.

Li, J, E J Helmerhorst, R F Troxler, and F G Oppenheim. 2004. "Identification of in Vivo Pellicle Constituents by Analysis of Serum Immune Responses." *Journal of Dental Research*. doi:10.1177/154405910408300112.

Li, Xueshu, Sean Parkin, Michael W. Duffel, Larry W. Robertson, and Hans Joachim Lehmler. 2010. "An Efficient Approach to Sulfate Metabolites of Polychlorinated Biphenyls." *Environment International* 36 (8). Elsevier Ltd: 843–48. doi:10.1016/j.envint.2009.02.005.

Lie, M a, B G Loos, Y M Henskens, M F Timmerman, E C Veerman, U van der Velden, and G a van der Weijden. 2001. "Salivary Cystatin Activity and Cystatin C in Natural and Experimental Gingivitis in Smokers and Non-Smokers." *Journal of Clinical Periodontology* 28 (10): 979–84. doi:281012.

Lim, Su Yin, Mark J Raftery, Jesse Goyette, Kenneth Hsu, and Carolyn L Geczy. 2009. "Oxidative Modifications of S100 Proteins: Functional Regulation by Redox." *Journal of Leukocyte Biology* 86 (3): 577–87. doi:10.1189/jlb.1008608.

Liotta, Lance A., and Emanuel Petricoin. 2011. "Cancer Biomarkers: Closer to Delivering on Their Promise." *Cancer Cell* 20 (3). Elsevier Inc.: 279–80. doi:10.1016/j.ccr.2011.08.021.

Lyons, K. M., J. H. Stein, and O. Smithies. 1988. "Length Polymorphisms in Human Proline-Rich Protein Genes Generated by Intragenic Unequal Crossing over." *Genetics* 120 (1): 267–78.

Maher, Katarina, Barbara Jerič Kokelj, Miha Butinar, Georgy Mikhaylov, Mateja Manček-Keber, Veronika Stoka, Olga Vasiljeva, Boris Turk, Sergei A Grigoryev, and Natasa Kopitar-Jerala. 2014. "A Role for Stefin B (Cystatin B) in Inflammation and Endotoxemia." *Journal of Biological Chemistry* 289 (46): 31736–50. doi:10.1074/jbc.M114.609396.

Majchrzak-Gorecka, Monika, Pawel Majewski, Beata Grygier, Krzysztof Murzyn, and Joanna Cichy. 2016. "Secretory Leukocyte Protease Inhibitor (SLPI), a Multifunctional Protein in the Host Defense Response." *Cytokine and Growth Factor Reviews*. doi:10.1016/j.cytogfr.2015.12.001.

Mamta, Singh, Singhal Udita, Bhasin Gk, Panday Rajesh, and Aggarwal Sk. 2013. "Oral Fluid: Biochemical Composition and Functions: A Review." *Journal of Pharmaceutical and Biomedical Sciences* 27 (14): 508–14.

Manconi, Barbara, Tiziana Cabras, Monica Sanna, Valentina Piras, Barbara Liori, Elisabetta Pisano, Federica

Iavarone, et al. 2016. “N- and O-Linked Glycosylation Site Profiling of the Human Basic Salivary Proline-Rich Protein 3M.” *Journal of Separation Science*. doi:10.1002/jssc.201501306.

Manconi, Barbara, Massimo Castagnola, Tiziana Cabras, Alessandra Olianas, Alberto Vitali, Claudia Desiderio, Maria Teresa Sanna, and Irene Messana. 2016. “The Intriguing Heterogeneity of Human Salivary Proline-Rich Proteins: Short Title: Salivary Proline-Rich Protein Species.” *Journal of Proteomics* 134. Elsevier B.V.: 47–56. doi:10.1016/j.jprot.2015.09.009.

Mankan, A K, A Kubarenko, and V Hornung. 2012. “Immunology in Clinic Review Series; Focus on Autoinflammatory Diseases: Inflammasomes: Mechanisms of Activation.” *Clinical and Experimental Immunology* 167 (3). England: 369–81. doi:10.1111/j.1365-2249.2011.04534.x.

Mann, Matthias, and Ole N Jensen. 2003. “Proteomic Analysis of Post-Translational Modifications.” *Nat Biotech* 21 (3). Nature Publishing Group: 255–61. http://dx.doi.org/10.1038/nbt0303-255.

Mansfield, Elizabeth, Jae Jin Chae, Hirsh D. Komarow, Tilmann M. Brotz, David M. Frucht, Ivona Aksentijevich, and Daniel L. Kastner. 2001. “The Familial Mediterranean Fever Protein, Pyrin, Associates with Microtubules and Colocalizes with Actin Filaments.” *Blood* 98 (3): 851–59. doi:10.1182/blood.V98.3.851.

Marek-Yagel, Dina, Ifat Bar-Joseph, Elon Pras, and Yackov Berkun. 2009. “Is E148Q a Benign Polymorphism or a Disease-Causing Mutation?” *The Journal of Rheumatology*. Canada. doi:10.3899/jrheum.090250.

Marek-Yagel, Dina, Yackov Berkun, Shai Padeh, Merav Lidar, Yael Shinar, Ifat Bar-Joseph, Haike Reznik-Wolf, Pnina Langevitz, Avi Livneh, and Elon Pras. 2010. “Role of the R92Q TNFRSF1A Mutation in Patients with Familial Mediterranean Fever.” *Arthritis Care & Research* 62 (9). United States: 1294–98. doi:10.1002/acr.20213.

Marenholz, Ingo, Ruth C. Lovering, and Claus W. Heizmann. 2006. “An Update of the S100 Nomenclature.” *Biochimica et Biophysica Acta - Molecular Cell Research*. doi:10.1016/j.bbamcr.2006.07.013.

McCormick, Michelle M., Farid Rahimi, Yuri V. Bobryshev, Katharina Gaus, Hala Zreiqat, Hong Cai, Reginald S A Lord, and Carolyn L. Geczy. 2005. “S100A8 and S100A9 in Human Arterial Wall: Implications for Atherogenesis.” *Journal of Biological Chemistry* 280 (50): 41521–29. doi:10.1074/jbc.M509442200.

McKiernan, Paul J, Noel G McElvaney, and Catherine M Greene. 2011. “SLPI and Inflammatory Lung Disease in Females.” *Biochemical Society Transactions* 39 (5): 1421–26. doi:10.1042/BST0391421.

Medlej-Hashim, Myrna, Valérie Delague, Eliane Chouery, Nabiha Salem, Mohammed Rawashdeh, Gérard Lefranc, Jacques Loiselet, and André Mégarbané. 2004. “Amyloidosis in Familial Mediterranean Fever Patients: Correlation with MEFV Genotype and SAA1 and MICA Polymorphisms Effects.” *BMC Medical Genetics* 5: 4. doi:10.1186/1471-2350-5-4.

Melis, Melania, Massimiliano Arca, Maria Carla Aragoni, Tiziana Cabras, Claudia Caltagirone, Massimo Castagnola, Roberto Crnjar, M. Messna, Beverly J. Tepper, and Iole Tomassini Barbarossa. 2015. "Dose-Dependent Effects of L-Arginine on PROP Bitterness Intensity and Latency and Characteristics of the Chemical Interaction between PROP and L-Arginine." *PLoS ONE* 10 (6). doi:10.1371/journal.pone.0131104.

Messana, Irene, Tiziana Cabras, Federica Iavarone, Barbara Manconi, Liling Huang, Claudia Martelli, Alessandra Olianas, et al. 2015. "Chrono-Proteomics of Human Saliva: Variations of the Salivary Proteome during Human Development." *Journal of Proteome Research* 14 (4): 1666–77. doi:10.1021/pr501270x.

Messana, Irene, Tiziana Cabras, Federica Iavarone, Federica Vincenzoni, Andrea Urbani, and Massimo Castagnola. 2013. "Unraveling the Different Proteomic Platforms." *Journal of Separation Science*. doi:10.1002/jssc.201200830.

Messana, Irene, Tiziana Cabras, Rosanna Inzitari, Alessandro Lupi, Cecilia Zuppi, Chiara Olmi, Maria Benedetta Fadda, Massimo Cordaro, Bruno Giardina, and Massimo Castagnola. 2004. "Characterization of the Human Salivary Basic Proline-Rich Protein Complex by a Proteomic Approach." *Journal of Proteome Research* 3 (4): 792–800. doi:10.1021/pr049953c.

Messana, Irene, Tiziana Cabras, Elisabetta Pisano, Maria Teresa Sanna, Alessandra Olianas, Barbara Manconi, Mariagiuseppina Pellegrini, et al. 2008. "Trafficking and Postsecretory Events Responsible for the Formation of Secreted Human Salivary Peptides: A Proteomics Approach." *Molecular & Cellular Proteomics : MCP* 7 (5): 911–26. doi:10.1074/mcp.M700501-MCP200.

Messana, Irene, Rosanna Inzitari, Chiara Fanali, Tiziana Cabras, and Massimo Castagnola. 2008. "Facts and Artifacts in Proteomics of Body Fluids. What Proteomics of Saliva Is Telling Us?" *Journal of Separation Science*. doi:10.1002/jssc.200800100.

Meyer, Kaspar. 2011. "Primary Sensory Cortices, Top-down Projections and Conscious Experience." *Progress in Neurobiology* 94 (4). Elsevier Ltd: 408–17. doi:10.1016/j.pneurobio.2011.05.010.

Mimouni, a, N Magal, N Stoffman, T Shohat, a Minasian, M Krasnov, G J Halpern, et al. 2000. "Familial Mediterranean Fever: Effects of Genotype and Ethnicity on Inflammatory Attacks and Amyloidosis." *Pediatrics* 105 (5): E70. doi:10.1542/peds.105.5.e70.

Mischak, Harald, Günter Allmaier, Rolf Apweiler, Teresa Attwood, Marc Baumann, Ariela Benigni, Samuel E Bennett, et al. 2010. "Recommendations for Biomarker Identification and Qualification in Clinical Proteomics." *Science Translational Medicine* 2 (46): 46ps42 LP-46ps42.

Mizuno, Takahisa, Hidemasa Sakai, Ryuta Nishikomori, Koichi Oshima, Osamu Ohara, Ikue Hata, Yosuke Shigematsu, Takashi Ishige, Kazushi Tamura, and Hirokazu Arakawa. 2012. "Novel Mutations of MVK Gene in

Japanese Family Members Affected with Hyperimmunoglobulinemia D and Periodic Fever Syndrome." *Rheumatology International* 32 (12): 3761–64. doi:10.1007/s00296-011-2225-z.

Moore, B W. 1965. "A Soluble Protein Characteristic of the Nervous System." *Biochemical and Biophysical Research Communications* 19 (6): 739–44. doi:10.1016/0006-291X(65)90320-7.

Moreno, E. C., M. Kresak, and D. I. Hay. 1982. "Adsorption Thermodynamics of Acidic Proline-Rich Human Salivary Proteins onto Calcium Apatites." *Journal of Biological Chemistry* 257 (6): 2981–89.

Morzel, Martine, Aline Jeannin, Géraldine Lucchi, Caroline Truntzer, Delphine Pecqueur, Sophie Nicklaus, Christophe Chambon, and Patrick Ducoroy. 2012. "Human Infant Saliva Peptidome Is Modified with Age and Diet Transition." *Journal of Proteomics* 75 (12). Elsevier B.V.: 3665–73. doi:10.1016/j.jprot.2012.04.028.

MP, Washburn, D Wolters, and Yates J R 3rd. 2001. "Large-Scale Analysis of the Yeast Proteome by Multidimensional Protein Identification Technology." *Nature Biotechnology* 19 (3): 242–47.

Naimushin, Alexey, Mirav Lidar, Ilan Ben Zvi, and Avi Livneh. 2011. "The Structural Effect of the E148Q MEFV Mutation on the Pyrin Protein: A Study Using a Quantum Chemistry Model." *Israel Medical Association Journal* 13 (4): 199–201.

Nakamura, Yu, Masatoshi Takeda, Hideo Suzuki, Hideyuki Hattori, Kunitoshi Tada, Shiro Hariguchi, Shigeo Hashimoto, and Tsuyoshi Nishimura. 1991. "Abnormal Distribution of Cathepsins in the Brain of Patients with Alzheimer's Disease." *Neuroscience Letters* 130 (2): 195–98. doi:10.1016/0304-3940(91)90395-A.

Ngo, Luan H, Paul D Veith, Yu-Yen Chen, Dina Chen, Ivan B Darby, and Eric C Reynolds. 2009. "Mass Spectrometric Analyses of Peptides and Proteins in Human Gingival Crevicular Fluid." *J Proteome Res* 9 (4): 1683–93. doi:10.1021/pr900775s.

Nikolov, Miroslav, Carla Schmidt, and Henning Urlaub. 2012. "Quantitative Mass Spectrometry-Based Proteomics: An Overview." In *Quantitative Methods in Proteomics*, edited by Katrin Marcus, 85–100. Totowa, NJ: Humana Press. doi:10.1007/978-1-61779-885-6_7.

Nilsson, Tommy, Matthias Mann, Ruedi Aebersold, John R Yates, Amos Bairoch, and John J M Bergeron. 2010. "Mass Spectrometry in High-Throughput Proteomics: Ready for the Big Time." *Nat Meth* 7 (9). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 681–85. http://dx.doi.org/10.1038/nmeth0910-681.

Nystrom, M, M Bergenfeldt, I Ljungcrantz, A Lindeheim, and K Ohlsson. 1999. "Production of Secretory Leucocyte Protease Inhibitor (SLPI) in Human Pancreatic Beta-Cells." *Mediators of Inflammation* 8 (3). United States: 147–51. doi:10.1080/09629359990478.

Ohlsson, K, and H Tegner. 1976. “Inhibition of Elastase from Granulocytes by the Low Molecular Weight Bronchial Protease Inhibitor.” *Scandinavian Journal of Clinical and Laboratory Investigation* 36 (5). England: 437–45.

Onen, Fatos. 2006. “Familial Mediterranean Fever.” *Rheumatology International* 26 (6). Germany: 489–96. doi:10.1007/s00296-005-0074-3.

Oppenheim, F. G., T. Xu, F. M. McMillian, S. M. Levitz, R. D. Diamond, G. D. Offner, and R. F. Troxler. 1988. “Histatins, a Novel Family of Histidine-Rich Proteins in Human Parotid Secretion. Isolation, Characterization, Primary Structure, and Fungistatic Effects on Candida Albicans.” *Journal of Biological Chemistry* 263 (16): 7472–77.

Oudhoff, Menno J, Jan G M Bolscher, Kamran Nazmi, Hakan Kalay, Wim van ’t Hof, Arie V Nieuw Amerongen, and Enno C I Veerman. 2008. “Histatins Are the Major Wound-Closure Stimulating Factors in Human Saliva as Identified in a Cell Culture Assay.” *The Federation of American Societies for Experimental Biology* 22 (11): 3805–12. doi:10.1096/fj.08-112003.

Ozen, S., D. Uckan, E. Baskin, N. Besbas, H. Okur, U. Saatci, and A. Bakkaloglu. 2001. “Increased Neutrophil Apoptosis during Attacks of Familial Mediterranean Fever.” *Clinical and Experimental Rheumatology* 19 (5 SUPPL. 24).

Padeh, Shai, Avi Livneh, Elon Pras, Yael Shinar, Merav Lidar, Olga Feld, and Yackov Berkun. 2010. “Familial Mediterranean Fever in the First Two Years of Life: A Unique Phenotype of Disease in Evolution.” *The Journal of Pediatrics* 156 (6). United States: 985–89. doi:10.1016/j.jpeds.2009.12.010.

Papin, S, S Cuenin, L Agostini, F Martinon, S Werner, H D Beer, C Grutter, M Grutter, and J Tschopp. 2007. “The SPRY Domain of Pyrin, Mutated in Familial Mediterranean Fever Patients, Interacts with Inflammasome Components and Inhibits proIL-1beta Processing.” *Cell Death Differ* 14 (8): 1457–66. doi:10.1038/sj.cdd.4402142.

Patterson, Scott D, and Ruedi H Aebersold. 2003. “Proteomics: The First Decade and Beyond.” *Nature Genetics* 33 Suppl (march): 311–23. doi:10.1038/ng1106.

Peluso, G, M. De Santis, R Inzitari, C Fanali, T Cabras, I Messana, M Castagnola, and G F Ferraccioli. 2007. “Proteomic Study of Salivary Peptides and Proteins in Patients with Sjögren’s Syndrome before and after Pilocarpine Treatment.” *Arthritis and Rheumatism* 56 (7): 2216–22. doi:10.1002/art.22738.

Pisano, Elisabetta, Tiziana Cabras, Caterina Montaldo, Vincenzo Piras, Rosanna Inzitari, Chiara Olmi, Massimo Castagnola, and Irene Messana. 2005. “Peptides of Human Gingival Crevicular Fluid Determined by HPLC-ESI-MS.” *European Journal of Oral Sciences* 113 (6): 462–68. doi:10.1111/j.1600-0722.2005.00246.x.

Poste, George. 2011. “Bring on the Biomarkers.” *Nature* 469 (7329). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 156–57. http://dx.doi.org/10.1038/469156a.

Pras, M, E Pras, and D Kastner. 1995. “The Origin of the FMF Gene.” *Israel Journal of Medical Sciences*. Israel.

Raj, Periathamby Antony, Mira Edgerton, and Michael J. Levine. 1990. “Salivary Histatin 5: Dependence of Sequence, Chain Length, and Helical Conformation for Candidacidal Activity.” *Journal of Biological Chemistry* 265 (7): 3898–3905.

Ravasi, Timothy, Kenneth Hsu, Jesse Goyette, Kate Schroder, Zheng Yang, Farid Rahimi, Les P. Miranda, Paul F. Alewood, David A. Hume, and Carolyn Geczy. 2004. “Probing the S100 Protein Family through Genomic and Functional Analysis.” *Genomics* 84 (1): 10–22. doi:10.1016/j.ygeno.2004.02.002.

Reutens, Anne T., Fabrice Bonnet, Olivier Lantieri, Ronan Roussel, and Beverley Balkau. 2013. “The Association between Cystatin C and Incident Type 2 Diabetes Is Related to Central Adiposity.” *Nephrology Dialysis Transplantation* 28 (7): 1820–29. doi:10.1093/ndt/gfs561.

Reviglio, Victor E., Ruben H. Sambuelli, Alejandra Olmedo, Micaela Falco, Jose Echenique, Terrence P. O’Brien, and Irene C. Kuo. 2007. “Secretory Leukocyte Protease Inhibitor Is an Inducible Antimicrobial Peptide Expressed in Staphylococcus Aureus Endophthalmitis.” *Mediators of Inflammation* 2007. doi:10.1155/2007/93857.

Rudney, J. D., H. Xie, N. L. Rhodus, F. G. Ondrey, and T. J. Griffin. 2010. “A Metaproteomic Analysis of the Human Salivary Microbiota by Three-Dimensional Peptide Fractionation and Tandem Mass Spectrometry.” *Molecular Oral Microbiology* 25 (1): 38–49. doi:10.1111/j.2041-1014.2009.00558.x.

Ruhl, Stefan. 2012. “The Scientific Exploration of Saliva in the Post-Proteomic Era: From Database back to Basic Function.” *Expert Review of Proteomics* 9 (1). Taylor & Francis: 85–96. doi:10.1586/epr.11.80.

Ruzindana-Umunyana, Angelique, and Joseph M. Weber. 2001. “Interactions of Human Lacrimal and Salivary Cystatins with Adenovirus Endopeptidase.” *Antiviral Research* 51 (3): 203–14. doi:10.1016/S0166-3542(01)00154-1.

Ryckman, Carle, Karen Vandal, Pascal Rouleau, Marième Talbot, and Philippe A Tessier. 2003. “Proinflammatory Activities of S100: Proteins S100A8, S100A9, and S100A8/A9 Induce Neutrophil Chemotaxis and Adhesion.” *Journal of Immunology (Baltimore, Md. : 1950)* 170 (6): 3233–42. doi:10.4049/jimmunol.170.6.3233.

Sabatini, L. M., and E. A. Azen. 1989. “Histatins, a Family of Salivary Histidine-Rich Proteins, Are Encoded by at Least Two Loci (HIS1 and HIS2).” *Biochemical and Biophysical Research Communications* 160 (2): 495–502. doi:10.1016/0006-291X(89)92460-1.

Sabatini, L M, L R Carlock, G W Johnson, and E A Azen. 1987. “cDNA Cloning and Chromosomal Localization (4q11-13) of a Gene for Statherin, a Regulator of Calcium in Saliva.” *Am J Hum Genet* 41 (6): 1048–60.

Saitoh, Eiichi, Hyung Suk Kim, Oliver Smithies, and Nobuyo Maeda. 1987. “Human Cysteine-Proteinase Inhibitors: Nucleotide Sequence Analysis of Three Members of the Cystatin Gene Family.” *Gene* 61 (3): 329–38. doi:10.1016/0378-1119(87)90196-X.

Salehzadeh, Farhad, Maryam Vahedi, Saeid Hosseini-Asl, Sepideh Jahangiri, Shahram Habibzadeh, and Mahsa Hosseini-Khotbesara. 2014. “PFAPA and 12 Common MEFV Gene Mutations Our Clinical Experience.” *Iranian Journal of Pediatrics* 24 (1): 64–68.

Samuels, Jonathan, and Seza Ozen. 2006. “Familial Mediterranean Fever and the Other Autoinflammatory Syndromes: Evaluation of the Patient with Recurrent Fever.” *Current Opinion in Rheumatology* 18 (1): 108–17. doi:10.1097/01.bor.0000198006.65697.5b.

Santamaria-Kisiel, Liliana, Anne C Rintala-Dempsey, and Gary S Shaw. 2006. “Calcium-Dependent and -Independent Interactions of the S100 Protein Family.” *The Biochemical Journal* 396 (2): 201–14. doi:10.1042/BJ20060195.

Savic, Sinisa, Laura J. Dickie, Michele Battellino, and Michael F. McDermott. 2012. “Familial Mediterranean Fever and Related Periodic Fever Syndromes/autoinflammatory Diseases.” *Current Opinion in Rheumatology* 24 (1): 103–12. doi:10.1097/BOR.0b013e32834dd2d5.

Schäfer, Beat W., Roland Wicki, Dieter Engelkamp, Marie geneviève Mattei, and Claus W. Heizmann. 1995. “Isolation of a YAC Clone Covering a Cluster of Nine S100 Genes on Human Chromosome 1q21: Rationale for a New Nomenclature of the S100 Calcium-Binding Protein Family.” *Genomics* 25 (3): 638–43. doi:10.1016/0888-7543(95)80005-7.

Schlesinger, D. H., and D. I. Hay. 1977. “Complete Covalent Structure of Statherin, a Tyrosine Rich Acidic Peptide Which Inhibits Calcium Phosphate Precipitation from Human Parotid Saliva.” *Journal of Biological Chemistry* 252 (5): 1689–95.

Schulz, Benjamin L, Justin Cooper-White, and Chamindie K Punyadeera. 2013. “Saliva Proteome Research: Current Status and Future Outlook.” *Critical Reviews in Biotechnology* 33 (3). Taylor & Francis: 246–59. doi:10.3109/07388551.2012.687361.

Schwartz, Steven S., Donald I. Hay, and Susan K. Schluckebier. 1992. “Inhibition of Calcium Phosphate Precipitation by Human Salivary Statherin: Structure-Activity Relationships.” *Calcified Tissue International* 50 (6): 511–17. doi:10.1007/BF00582164.

Seshadri, Sudarshan, Michelle D Duncan, Judith M Hart, Mikhail a Gavrilin, and Mark D Wewers. 2007. “Pyrin Levels in Human Monocytes and Monocyte-Derived Macrophages Regulate IL-1beta Processing and Release.” *Journal of Immunology (Baltimore, Md. : 1950)* 179: 1274–81. doi:179/2/1274.

Shang, Xuan, Hanhua Cheng, and Rongjia Zhou. 2008. “Chromosomal Mapping, Differential Origin and Evolution of the S100 Gene Family.” *Genetics, Selection, Evolution : GSE* 40 (4): 449–64. doi:10.1051/gse:2008013.

Shintani, S, H Hamakawa, Y Ueyama, M Hatori, and T. Toyoshima. 2010. “Identification of a Truncated Cystatin SA-I as a Saliva Biomarker for Oral Squamous Cell Carcinoma Using the SELDI ProteinChip Platform.” *International Journal of Oral and Maxillofacial Surgery* 39 (1): 68–74. doi:10.1016/j.ijom.2009.10.001.

Shlipak, Michael G, Mark J Sarnak, Ronit Katz, Linda F Fried, Stephen L Seliger, Anne B Newman, David S Siscovick, and Catherine Stehman-Breen. 2005. “Cystatin C and the Risk of Death and Cardiovascular Events among Elderly Persons.” *The New England Journal of Medicine* 352 (20): 2049–60. doi:10.1056/NEJMoa043161.

Singh, Ava, Andrew Bateman, Qinzhang Zhu, Shunichi Shimasaki, Fred Esch, and Samuel Solomon. 1988. “Structure of a Novel Human Granulocyte Peptide with Anti-ACTH Activity.” *Biochemical and Biophysical Research Communications* 155 (1): 524–29. doi:10.1016/S0006-291X(88)81118-5.

Siqueira, W L, E Salih, D L Wan, E J Helmerhorst, and F G Oppenheim. 2008. “Proteome of Human Minor Salivary Gland Secretion.” *Journal of Dental Research* 87 (5): 445–50.

Sohar, E, J Gafni, M Pras, and H Heller. 1967. “Familial Mediterranean Fever. A Survey of 470 Cases and Review of the Literature.” *The American Journal of Medicine* 43 (2). United States: 227–53.

Sohar, Ezra, Joseph Gafni, Mordehai Pras, and Harry Heller. 2016. “Familial Mediterranean Fever.” *The American Journal of Medicine* 43 (2). Elsevier: 227–53. doi:10.1016/0002-9343(67)90167-2.

Song, X y, L Zeng, W Jin, J Thompson, D E Mizel, K Lei, R C Billinghurst, A R Poole, and S M Wahl. 1999. “Secretory Leukocyte Protease Inhibitor Suppresses the Inflammation and Joint Damage of Bacterial Cell Wall-Induced Arthritis.” *The Journal of Experimental Medicine* 190 (4): 535–42. doi:10.1084/jem.190.4.535.

Stetler, Gary, Michael T. Brewer, and Robert C. Thompson. 1986. “Isolation and Sequence of a Human Gene Encoding a Potent Inhibitor of Leukocyte Proteases.” *Nucleic Acids Research* 14 (20): 7883–96. doi:10.1093/nar/14.20.7883.

Streckfus, Charles F, Lenora R Bigler, and Michael Zwick. 2006. “The Use of Surface-Enhanced Laser

Desorption/ionization Time-of-Flight Mass Spectrometry to Detect Putative Breast Cancer Markers in Saliva: A Feasibility Study." *Journal of Oral Pathology & Medicine* 35 (5). Blackwell Publishing Ltd: 292–300. doi:10.1111/j.1600-0714.2006.00427.x.

Strupat, Kerstin, Hélène Rogniaux, Alain Van Dorsselaer, Johannes Roth, and Thomas Vogl. 2000. "Calcium-Induced Noncovalently Linked Tetramers of MRP8 and MRP14 Are Confirmed by Electrospray Ionization-Mass Analysis." *Journal of the American Society for Mass Spectrometry* 11 (9): 780–88. doi:10.1016/S1044-0305(00)00150-1.

Suarez-Carmona, Meggy, Pascale Hubert, Philippe Delvenne, and Michael Herfs. 2015. "Defensins: 'Simple' antimicrobial Peptides or Broad-Spectrum Molecules?" *Cytokine and Growth Factor Reviews*. doi:10.1016/j.cytogfr.2014.12.005.

Tan, Meng-Shan, Jin-Tai Yu, Teng Jiang, Xi-Chen Zhu, and Lan Tan. 2013. "The NLRP3 Inflammasome in Alzheimer's Disease." *Molecular Neurobiology* 48 (3): 875–82. doi:10.1007/s12035-013-8475-x.

Taskiran, Ekim Z., Arda Cetinkaya, Banu Balci-Peynircioglu, Yeliz Z. Akkaya, and Engin Yilmaz. 2012. "The Effect of Colchicine on Pyrin and Pyrin Interacting Proteins." *Journal of Cellular Biochemistry* 113 (11): 3536–46. doi:10.1002/jcb.24231.

Tegner, H. 1978. "Quantitation of Human Granulocyte Protease Inhibitors in Non-Purulent Bronchial Lavage Fluids." *Acta Oto-Laryngologica* 85 (3–4). England: 282–89.

Theodoropoulou, Katerina, Federica Vanoni, and Michaël Hofer. 2016. "Periodic Fever, Aphthous Stomatitis, Pharyngitis, and Cervical Adenitis (PFAPA) Syndrome: A Review of the Pathogenesis." *Current Rheumatology Reports* 18 (4): 18. doi:10.1007/s11926-016-0567-y.

Thomadaki, K, Ja Bosch, Fg Oppenheim, and Ej Helmerhorst. 2013. "The Diagnostic Potential of Salivary Protease Activities in Periodontal Health and Disease." *Oral Diseases* 19 (8): 781–88. doi:10.1111/odi.12069.

Thomadaki, K, E J Helmerhorst, N Tian, X Sun, W L Siqueira, D R Walt, and F G Oppenheim. 2011. "Whole-Saliva Proteolysis and Its Impact on Salivary Diagnostics." *Journal of Dental Research* 90 (11): 1325–30. doi:10.1177/0022034511420721.

Thompson, R C, and K Ohlsson. 1986. "Isolation, Properties, and Complete Amino Acid Sequence of Human Secretory Leukocyte Protease Inhibitor, a Potent Inhibitor of Leukocyte Elastase." *Proc Natl Acad Sci U S A* 83 (18): 6692–96.

Tipton, Jeremiah D., John C. Tran, Adam D. Catherman, Dorothy R. Ahlf, Kenneth R. Durbin, and Neil L. Kelleher. 2011. "Analysis of Intact Protein Isoforms by Mass Spectrometry." *Journal of Biological Chemistry* 286 (29): 25451–58. doi:10.1074/jbc.R111.239442.

Touitou, Isabelle, Suzanne Lesage, Michael McDermott, Laurence Cuisset, Hal Hoffman, Catherine Dode, Nitza Shoham, et al. 2004. “Infevers: An Evolving Mutation Database for Auto-Inflammatory Syndromes.” *Human Mutation* 24 (3): 194–98. doi:10.1002/humu.20080.

Trabandt, Andreas, Renate E. Gay, Hans???Georg ???G Fassbender, and Steffen Gay. 1991. “Cathepsin B in Synovial Cells at the Site of Joint Destruction in Rheumatoid Arthritis.” *Arthritis & Rheumatism* 34 (11): 1444–51. doi:10.1002/art.1780341116.

Tunca, M, S Akar, F Onen, H Ozdogan, O Kasapcopur, F Yalcinkaya, E Tutar, et al. 2005. “Familial Mediterranean Fever (FMF) in Turkey: Results of a Nationwide Multicenter Study.” *Medicine (Baltimore)* 84 (1): 1–11. doi:10.1097/01.md.0000152370.84628.0c.

Turk, Vito, and Wolfram Bode. 1991. “The Cystatins: Protein Inhibitors of Cysteine Proteinases.” *FEBS Letters*. doi:10.1016/0014-5793(91)80804-C.

Venkatesan, Kavitha, Jean-françois Rual, Alexei Vazquez, Irma Lemmens, Tomoko Hirozane-kishikawa, Tong Hao, Xiaofeng Xin, et al. 2010. “NIH Public Access.” *Systems Biology* 6 (1): 83–90. doi:10.1038/nmeth.1280.An.

Vitorino, Rui, Maria João Calheiros-Lobo, José A. Duarte, Pedro M. Domingues, and Francisco M L Amado. 2008. “Peptide Profile of Human Acquired Enamel Pellicle Using MALDI Tandem MS.” *Journal of Separation Science* 31 (3): 523–37. doi:10.1002/jssc.200700486.

Vitorino, Rui, Maria João Calheiros-Lobo, Jason Williams, António J. Ferrer-Correia, Kenneth B. Tomer, José A. Duarte, Pedro M. Domingues, and Francisco M L Amado. 2007. “Peptidomic Analysis of Human Acquired Enamel Pellicle.” *Biomedical Chromatography* 21 (11): 1107–17. doi:10.1002/bmc.830.

Vitorino, Rui, Sofia De Morais Guedes, Rita Ferreira, Maria João C Lobo, José Duarte, António J. Ferrer-Correia, Kenneth B Tomer, Pedro M. Domingues, and Francisco M L Amado. 2006. “Two-Dimensional Electrophoresis Study of in Vitro Pellicle Formation and Dental Caries Susceptibility.” *European Journal of Oral Sciences* 114 (2): 147–53. doi:10.1111/j.1600-0722.2006.00328.x.

Vogelmeier, Claus, Adrian Gillissen, and Roland Buhl. 1996. “Use of Secretory Leukoprotease Inhibitor to Augment Lung Antineutrophil Elastase Activity.” *Chest*. doi:10.1378/chest.110.6.

Vogl, Thomas, Christian Pröpper, Michael Hartmann, Anke Strey, Kerstin Strupat, Christian Van Den Bos, Clemens Sorg, and Johannes Roth. 1999. “S100A12 Is Expressed Exclusively by Granulocytes and Acts Independently from MRP8 and MRP14.” *Journal of Biological Chemistry* 274 (36): 25291–96. doi:10.1074/jbc.274.36.25291.

Wen, Jie, Nikolaos G Nikitakis, Risa Chaisuparat, Teresa Greenwell-Wild, Maria Gliozzi, Wenwen Jin, Azita Adli, et al. 2011. "Secretory Leukocyte Protease Inhibitor (SLPI) Expression and Tumor Invasion in Oral Squamous Cell Carcinoma." *The American Journal of Pathology* 178 (6): 2866–78. doi:10.1016/j.ajpath.2011.02.017.

White, M. R., E. J. Helmerhorst, A. Ligtenberg, M. Karpel, T. Tecle, W. L. Siqueira, F. G. Oppenheim, and K. L. Hartshorn. 2009. "Multiple Components Contribute to Ability of Saliva to Inhibit Influenza Viruses." *Oral Microbiology and Immunology* 24 (1): 18–24. doi:10.1111/j.1399-302X.2008.00468.x.

Whiteaker, Jeffrey R, Chenwei Lin, Jacob Kennedy, Liming Hou, Mary Trute, Izabela Sokal, Ping Yan, et al. 2011. "A Targeted Proteomics-Based Pipeline for Verification of Biomarkers in Plasma." *Nat Biotech* 29 (7). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 625–34. http://dx.doi.org/10.1038/nbt.1900.

Wittkowski, Helmut, Michael Frosch, Nico Wulffraat, Raphaela Goldbach-Mansky, Tilmann Kallinich, Jasmin Kuemmerle-Deschner, Michael C. Frühwald, et al. 2008. "S100A12 Is a Novel Molecular Marker Differentiating Systemic-Onset Juvenile Idiopathic Arthritis from Other Causes of Fever of Unknown Origin." *Arthritis and Rheumatism* 58 (12): 3924–31. doi:10.1002/art.24137.

Wu, Y, R Shu, L. J. Luo, L. H. Ge, and Y. F. Xie. 2009. "Initial Comparison of Proteomic Profiles of Whole Unstimulated Saliva Obtained from Generalized Aggressive Periodontitis Patients and Healthy Control Subjects." *Journal of Periodontal Research* 44 (5): 636–44. doi:10.1111/j.1600-0765.2008.01172.x.

Yalçinkaya, Fatos, Seza Ozen, Zeynep Birsin Ozçakar, Nuray Aktay, Nilgün Cakar, Ali Düzova, Ozgür Kasapçopur, et al. 2009. "A New Set of Criteria for the Diagnosis of Familial Mediterranean Fever in Childhood." *Rheumatology (Oxford, England)* 48 (4): 395–98. doi:10.1093/rheumatology/ken509.

Yates, John R, Cristian I Ruse, and Aleksey Nakorchevsky. 2009. "Proteomics by Mass Spectrometry: Approaches, Advances, and Applications." *Annual Review of Biomedical Engineering* 11: 49–79. doi:10.1146/annurev-bioeng-061008-124934.

Yin, A., H. C. Margolis, Y. Yao, J. Grogan, and F. G. Oppenheim. 2006. "Multi-Component Adsorption Model for Pellicle Formation: The Influence of Salivary Proteins and Non-Salivary Phospho Proteins on the Binding of Histatin 5 onto Hydroxyapatite." *Archives of Oral Biology* 51 (2): 102–10. doi:10.1016/j.archoralbio.2005.06.003.

Zavasnik-Bergant, Tina. 2008. "Cystatin Protease Inhibitors and Immune Functions." *Frontiers in Bioscience : A Journal and Virtual Library* 13: 4625–37. doi:10.2741/3028.

Zeng, Qiong, Kun Lin, Mianxuan Yao, and Liling Wei. 2015. "Significant Correlation between Cystatin C, Cerebral Infarction, and Potential Biomarker for Increased Risk of Stroke." *Current Neurovascular Research* 12

(1): 40–46. doi:10.2174/156720261266615010215094l.

Zhang, Aihua, Hui Sun, Ping Wang, and Xijun Wang. 2013. “Salivary Proteomics in Biomedical Research.” *Clinica Chimica Acta* 415. Elsevier B.V.: 261–65. doi:10.1016/j.cca.2012.11.001.

Zimmer, Danna B, Patti Wright Sadosky, and David J Weber. 2003. “Molecular Mechanisms of S100-Target Protein Interactions.” *Microscopy Research and Technique* 60 (6): 552–59. doi:10.1002/jemt.10297.

PART 2

TOP-DOWN AND BOTTOM-UP PROTEOMICS APPROACHES ON THE FRACTION OF LOW MOLECULAR WEIGHT OF PROTEIN EXTRACTS FROM HUMAN COLONIC MUCOSA

1.0 *Introduction*

1.1 *Colorectal cancer*

Colorectal cancer (CRC), also known as bowel cancer, is the development of cancer from the colon or rectum (parts of the large intestine). It is due to the abnormal growth of cells that have the ability to invade or spread to other parts of the body. Most colorectal cancers are due to old age and lifestyle factors with only a small number of cases due to underlying genetic disorders. Some risk factors include diet, obesity and smoking. Some of the inherited genetic disorders that can cause colorectal cancer include familial adenomatous polyposis and hereditary non-polyposis colon cancer; however, these represent less than 5% of cases (Haggar & Boushey, 2009). It typically starts as a benign tumor, often in the form of a polyp, which over time becomes cancerous.

Approximately, there are 1,000,000 new cases of CRC and 500,000 deaths associated with CRC each year (Tanaka, Tanaka, Tanaka, & Ishigamori, 2010). CRC represents one of the primary causes of cancer deaths in Europe and the United States (Bingham & Riboli, 2004). In Asia, including Japan, CRC is the fourth leading cause of mortality by cancer, and its incidence is increasing (Sung, Lau, Goh, & Leung, 2005). Despite improved treatment strategies involving surgery and chemo- and radio-therapy have increased the overall survival rates in the early stages, 40-50% of patients with CRC present with metastasis either at the time of diagnosis or as recurrent disease upon intended curative therapy (Calon et al., 2012).

Most of these tumors are adenocarcinomas originating from adenomatous polyp arising from the glandular epithelium of the intestine. Adenomas are initiated by somatic mutation of the tumor suppressor gene *APC* (Lamlum et al., 2000). Other frequent genetic alterations include activating mutations in KRAS and BRAF, inactivation of TP53, alterations of the PI3K/Akt and the TGFβ signaling pathways (Jones et al., 2008). Specific chromosome copy number changes, such as a loss of chromosome 18q or a gain of chromosome 20q, have also been associated with progression (Hermsen et al., 2002).

These chromosomal gains and losses can also be detected in a minority of adenomas, and the adenomas that present these changes are therefore considered to be at high-risk of progression (Carvalho et al., 2012).

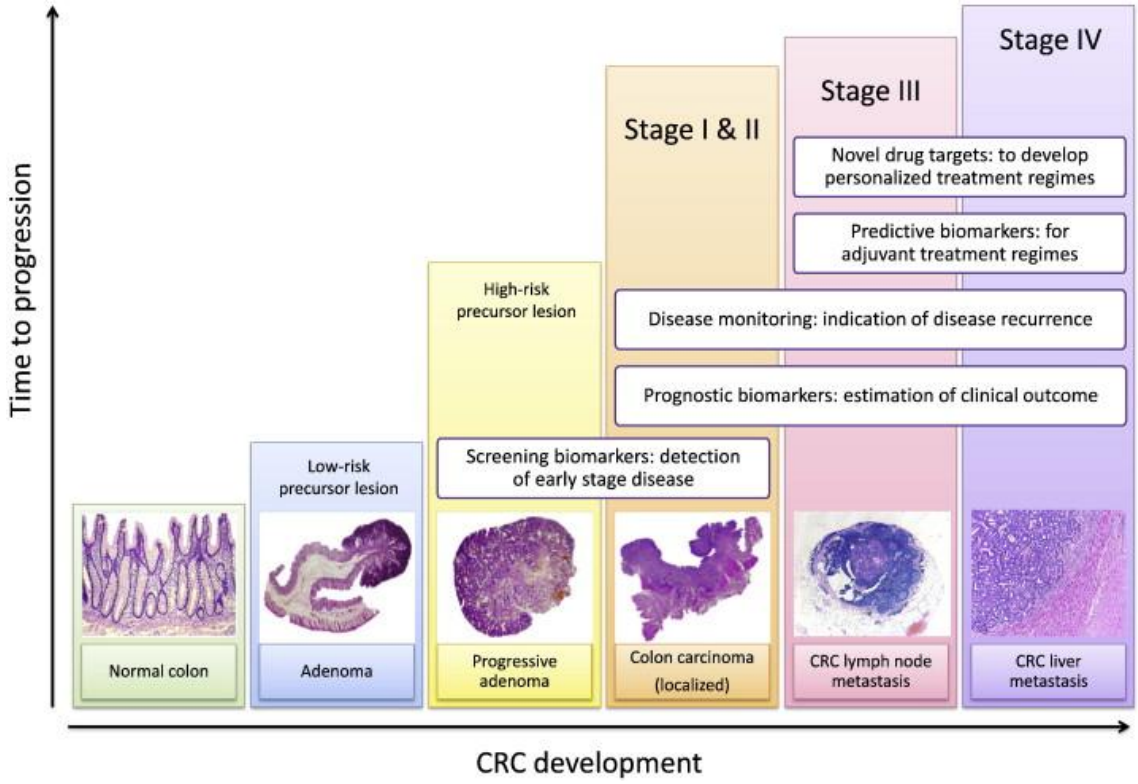
The accumulation of genetic mutations in accordance with chromosomal instability, shifts the normal intestinal lining to an adenomatous polyp, then high-grade adenoma and finally to a carcinoma (Markowitz & Bertagnolli, 2009). CRC can also arise from nonpolypoid and

depressed lesions. Although these lesions are less common than that of the polypoid adenoma, they manifest more aggressive behavior and more rapid growth, and they are more difficult to diagnose (Hurlstone et al., 2004). The early diagnosis is the most important predictor of survival for patients with colorectal cancer (Yu et al., 2008). Development of colorectal cancer has been viewed as an ordered process in which three main phases could be discerned: initiation, promotion and progression. The Fig. 1 shows the sequential step that lead to the onset and progression of the CRC: i) the first step is represented by pre-malignant precursor stage or adenoma ii) adenoma can progress into a carcinoma iii) the stages I and II are represented by no lymph node involvement, and in this phase patients receive surgery only and are not recommended for adjuvant chemotherapy iv) the next step is the stage III where the lymph node metastases is present v) finally, the stage IV where metastases have spread to distant organs (de Wit, Fijneman, Verheul, Meijer, & Jimenez, 2013).

The extent of the disease in terms of local invasion, spread to lymph nodes and distal organs at time of diagnosis, referred to as stage of disease, is an important prognostic factor, with five-years survival rates of more than 90% for localized CRC (stage I) and only about 10% for CRC that metastasized to distant organs (stage IV) (Barderas et al., 2010). Surgery remains the primary modality of treatment for malignancies of the lower gastrointestinal tract, and standard resection is the best therapy required for early-stage cancer. In relation to tumor progression in terms of depth of penetration and lymph node involvement, the chance of cure with surgery alone diminishes; in fact, rates of local recurrence and survival are dependent on the tumor–node–metastasis (TNM) stage (Nelson et al., 2001). Prognosis and treatment regimens are mainly dependent on tumor stage. Invasive cancers that are confined within the wall of the colon (TNM stages I and II) are curable, but if untreated, they spread to regional lymph nodes (stage III) and then metastasize to distant sites (stage IV) (Ginsberg RJ, Vokes EE, 1997). Stage I and II tumors are curable by surgical excision, and up to 73% of cases of stage III disease are curable by surgery combined with adjuvant chemotherapy (Concepts, 2012). Recent advances in chemotherapy have improved survival, but stage IV disease is usually incurable (Ginsberg RJ, Vokes EE, 1997); (cancer.org).

Tumor metastasis is a multi-step process by which tumor cells disseminate from their primary site and form secondary tumors at a distant site. Metastasis is the major cause of death in the vast majority of cancer patients. However, the mechanisms underlying each step remain obscure (Barderas, Babel, & Casal, 2010).

Fig. 1. The different stages of development are depicted by hematoxylin and eosin stained examples on the x-axis. On the y-axis time of progression is indicated by colored bars representing the different stages. Clinical needs for biomarkers are depicted within these bars indicating the stages for which they are relevant. Screening biomarkers are needed for early detection in screen-relevant lesions. Prognostic biomarkers are warranted from stage I on to predict disease outcome (a priori of adjuvant therapy) and to select for patients that would benefit from adjuvant chemotherapy. Disease surveillance biomarkers are also needed from stage I on to monitor disease recurrence. Predictive biomarkers are needed for more advanced stages to select patients for adjuvant targeted therapies. Finally, there is a need for novel drug targets, especially for late stage disease to improve clinical outcome (M. de Wit et al., 2013).



1.2 *Dukes classification*

A stage classification still adopted is that of Dukes. In 1932, Dukes (Dukes, 1980) stated that in its earliest stages, rectal cancer begins as an epithelial proliferation rising from the surface by a preexisting adenoma. The cancer metastasizes through the bowel wall to the lymphatics. Cases in which the carcinoma is limited to the wall of the rectum were designated A.

Those in which the cancer has spread by direct continuity to the extrarectal tissue were designated B. Cases in which metastases are present in the regional lymph nodes were called C (Fig. n° 2). A more advanced pathologic stage was associated with a worse prognosis.

In 1949, Kirklin, Dockerty, and Waugh (KIRKLIN, DOCKERTY, & WAUGH, 1949) proposed a modification of Dukes' classification.

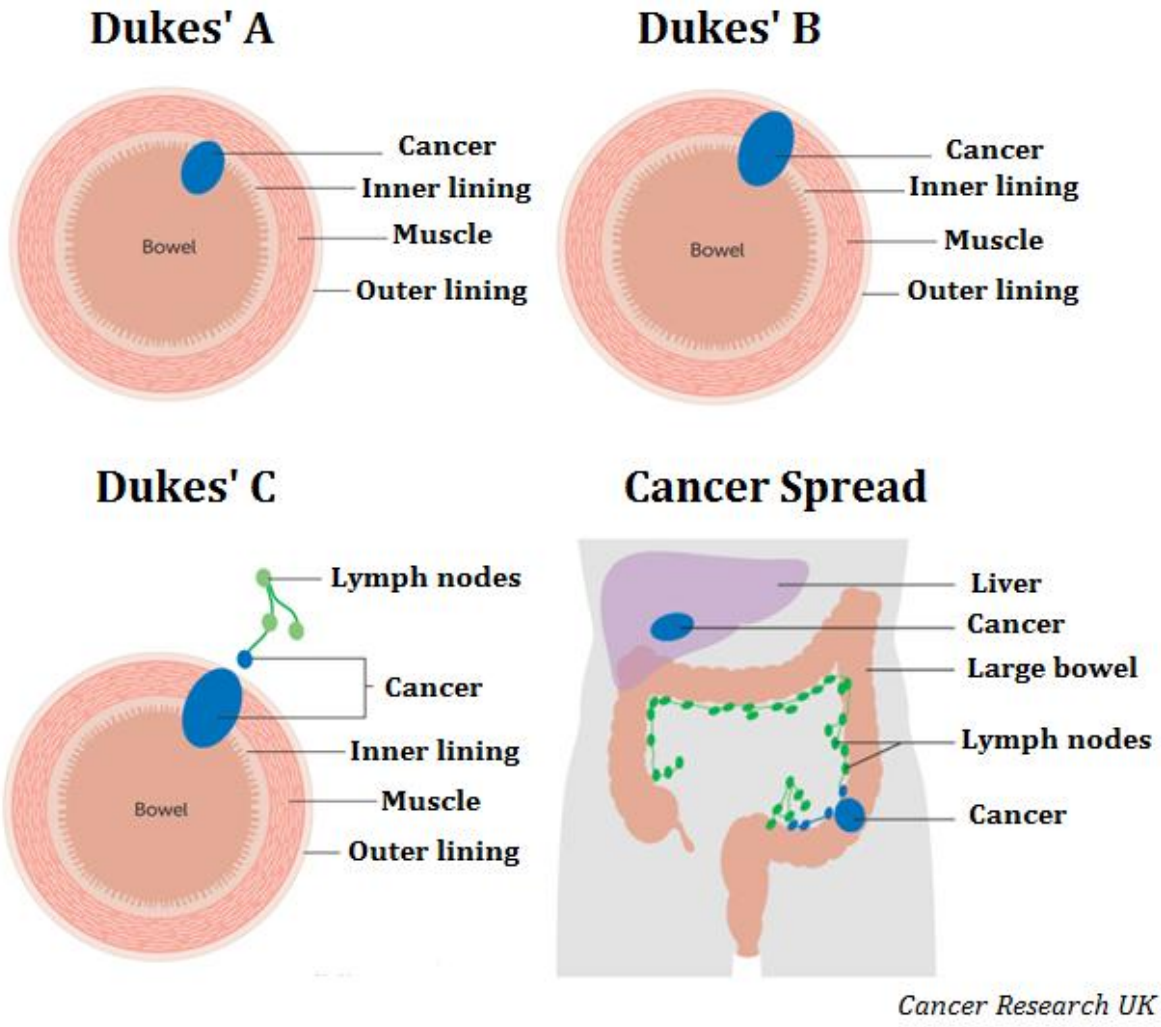
The authors preserved the A, B, C framework but added, for B lesions, the subscript designation “1” for lesions that have extended into, but not through the muscularis propria and “2” for tumors that have penetrated the muscularis propria.

Finally, in 1954, Astler and Coller (Astler & Coller, 1954) reported on specimens of the rectum and colon removed at surgery and classified them using Dukes' classification as modified by Kirklin et al. (KIRKLIN et al., 1949).

The Modified Astler–Coller (MAC) classification is shown below:

- Type A: when the lesion is limited to the mucosa;
- Type B1: when lesions extending into the muscularis propria, but not penetrating it, with negative nodes;
- Type B2: when lesions penetrating the muscularis propria, with negative nodes;
- Type C1: when lesions extending into the muscularis propria, but not penetrating it, with positive nodes;
- Type C2: when lesions penetrating the muscularis propria with positive nodes.

Fig. 2. Phases of Dukes' classification. Fig. taken from the World Health Organization Classification of Tumours.



1.3 *Biomarker discovery for colo-rectal cancer.*

Colo-rectal cancer is a major public health problem in many parts of the world and the second most common cause of cancer-related deaths in Europe and other Western countries (Álvarez-Chaver, Otero-Estévez, de la Cadena, Rodríguez-Berrocal, & Martínez-Zorzano, 2014). Emerging collections of genomics data for human cancers present the formidable challenge of understanding how genomic abnormalities drive the biological and clinical characteristics of cancer. The pathogenesis of CRC is a progressive accumulation of mutations in multiple genes, much less is known at the proteome level (Álvarez-Chaver et al., 2014). This task will be facilitated by proteomic analyses, which provides an intermediate layer of biological information that is more directly connected to phenotype.
Thank to recent advances in new technologies and approaches, immense efforts have been put in proteomics and genomics fields to deliver detailed analysis of the genes and proteins, to gain a more complete understanding of cellular systems at both genomic and proteomic levels, allowing a mechanistic understanding of the human diseases and opening avenues for identification of novel gene and protein based prognostic and therapeutic markers (Sethi, Hancock, & Fanayan, 2016). Firstly, early detection of the disease seems a realistic approach to reduce CRC mortality, since there is a well-defined benign precursor lesion (i.e. an adenoma) and there is a large time-span during which curative intervention can take place (de Wit et al., 2013). Therefore screening programs are being implemented in various countries (Hoff & Dominitz, 2010). These screening methods are frequently based on stool tests, although in some countries primarily colonoscopy based screening is provided (Sillars-Hardebol, Carvalho, Van Engeland, Fijneman, & Meijer, 2012). Current stool tests for CRC screening detect haem or hemoglobin in stool, which ends up in the stool because of a tumor bleeding. However, as hemoglobin originates from red blood cells it is not a marker specific for neoplastic cells, and therefore can also be present in the stool due to non-tumor lesions. Usually, presence of hemoglobin in the stool dictates a follow-up with colonoscopy (de Wit et al., 2013). Therefore, screening biomarkers should have high specificity to prevent a high number of false-positive subjects to be referred for further tests. New biomarkers are urgently needed to improve diagnosis and prognosis of cancer. Discovering biological markers for early detection, when treatment is most effective, is essential to prevention and long-term survival of patients. Development of reliable biomarkers requires an increased understanding of the CRC biology and the underlying molecular and cellular mechanisms of this disease (Sethi et al., 2016). Different proteomic tools are used for the discovery of

candidate protein markers for CRC, such as two-dimensional electrophoresis methods, quantitative mass spectrometry-based techniques or protein microarrays (Álvarez-Chaver et al., 2014). Moreover, colon cancer cell lines can be used to investigate specific (signaling) pathways in tumor biology, e.g. by gene knockdown or overexpression, and to address questions related to drug treatment response (de Wit et al., 2013). Yin X. and colleagues performed a label-free quantification of paraffin-embedded (FFPE) colorectal tissues; their results reported relative quantification information of 1017 proteins by label-free and 6294 proteins by iTRAQ. Three potential metastasis biomarkers 1) integrin alpha 5 (ITA5) 2) actin-related protein (ARP3) 3) proteins vitronectin (VTN) were evaluated by functional annotation and validated by Western blotting (Yin et al., 2015). Wisnieswskj JR and colleagues reported a proteomic analysis of microdissected material from formalin-fixed and FFPE colorectal cancer, quantifying >7500 proteins between patient matched normal mucosa, primary carcinoma, and nodal metastases. They found in stomatin-like 2 (STOML2), baculoviral IAP repeat-containing protein 6 (BIRC6), and glucose transporter type 1 (GLUT1) proteins, potential biomarkers because upregulated in cancer samples (Wisniewski et al., 2012). Finally, K. Sethi and colleagues proposed CRC glycoprotein biomarkers, such as carbohydrate antigen 19-9 (CA 19-9) (Sethi et al., 2016). In fact, cancer progression is accompanied by several parameters, including changes in the extent and nature of protein glycosylation and increased levels of blood glycoproteins. Numerous studies have demonstrated the link between aberrant glycosylation and tumor behavior, by promoting tumor progression, metastasis, and invasion (Kannagi, Sakuma, Cai, & Yu, 2015). The interest of scientific research on discovering potential biomarkers for colo-rectal cancer can be easily demonstrated by performing a simple search on PubMed, where is possible to find about 980 publications regarding this topic. Such a large number highlights the strong need for scientific community to find out possible biomarkers of interest for the disease and, at the same time, reveal that cancer necessitates to be profoundly explored in proteomic research field. Actually, only a few proteins are being used as biomarkers in CRC (Barderas et al., 2010); the first critical issues in proteomic analysis are the selection of the sample set (plasma, tissue, cell lines) and its manipulation (into paraffin, frozen or fresh tissue), and the different sampling tissue areas in the same tumor from which samples are collected. Cell lines and tissue samples have been used indistinctively for proteomic analysis. However, there are many doubts about how accurately cell lines reflect the cancer proteome, as they are reasonably heterogeneous and probably altered by culture conditions (Barderas et al., 2010); conversely, tumor tissue represents the most direct approach to

identify biomarkers as they are most likely present in cancer tissues at higher concentrations. Secondly, the different proteomic techniques can affect the search for possible biomarkers. Initial CRC proteomic studies were carried out mainly by 2-D PAGE on different colorectal tumor cell lines but in general, the number of tumor samples analyzed in these studies was relatively low (11%) (Sagynaliev et al., 2005). These results highlight the necessity of taking extreme care in data interpretation, avoiding excessive overstatements based on these results and underline some of the current limitations of 2-D technology, mainly related to the detection of relatively abundant proteins. Therefore, more sensitive techniques are needed to detect the presence of low-abundant proteins and their PTMs. Finally, another problem in biomarker search in CRC has been the limited number of samples analyzed in published reports (≤16 cancer samples per study), probably due to the difficulties in sample collection and the high price of the proteomic analysis.

Little work has been done so far to validate these candidate biomarkers (Álvarez-Chaver et al., 2014). Hence, it would be appropriate to collect samples of large cohorts of healthy subjects and CRC patients and analyze their peptidome and proteome in depth to fully appreciate the potential of proteomic biomarker changes. This kind of validation process would allow the transfer of new biomarkers to clinical use enabling a better detection and treatment of CRC (Álvarez-Chaver et al., 2014).

1.4 *Thymosin β-4 and β-10 as potential biomarker in colo-rectal cancer*

Thymosin β4 (Tβ4) is the major component of β-thymosin family, composed by 16 peptides, originally isolated in thymus, structurally correlated with aminoacidic sequence highly conserved. Tβ4 presence was subsequently identified in a variety of tissues and cells (Thomas Huff, Müller, Otto, Netzker, & Hannappel, 2001). A number of studies showed that Tβ4 is a ubiquitous peptide acting as a multifunctional bioactive peptide with an essential role in protecting and restoring functionality of many cells, tissues and organs. In humans, thymosin β4, β10 (Tβ10) and β15 (Tβ15) codified by different genes and functionally distinct have been identified (Hannappel, 2007) ; (Goldstein, Hannappel, & Kleinman, 2005). Tβ4 is a 43 aminoacid peptide, N-term acetylated after methionine removal, codified by *TMSB4X* gene; it shows a dynamic conformation in reason of its non-structured and flexible configuration that allows the interaction with various proteins and to act with multiple functions, both intra- and extra-cellular, such as nuclear transcription factor for instance (Mannherz & Hannappel, 2009) ; (Thomas Huff et al., 2004). Various

post-translational modification can affect Tβ4 structure and function (proteolytic cleavage, generation of isopeptidic cross-linking, phosphorylation or oxidation processes) however, the functional meaning of those has not been completely elucidated (Hannappel, 2010).

As mentioned before, presence of Tβ4 peptide has been revealed in major cellular types such as leucocytes, macrophages and platelets, in tissues like spleen and thymus, but not in erythrocytes (Mannherz & Hannappel, 2009). Tβ4 blood concentration is between 10-200 nM (both serum and plasma), 100 nM in saliva and 200 µM in gingival-crevicular fluid (Inzitari et al., 2009). Release mechanism of Tβ4 peptide through secretion, cellular lysis or necrosis, remains to be elucidated however, because Tβ4 precursor lack of release signal peptide, its release should follow a non-canonical pathway of secretion. Equally enigmatic are the mechanism and the receptor system with which Tβ4 would induce specific answers. Immunohistochemical and biochemical approaches have been applied by our research group to clarify the secretion pathway of the peptide. Salivary glands samples collected at different gestational stages showed that, during intrauterine life, expression levels of the peptide is maximal and localized in ducts, beads, in lumen of the thymus glands. Tβ4 concentration tends to reduce in the newborn, mostly localizing in ducts and almost disappearing in children and then in adult stage (Sonia Nemolato et al., 2009). Comparison of Tβ4 expression levels in the various districts of gastrointestinal and genitourinary tracts of the fetus and adult showed wide differences, supporting the hypothesis of the different expression stages of the peptide through the various phases of life: tongue, esophagus, stomach, ileum and colon show a Tβ4 immunoexpression greater in the fetus than in adults, with a different tissue localization. The pancreas is immunoreactive for the peptide both in the fetus than in adults, with a significant difference from the liver, which appears totally negative for Tβ4 in the fetus and, surprisingly, positive in adults (Sonia Nemolato et al., 2010). Bladder, prostate, endometrium, ovary and testis show positivity for the peptide in the fetus as well in adults, with similar locations but different intensity of reaction. The fetal kidney has strong positivity for Tβ4 in the tubules in formation and it is maintained even in the adult kidney glomeruli, with consistently negative in both age groups (S Nemolato, Cabras, Fanari, Cau, Fanni, et al., 2010). Extensive research on Tβ4 immunoreactivity in adult liver shown that the peptide accumulates in large grains with a well-defined zonal localization in mature hepatocytes, counting the liver among the major organs of Tβ4 synthesis in the adult (S Nemolato et al., 2011). Recent interesting data on Tβ4 expression revealed its presence in tissue mast cells, both far and next to the tumor site (S Nemolato, Cabras, Fanari, Cau, Fraschini, et al., 2010), suggesting the involvement of the peptide in

inflammatory processes and in peritumoral tissue reaction sites. The main role is thought to control the assembly and disassembly of cytoskeletal actin filaments, fundamental process for differentiation, migration and cell adhesion and organogenesis (Mannherz & Hannappel, 2009) ; (Goldstein et al., 2005); in fact, Tβ4 can bind G-actin monomers establishing complexes with stoichiometry 1: 1, also through the formation of isopeptidic cross-linking (Safer, Sosnick, & Elzinga, 1997). In addition to its role as actin polymerization regulator, Tβ4 is involved in many critical biological processes, including angiogenesis, wound healing, the inflammatory response, cell migration and intracellular signaling through the AKT pathway (Malinda et al., 1999). Tβ4 is also capable to interact with fibrin, promoting a fast tissue remodeling and contemporary stimulating wound healing through collagen deposition, endothelial cells and keratinocytes migration (Thomas Huff, Otto, Muller, Meier, & Hannappel, 2002). Moreover, it has been demonstrated that Tβ4 sulfoxide owns an important anti-inflammatory role through neutrophils chemotaxis inhibition (Young et al., 1999). Tβ4 peptide also promotes angiogenesis, stimulating stem cells and/or progenitor cells differentiation leading to generation of new blood vessels (Goldstein, Hannappel, Sosne, & Kleinman, 2012); a proteolytic cleavage at the N-terminal end of Tβ4 produces tetrapeptide ac-SDKP (or seraspenide) that possess a pro-angiogenic and antifibrotic action (Hannappel, 2010). It is known that β-thymosins, including Tβ4 and Tβ10 are involved not only in normal cell migration, but also in tumor metastasis (Sribenja et al., 2016).

Cell migration is one of the fundamental cellular events of life and triggers many physiological and pathological processes such as embryonic development, wound healing, tissue repair, angiogenesis, vascular remodeling, inflammation, and neuronal out growth. Aberrant cell migration contributes to pathologies such as cardiovascular diseases, tumor metastatic cascade including tumor angiogenesis, invasion, and metastasis (Sribenja et al., 2016). Expression of Tβ4 characterizes many malignant tumors and such expression has been proposed to contribute to the malignant phenotype (Hong, Lee, Hong, & Hong, 2016); (Fu et al., 2015) ; (Yoon et al., 2011).

In colorectal cancer it was observed a strong immunoreactivity for Tβ4 and a positivity of tumor cells in the process of epithelial-mesenchymal transformation on the front of tumor infiltration. This preliminary observation led to propose a major new feature of Tβ4 in promoting the infiltration and tumor progression (Sonia Nemolato et al., 2012). The process of epithelial-mesenchymal transformation seems the link between tumor progression and fetal development program, explaining in part how the peptide is re-expressed in pathological conditions in adults using their own mechanisms of fetal life (Faa et al., 2012).

The Tβ4 could be a new molecular target in some types of cancer. In particular, the data suggest that the Tβ4 is over-expressed in metastatic cells and that tumors with high expression of the peptide are more aggressive and tend to metastasize more easily and to be typically more resistant to chemotherapy (Goldstein, 2003). Recently it has been observed in studies conducted by my research group, that in cell cultures of liver and colon cancers (HepG2 e Caco2), the Tβ4 may, under certain conditions, to move from the cytoplasm to the nucleus. So, not only varies the expression of the peptide but also its localization. The nuclear translocation of the peptide in the tumor cells was also observed by other researchers (Thomas Huff et al., 2002) ; (Cha, Jeong, & Kleinman, 2003).

Tβ10 has recently been recognized as being an important player in the metastatic cascade including tumor angiogenesis, invasion, and metastasis. In a study on β-thymosins expression in hepatocellular carcinoma, both Tβ4 and Tβ10 were detected in tumor cells. Moreover, Tβ10 showed a strong expression in cells undergoing stromal invasion, in contrast with the absence of reactivity for Tβ4 (Theunissen et al., 2014). Tβ10 has been associated to several cancer: it was detected in the majority of the goiters, hyperproliferative cancer tissue, and thyroid adenoma, but not in normal thyroid (Chiappetta et al., 2004); Tβ10 expression levels correlated significantly with the stage of lung cancer, distant metastases, lymph node metastases, and degree of differentiation of lung cancer (McDoniels-Silvers, Nimri, Stoner, Lubet, & You, 2002); (Y. J. Lee et al., 2011). In breast cancer, Tβ10 was detected mainly in the malignant tissues (Verghese-Nikolakaki, Apostolikas, Livaniou, Ithakissios, & Evangelatos, 1996), moreover, the expression of Tβ10 was down-regulated in ovarian cancer compared with normal ovary tissues (S. H. Lee et al., 2001). About colorectal cancer still poorly is known about this peptide.

2.0 Objective of the study

Differently from the several studies on CRC proteomics performed before, the main objective of the study described in this thesis has been to highlight significant quantitative and qualitative differences for the Tβ4 and Tβ10 associable to CRC invasion and stadiation. To this aim, healthy mucosa, and tumoral tissues from the surface and the deep layer of the tumor from a same patient were collected, and a method of protein extraction was optimized to separate tissue extracts containing the low-molecular weight protein fraction, which were analyzed by HPLC-ESI-MS. A second objective, was to investigate qualitative and quantitative variations concerning other peptides and proteins present in the protein extracts by proteomic platforms, as a function of the kind of tumoral tissue and of the tumor stadium. The results of the study may be useful to individuate potential markers correlated to the progression or to the stadium of the tumor and suggestive for therapeutic address.

3.0 Experimental

3.1 *Reagents and apparatus*

All chemicals and reagents used were of analytical grade and were purchased from Sigma Aldrich (St. Louis, MI), Merck (Darmstadt, Germany) and Bio-Rad (Hercules, CA). Low-resolution HPLC-ESI-MS measurements were carried out by means of a Surveyor HPLC system (ThermoFisher, San Jose, CA, USA) connected to a LCQ Advantage mass spectrometer (ThermoFisher Scientific San Jose, CA). The chromatographic column was a reversed-phase Vydac 208MS-C8 (Hesperia, CA, USA) with 5 μm particle diameter (column dimensions was 150x2.1 mm). High-resolution HPLC-ESI-MS/MS experiments were carried out using an Ultimate 3000 Micro HPLC apparatus (Dionex, Sunnyvale, CA, USA) equipped with a FLM-3000-Flow manager module and coupled to an LTQ Orbitrap XL apparatus (Thermofisher). The columns were a Zorbax 300SB-C8 column (3.5 μm particle diameter; 1.0 x 150 mm) for the top-down analysis, and a Zorbax 300SB-C18 column (5 μm particle diameter; 1.0 x 150 mm) for the bottom-up.

3.2 *Samples and Study subjects*

The study included 22 patients submitted to surgical resection of colo-rectal tumors or adenomas by the unit of Colorectal Surgery of the Department of Surgery (Cagliari University). Ethics Committee approval was obtained for the study and full written consent forms were obtained from the donors. The tissue extracts were prepared from biopsies of intestinal mucosa of the patients with colo-rectal cancer (CRC) or adenomas (NON CRC). Colon cancers were included when characterized by budding margins with evident morphological signs of epithelialmesenchymal transition. Three different tissue samples of different sizes (but not exceeding 1 x 1 cm) were provided from each patients. They corresponded to: i) the surface layer of the tumor, ii) the deep layer of the tumor, iii) the normal colon mucosa, respectively named "*S*" superficial, "*D*" deep and "*H*" health. For case 1 only *S* and *D* tumoral tissues were provided.

The clinical diagnosis was performed by the unit of Pathologic Anatomy of Department of Surgery (University of Cagliari). Patients were classified in two groups: 18 patients were CRC, 4 NON-CRC. In Table 1 the type of CRC (expansive or infiltrating) and the Dukes Stadium for each case included in the study are reported.

Table 1. Type of CRC (expansive or infiltrating), Dukes Stadium for each patient.

Patients	Type of CRC	Dukes Stadium
#1	Unknown	Adenomas Low grade
#2	Infiltrating	C
#3	Unknown	Adenomas Low grade
#4	Infiltrating	B
#5	Infiltrating	B
#6	Infiltrating	C
#7	Infiltrating	B
#8	Infiltrating	B
#9	Infiltrating	A
#10	Unknown	Adenomas Low grade
#11	Infiltrating	C
#12	Infiltrating	B
#13	Infiltrating	C
#14	Infiltrating	C
#15	Infiltrating	C
#17	Expansive	A
#18		C
#19	Infiltrating	C
#20	Unknown	Adenomas High grade
#22	Infiltrating	A
#23	Expansive	B
#24	Expansive	A

A= low stadium B= medium stadium C= high stadium

3.3 *Protein extraction*

After surgical resection, tissue samples were immediately washed from blood residues with a physiological solution, and dipped in 600 µL of extraction buffer to be homogenized in an ice bath. The extraction buffer was composed as follows: 25 mM Hepes pH 7.8, 50 mM KCl, beta-Octyl-glucopyranoside 0.2%, 1 mM dithiothreitol. To inhibit proteases one Mini-Complete ™ pill (Roche Diagnostics) was added to 10 mL of buffer. The homogenization was performed with Ultra Turrax apparatus and followed by three cycles of 5 min in a sonication bath.

The homogenized tissues, were centrifuged at 13000 rpm, 4°C for 10 min and an aliquot of 5µl of the supernatant was used for determining the total protein concentration by Bicinchoninic Acid (BCA) assay (QuantiPro BCA assay kit Sigma-Aldrich); the resulting concentrations were corrected on the basis of initial volumes of raw extract recovered after homogenization and expressed as mg of total protein in 1 mL of solution. Two different protein fractionating procedures were utilized and compared: a) ultrafiltration with 30 KDa cut-off membranes, b) treatment with a 0.05% TFA, 20% ACN hydro-organic solution.

3.3 a) *Ultrafiltration of the raw extract*

The supernatants were submitted to ultra-filtration with 30 kDa cut-off membranes (Amicon® Ultra Centrifugal Filters della Merck-Millipore Corporation). In this way it were obtained two different fractions: 1) fraction <30KDa, 2) fraction >30KDa. The fraction <30KDa was recovered and subjected to dialysis in 25 mM sodium acetate buffer pH 4.3, under stirring at 4 °C for 3 hours and using the dialysis devices with a molecular cut-off of 500 Da (Float-A-Lyzer G2; Spectrum Laboratories). The fraction obtained from dialysis was lyophilized and suspended in 100µl of 0.1% aqueous TFA. 33µl of these were immediately analyzed by low resolution RP-HPLC-ESI-MS, and 20µl were used for structural characterization, which was performed by high resolution HPLC-ESI-MS/MS analysis (with an LTQ Orbitrap XL apparatus).

The remaining volumes were stored at -80°C for the further analysis.

The fraction >30KDa retained by filter membrane was solubilized in 200µl of a hydro-organic solution 0.05% TFA, 20% ACN.

The samples were sonicated with 2 cycles of 5 min and, then, centrifuged at 13000 rpm, 4°C for 10 minutes. The volume of the clear supernatant solution was reduced to 100 µL by

partial lyophilization, 33µl of these were immediately analyzed by RP-HPLC-ESI-MS whereas the remaining volumes were stored at -80 °C for the further analysis.

3.3 b) *Hydro-organic treatment of the raw extract*

This procedure was performed on protein extracts from four patients (cases #3, #4, #5, #6) for each of the tissue types *S*, *D* and *H*. Samples were in part treated with a method based on the use of a hydro-organic mixture consisting of 0.05% TFA, 20% ACN.
For this purpose, the tissue samples were divided with a scalpel into two halves of similar weight and the two parts were homogenized in an equal volume of extraction buffer using the Ultraturrax apparatus. For each homogenate 5 µl were used for the determination of protein concentration. Homogenates from *S, D*, and *H* tissues destined to ultra-filtration were treated with the procedure described in paragraph 3.3, the other corresponding homogenates were diluted with an equal volume of 0.05% TFA, 20% ACN (hydro-organic treatment). After mixing and sonication, these samples were centrifuged at 13000 rpm for 10 minutes, at 4°C, and the supernatant was lyophilized. The powder was subsequently suspended in 100 µl of TFA 0.1%, and 33 µl of this volume was immediately injected in HPLC-ESI-MS, whereas the remaining volumes were stored at -80° C for any further analysis.

3.4 *Low resolution HPLC-ESI-MS analysis*

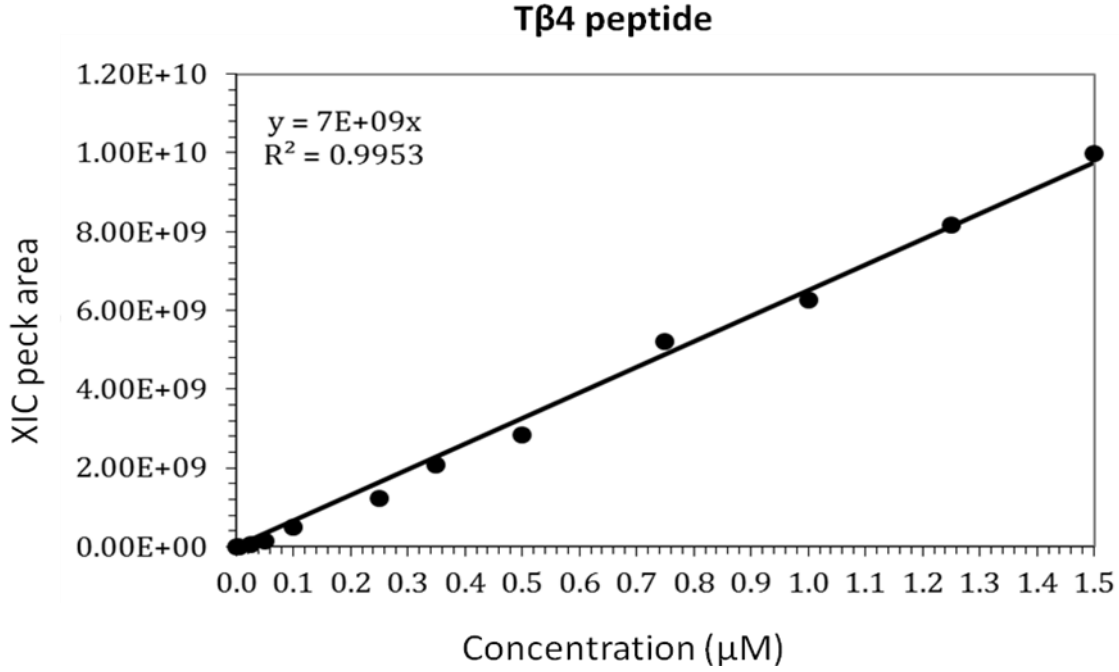
The experimental conditions used for the RP-HPLC/low resolution-ESI-MS were the same described in the experimental section of the first part of the thesis (paragraph 6.5).
Experimental mass values of each protein and peptide was obtained using the MagTran 1.0 software (Zhang & Marshall, 1998). The experimental values were compared with the theoretical masses of the proteins present at the human UniProtKB Data Bank available on Exspasy website (http://www.expasy.org/), by using the TagIdent tool (http://web.expasy.org/tagident/), in order to hypothesize an identification. Proteins and peptides observed by low-resolution MS were characterized by high resolution MS/MS as described in the paragraph 3.6.

3.5 *Quantification and statistical analysis*

The quantification was performed by HPLC/low-resolution-ESI-MS by a label-free method based on XIC peak areas analysis, as described in the experimental section of the first part of the thesis (paragraph 6.5).

The m/z of the multi-charged ions selected to quantify the peptides and proteins investigated in this study are reported in the Supplemental Table S1 (Supporting Information section), where elution time, average and monoisotopic mass values (experimental and theoretical) are indicated for each component. The areas of XIC peaks, measured in the <30 KDa and in >30KDa fractions, have been corrected on the base of the total protein concentrations measured in the total extract of the same samples. Furthermore, in the case of the β-thymosins the areas of XIC peaks were correlated to the concentrations expressed in μM, by LC-ESI-MS analysis of known concentration solutions of the standard Tβ4 peptide (0.005, 0.025, 0.05, 0.10, 0.25, 0.35, 0.5, 0.75, 1.00, 1.25, 1.50 μM). The linear regression analysis is shown in Fig. 3. The same slope was used to estimate the Tβ10 concentration, due to the high structural similarity of these two peptides.

Fig. 3. Linear regression analysis.



To compare *S*, *D* and *H* samples statistical analysis have been performed with GraphPad Prism (5.0 version). The statistical tests was according to data distribution and variances: parametric t test (variance homogeneous); t test with Welch correction (normal distribution, variance unequal), and the nonparametric Mann-Whitney test (skewed distribution, variance unequal). Statistical analysis was considered to be significant when the p value was <0.05. Spearman or Pearson tests were used for correlation analysis, accordingly with skewed or normal distributions.

3.6 *RP-HPLC/high-resolution-ESI-MS/MS analysis*

The >30KDa and <30KDa fractions from 3 cases were submitted to RP-HPLC/high-resolution-ESI-MS/MS analysis by using both top-down and bottom-up approaches. In the first approach experiments were performed on entire peptides and proteins present in >30KDa and <30KDa fractions, in the second one, tryptic digests of these fractions were prepared and analyzed by LC-MS/MS. Eluents were: (eluent A) 0.1% aqueous formic acid solution and (eluent B) 0.1% aqueous formic acid solution in acetonitrile/water 80:20 v/v. The applied gradient was from 5 to 55% B in 40 min, from 55 to 100% B in 2 min, from 100 to 5% B in 2 min for a total acquisition time of 61 min at a flow rate of 80 L/min. The MS spectra were acquired in data-dependent mode in them/z range from 150 to 2000. For each MS scan, the three most intense multiply charged ions were selected and fragmented by collision induced dissociation (35% normalized collision energy). MS and MS/MS scans were acquired at a resolution of 60000. Alternatively, fragmentation was carried out using the same conditions on selected multiply charged ions corresponding to specific protein masses. The capillary temperature was set to 250°C, source voltage 4 kV**.**

Data were generated by Xcalibur 2.2 SP1.48 (Thermo Fisher Scientific) using default parameters of the Xtract program for the deconvolution. MS/MS data were analyzed by the Proteome Discoverer software (version 1.4.1.14, Thermo Fisher Scientific), based on SEQUEST HT cluster as search engine against UniProtKB/Swiss-Prot human database (released on 7rd of September 2016, 2,365,638 entries). For peptide matching the limits were Xcorr scores greater than 1.5 for singly charged ions, 2.0 and 2.5 for doubly and triply charged ions, respectively. Furthermore, the cleavage specificity was set to trypsin with two missed cleavages in the bottom-up analysis. Precursor mass search tolerance was 10 ppm and fragment mass tolerance 0.8 Da. The following modifications were searched: phosphorylation, acetylation, oxidation of methionine residues. Peptide sequences and sites

of covalent modifications were also validated by manual inspection of the deconvoluted fragmentation spectra. For some proteins the manual analysis was the only feasible, and the identification was obtain by using Protein BLAST software (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and MS-Product software available at the ProteinProspector website (http://prospector.ucsf.edu/prospector/msh-ome.htm).

3.7 *Trypsin digestion*

Aliquots of 20 μL of each of the protein fractions >30 KDa and <30 KDa from *S*, *D*, and *H* tissues of three different cases were submitted to trypsin digestion using the kit "Trypsin Singles Proteomic Grade" (Sigma-Aldrich) according to the manufacturer's instructions. Digestion was stopped after 12 h by acidification with 0.1% formic acid (final concentration), and the solution stored at -80 °C until the analysis by high-resolution HPLC-ESI-MS.

4.0 Results

4.1 *Top-down and Bottom-up characterization of proteins and peptides.*

For the characterization of proteins and peptides, RP-HPLC-high resolution ESI-MS/MS analyses were performed both on intact proteins and peptides and on their tryptic fragments present in the >30KDa and <30KDa fractions. The top-down approach allowed identifying several proteins and peptides, as well as their derivatives from post-translational modifications. Their identification was confirmed by the bottom-up approach, which was mainly useful to characterize proteins with mass value greater than 10000 Da.

Table 2 reports the UniProt-KB codes, the elution times, monoisotopic and average mass values (experimental and theoretical), and the m/z ions utilized to perform the top-down high-resolution MS/MS analysis, of a part of the components characterized, only those considered significant for this study.

Results obtained by the bottom-up approach are summarized in Table 3, where the mass values and the sequence position of tryptic fragments originated from each parent proteins, such as their m/z ions used for the MS/MS analysis, were also reported.

The analysis of MS/MS spectra, performed manually or by the "Proteome Discoverer" software, are reported in the *Supporting information section I* and *II (SIS-I, SIS-II)* that report supplemental Figures showing top-down and bottom-up analysis respectively.

4.2 Thymosyns β

Thymosin β4 and β10 were detected at the elution times expected for these peptides at the chromatographic conditions used, indeed, the two peptides were previously detected and characterized in other tissues and biological fluids by my research group (Cabras et al., 2015). MS/MS analysis performed on several m/z multiply-charged ions of the two peptides confirmed their identification (Table 2), the Figures S1 and S2, in the *SIS-I*, show the results of MS/MS analysis performed on the 1241.13 $[M + 4H]^{4+}$ m/z of Tβ4 and on the 823.26 $[M + 6H]^{6+}$ m/z of Tβ10, respectively. Beyond the N-terminal acetylation, the most common PTM found for thymosins β that occurs after Met1 removal, other modifications were characterize (Table 2): the proteolytic fragments originated by removal of the C-terminal amino acid residues, Tβ4 1-41 (des-Glu-Ser), and Tβ10 1-41 (des-Ile-Ser) (Fig. S3a-b respectively, *SIS-I*); in protein extracts from human intestinal mucosa biopsies, it was

possible to identify Nε-lysine acetylated derivative of Tβ4 diacetylated on Lys16 and Lys25 (Fig. S4, *SIS-I*). Moreover, oxidized species on Met6 of both thymosins β4 and β10 were characterized by high resolution LC-MS/MS analysis (Fig. S5a-b, respectively, *SIS-I*). Other known proteolytic fragments of the two β-thymosins (Cabras et al., 2015), other Nα-lysine acetylated Tβ4 species as well as Nα-lysine acetylated Tβ10 species were searched along the TIC profile, by XIC procedure, but none of these species was detected in our samples. Bottom-up experiments have confirmed the presence of Tβ4 species identifying the unique fragment 21-39 (Table 3 and Fig. S1 *SIS-II)*.

4.3 Pro-thymosyn α and parathymosin.

Table 2 reports the detection of the isoform II of pro-thymosin α missing the first Met residue, the fragment 1-36 of the pro-thymosin α, called Tα11, and the parathymosin missing the initial Met residue. All these components were N-terminally acetylated. The XIC search of other protein species deriving from the pro-thymosin α and parathymosin, detected in other tissue extracts in previous studies (Cabras T expert opinion) were not observed in our samples.

MS/MS analysis of the intact components allowed confirm their identification, as well as the Nα-acetylation, and are reported as supplemental figures (Fig. S6-7, pro-thymosin α and Tα11, S8 for parathymosin, *SIS-I*). Top-down MS/MS analysis of pro-thymosin α and parathymosin were performed on more m/z ions of these proteins in different samples. When the tryptic digest obtained from <30 KDa fraction was analyzed two unique fragment of pro-thymosin α, isof. II, were characterized confirming its identification (Table 3, Fig. S2a-2b, *SIS-II*).

4.4 Ubiquitin.

The characterization of the monomeric ubiquitin in our samples was performed only by top-down experiments performed on the ion $[M+9H]^{9+}$ at m/z 952.08 in two different samples (Fig. S9, *SIS-I*).

Table 2. Swiss-Prot code, elution times, experimental and theoretical monoisotopic and average mass values, m/z of the multi-charged ions of proteins and peptides investigated.

Peptide / Protein (SwissProt code)	Post-translational modifications	Time elution (min)	Exp. (Theor.) Monoisotopic Mass value (Da)	Exp. (Theor.) Average Mass value (Da)	MS/MS top-down analysis m/z and charge
Tβ4 (P62328)	Nα-acetylated	19.3-19.8	4960.49 ± 0.08 (4960.48)	4962.8 ± 0.6 (4963.5)	*1241.13 (+4); 993.11 (+5); 827.76 (+6)*
	Nα-acetylated; Fragm. 1-41	19.6-20.1	4744.43 ± 0.08 (4743.41)	4746.7 ± 0.6 (4747.31)	*792.08 (+6)*
	Nα-acetylated M6-sulfoxide	17.4-18.0	4976.49 ± 0.08 (4976.48)	4978.8 ± 0.6 (4979.46)	*830.42 (+6)*
	Nα-acetylated; Nε-acetylated K16/K25-diacetylated	21.2-21.6	5044.51 ± 0.08 (5044.51)	5046.93 ± 0.6 (5047.58)	*1010.51 (+5); 842.09 (+6)*
Tβ10 (P63313)	Nα-acetylated	20.3-20.6	4933.54 ± 0.08 (4933.52)	4935.9 ± 0.6 (4936.48)	*823.26 (+6) 705.79 (+7)*
	Nα-acetylated; Fragm. 1-41	20.0-20.7	4733.42 ± 0.08 (4733.41)	4735.8 ± 0.6 (4736.28)	*790.24 (+6)*
	Nα-acetylated; M6-sulfoxide	20.3-20.9	4949.53 ± 0.08 (4949.52)	4952.12 ± 0.6 (4952.48)	*825.93 (+6)*
Pro-Tα, Isof. 2 (P06454)	M1 missing; Nα-acetylated;	20.7-21.2	11977.9 ± 0.2 (11977.9)	11984 ± 1 (11984.7)	*922.85 (+13); 857.00 (+14); 999.67 (+12)*
Tα11	Nα-acetylated; Frag. 1-36 of pro-Tα	19.5-20.2	3787.83 ± 0.06 (3787.82)	3789.59 ± 0.5 (3790.02)	*947.96 (+4)*
Parathymosin (P20962)	Nα-acetylated; Fragm. 2-102	20.5-20.7	11434.2 ± 0.2 (11434.2)	11440 ± 1 (11440.8)	*1041.12 (+11); 1272.14 (+9)*
Ubiquitin (P0CG48)		29.5-30.2	8560.63 ± 0.1 (8559.62)	8564.2 ± 1 (8564.8)	*952.08 (+9)*
SH3BP-1 (Q9H299)	M1 missing; Nα-acetylated	38.5-38.9	10342.3 ± 0.2 (10342.241)	10347.5 ± 1 (10348.5)	*1294.41 (+8); 1035.64 (+10).*
FABP1 (P07148)	M1 missing; Nα-acetylated	40.5-40.9	14110.5 ± 0.2 (14110.4)	14118 ± 2 (14119.2)	*1412.95 (10); 1086.97 (+13); 942.10 (+15); 883.47 (+16);*
FABP1 94:T>A variant, (P07148)	M1 missing; Nα-acetylated	40.5-40.9	14080.4 ± 0.2 (14080.4)	14088 ± 2 (14089.2)	1007.32 (+14); 1084.73 (+13); 940.17 (+15); 1409.95 (+10)
Carbonic anhydrase 1, (P00915)	M1 missing; Nα-acetylated	40.3-40.7	--	28778 ± 3 (28781)	

Peptide / Protein (SwissProt code)	Post-translational modifications	Time elution (min)	Exp. (Theor.) Monoisotopic Mass value (Da)	Exp. (Theor.) Average Mass value (Da)	MS/MS top-down analysis m/z and charge
9955 Da protein		29.6-29.9	9950.0 ± 0.1	9955 ± 1	**Pending for characterization**

4.5 SH3BP-1 protein.

The monoisotopic mass value of 10342.3 ± 0.2 Da detected in the range of 38.5-38.9 min, was identified as the SH3 domain-Binding Glutamic acid-Rich-Like protein 3, known also as SH3 domain-binding protein 1 (SH3BP-1) (Table 2), with the expected Nα-acetylation, following the Met1 removal, by mean of the high resolution top-down MS/MS analysis performed on the ions $[M+8H]^{8+}$ and $[M+10H]^{10+}$ at m/z 1294.54, and 1035.93 values, respectively (Fig. S10, *SIS-I*). The identification of SH3BP-1 protein was confirmed by the bottom-up approach with the characterization of three unique tryptic peptides of this protein (Table 3, Fig. S3a-c, *SIS-II*).

4.6 FABP1.

The Fatty Acid-Binding Protein 1 (FABP1) and its natural variant 94: Thr→Ala were characterized in our samples. This protein is known also as Liver-type Fatty Acid-Binding Protein (L-FABP). The determination of the monoisotopic mass values (Table 2), and the high resolution MS/MS analysis performed on both entire proteins (Fig. S10-11, *SIS-I*) were in accordance with the removal of the Met1 residue and the presence of a Nα-acetylation on the Ser residue in the position 2. These were novel modifications, indeed structural information reported in the UniProt-KB data bank indicate a Nα-acetylation on Met1, on the base of the results published by Chan L. and colleagues (Chan et al., 1985). In this paper a partial amino acid sequencing was obtained on the L-FABP purified from liver that individuate the first residue as methionine, while the Nα-acetylation was detected on the translation product of the mRNA L-FABP obtained in a reticulocyte lysate system.
Bottom-up experiments confirmed the identification of the FABP1 (Fig. S4a-d, *SIS-II*).

4.7 Carbonic anhydrase 1

The mass value of 28778 ± 3 Da, determined by low- resolution MS analysis, was attributed to the Carbonic Anhydrase 1 (CA-1), Met1 missing and Nα-acetylated (expected average mass 28781 Da). The characterization was obtained by bottom-up experiments (Table 3), which allowed reaching the 58% of coverage with the ten tryptic fragments sequenced by high-resolution MS/MS analysis (Fig. S5a-l, *SIS-II*).

4.8 9955 Da uncharacterized protein.

In Table 2 is reported the mass values, monoisotopic and average, of one protein detected in the extracts that appeared interesting as regards its level variations in the different kind of colonic tissues investigated, as described in the following section. The characterization of this protein was not feasible, probably due to its low concentrations in the samples, not useful MS/MS spectra was obtained for it.

Table 3. Swiss-Prot code, unique peptides, sequence of fragments, experimental and theoretical monoisotopic [M+H$^+$] value, m/z MS/MS charged ions of proteins and peptides characterized.

Peptide / Protein (SwissProt code)	Unique peptides	Sequence of fragments (position in the pro-protein)	Exp. (Theor.) Monoisotopic [M+H]$^+$ value	Elution time (min.)	MS/MS analysis m/z and charge
Tβ4 (P62328)	1	TETQEKNPLPSKETIEQEK (21-39)	2229.12 ± 0.04 (2229.12)	15.93	*743.71(+3)*
Pro-Tα, Isof. 2 (P06454)	*2*	AAEDDEDDDVDTK (91-101)	1437.55± 0.03 (1437.55)	9.32	*719.28 (+2)*
		AAEDDEDDDVDTKK (91-102)	1565.64 ± 0.03 (1565.64)	7.05	*783.33 (+2)*
SH3BP-1 (Q9H299)		VYSTSVTGSR (6-15)	1056.53 ± 0.01 (1056.53)	14.84	*528.77 (+2)*
	3	IQYQLVDISQDNALRDEMR (33-51)	2307.14 ± 0.04 (2307.14)	25.66	*769.72 (+3)*
		ATPPQIVNGDQYcGDYELFVEAVEQN TLQEFLK (Carbamidomethyl-Cys) (59-91)	3815.80 ± 0.06 (3815.81)	36.10	*1272.60 (+3)*
FABP1 (P07148)		AIGLPEELIQK (21-31)	1210.70 ± 0.01 (1210.70)	25.12	*605.85 (+2)*

Peptide / Protein (SwissProt code)	Unique peptides	Sequence of fragments (position in the pro-protein)	Exp. (Theor.) Monoisotopic $[M+H]^+$ value	Elution time (min.)	MS/MS analysis m/z and charge
		FTITAGSK (50-57)	824.45 ± 0.01 (824.45)	17.92	*412.73 (+3)*
		TVVQLEGDNK (81-90)	1102.58 ± 0.01 (1102.57)	17.30	*551.79 (+2)*
		TVVQLEGDNKLVTTFK (81-97)	1791.99 ± 0.03 (1791.98)	23.81	*598.00 (+3)*
Carbonic anhydrase 1, (P00915)		LYPIANGNNQSPVDIK (20-35)	1742.91 ± 0.03 (1742.91)	21.45	*871.96 (+2)*
		TSETKHDTSLKPISVSYNPATAK (36-58)	2475.27 ± 0.04 (2475.27)	18.20	*825.76 (+3)* *619.57 (+4)*
		HDTSLKPISVSYNPATAK (41-58)	1929.01 ± 0.04 (1929.01)	18.95	*643.67 (+3)*
		EIINVGHSFHVNFEDNDNR (59-77)	2256.04 ± 0.04 (2256.04)	24.04	*752.68 (+3)*
		GGPFSDSYR (82-90)	985.43956 985.4374	18.65	493.22342
	10	VLDALQAIK (161-169)	970.59496 970.5931	23.37	485.80112
		RAPFTNFDPSTLLPSSLDFWTYPGSLT HPPLYESVTWIIcK (Carbamidomethyl-Cys) (174-214)	4754.36098 4754.3639	36.65	*1189.35 (+4)*
		ESISVSSEQLAQFR (215-228)	1580.79 ± 0.03 (1580.79)	23.41	*790.90 (+2)*
		SLLSNVEGDNAVPMQHNNRPTQPLK (229-253)	2759.39 ± 0.05 (2759.39)	22.31	920.47 (+3)
		YSSLAEAASK (229-238)	1026.51116 1026.5102	16.49	513.75922

4.9 *Quantification of peptides and proteins in tumoral and healthy intestinal mucosa.*

We have optimized the procedure of protein tissue extraction in order to have the highest yield of thymosins β4 and 10. Through HPLC-ESI-MS analysis, tissue protein extracts obtained with different procedures (as reported in paragraphs 3.3a and 3.3b) from four patients have been compared (cases #3, #4, #5 and #6). For each of these patients were prepared protein extracts from tumor tissues (deep and superficial) and healthy colon mucosa, and for each of tissue samples were performed two different protein extraction procedures based on ultrafiltration and hydro-organic treatment.

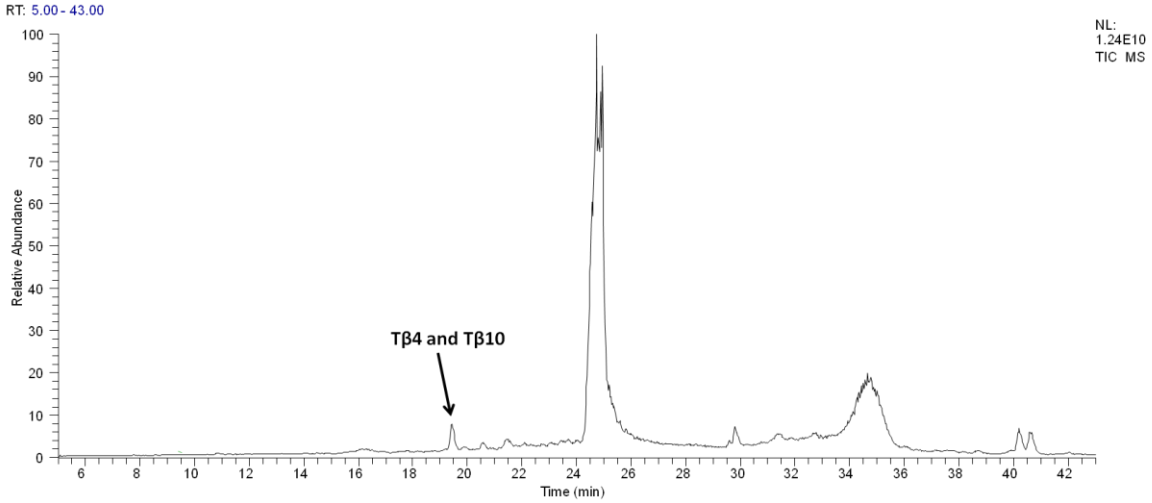
From the ultrafiltration of tissue samples (three, *S, D, H,* for each case, 12 in total), two protein fraction were obtained: <30KDa and >30KDa. Despite various washing and filtration cycles have been made, HPLC-ESI-MS analysis of the <30 KDa and >30KDa protein fractions showed that the Tβ4 and Tβ10 were retained by the ultrafiltration membrane (>30KDa) (Fig. 4, panel A and B), even if their concentration in the fractions >30KDa were lower than in <30KDa fractions, the difference was statistically significant for both Tβ4 (p value = 0.007) and Tβ10 (p value = 0.01). The statistical analysis were made comparing all the <30KDa protein fractions with all the >30KDa fractions (Fig. 5).

Modified species of both thymosins were observed roughly in the <30 KDa or in the >30KDa fractions, not differences were determined in their levels. When the total Tβ4 (Tβ4 + its modified species) and total Tβ10 (Tβ10 + its modified species) were compared in the two protein fraction series a results similar to that described before was obtained: concentration of total Tβ4 and Tβ10 were more abundant in the low-molecular weight fraction (p values = 0.007 and 0.01, respectively). According these results, for the subsequent quantitative comparison (paragraph 4.10) it was decided to sum the areas of XIC peak of the peptide measured in the two protein fractions obtained from the same tissue, from this total XIC peak area it was calculated the peptide concentration.

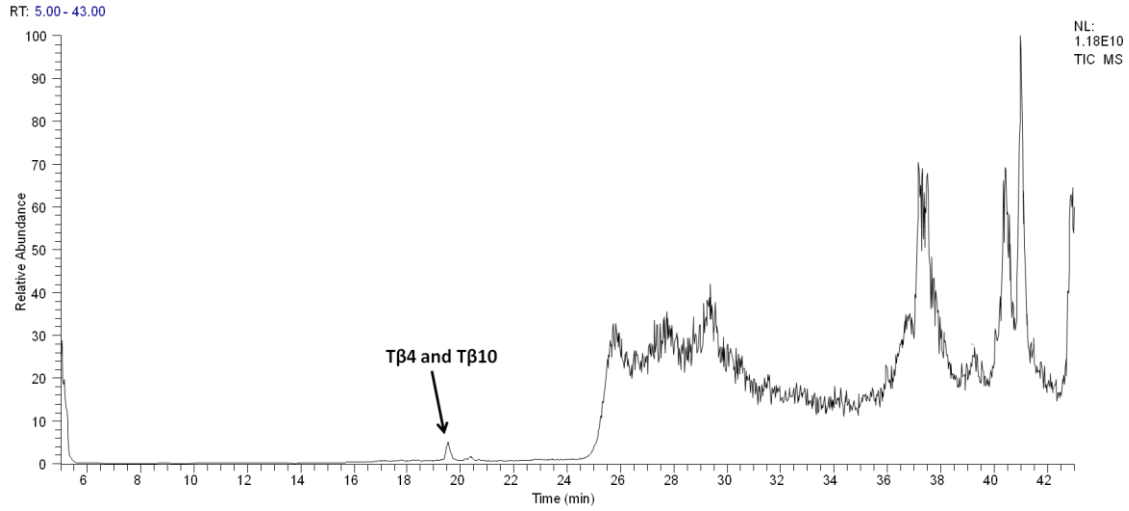
Comparing the results obtained from the parallel treatment (ultrafiltration and hydro-organic treatment) it was evident that the treatment with TFA 0.05%/20% ACN, despite the advantage of being simple and fast, it does not bring in solution the same concentration of Tβ4 and Tβ10 obtained by the ultrafiltration (Fig. 4C). To evaluate the total quantity of Tβ4 and Tβ10 recovered by the two different treatment, the nanomoles were calculated in all the 36 samples and the comparison was performed between hydro-organic extracts and the corresponding ultrafiltration fractions (<30KDa + >30KDa).

Fig. 4. TIC profiles of the fraction <30 KDa (panel A), the fraction >30KDa (panel B) and the acid extract (panel C).

A



B



C

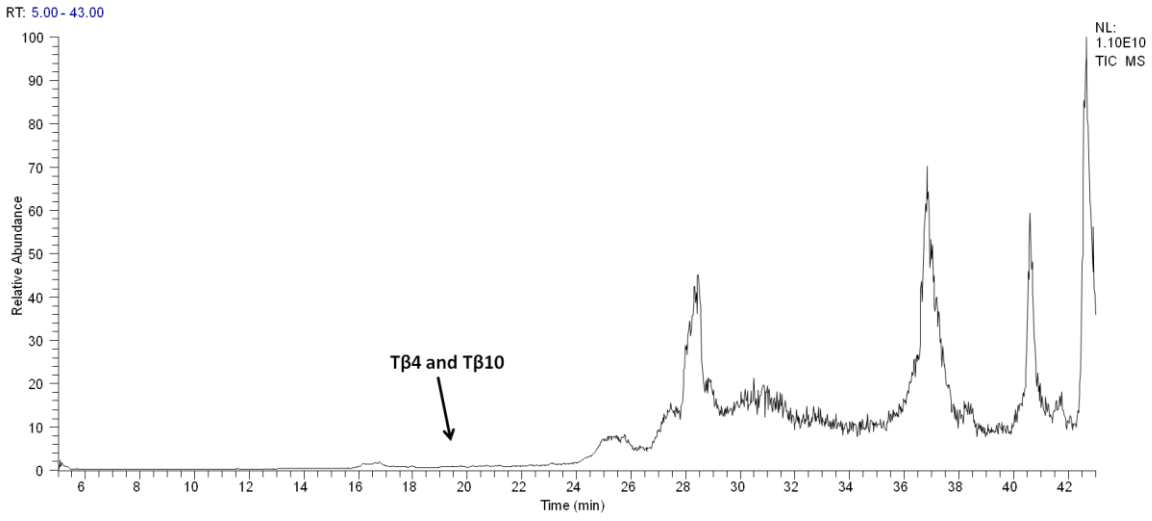
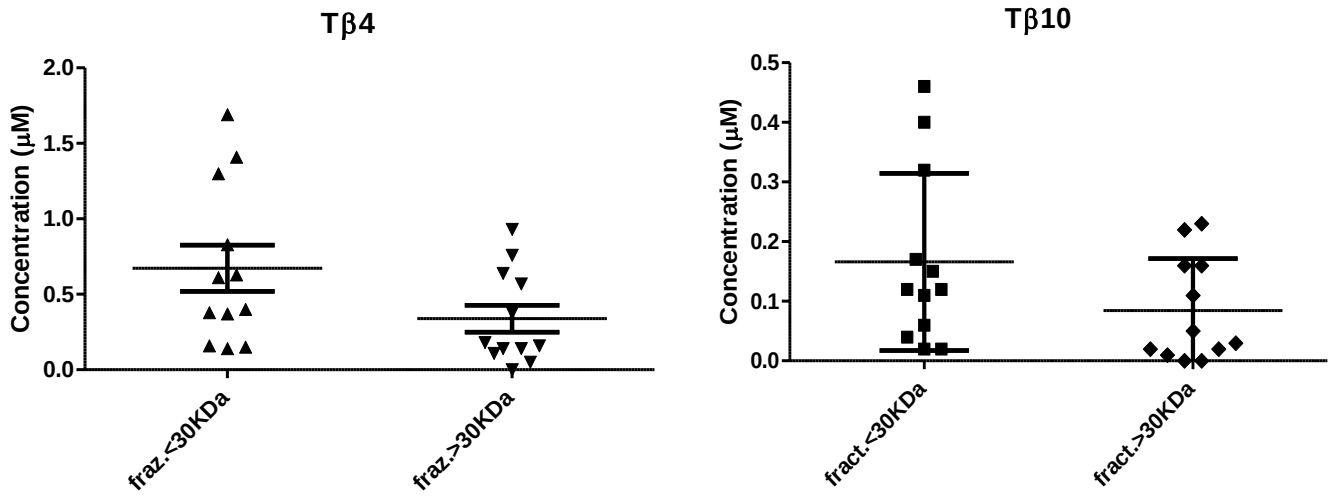
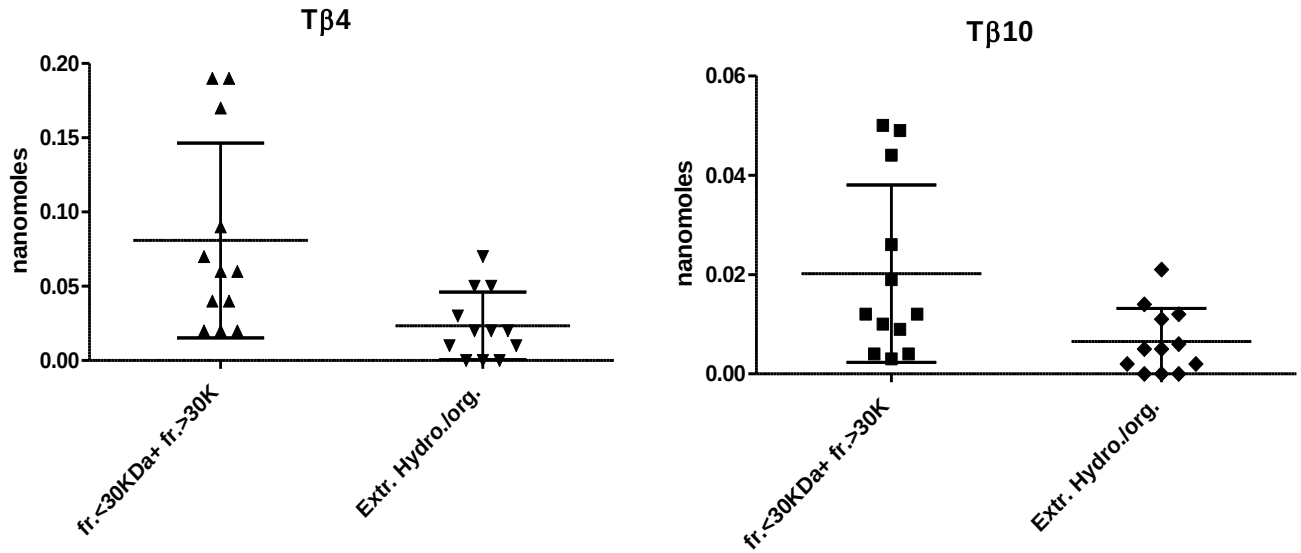


Fig. 5. Comparison between the <30KDa and >30KDa fractions for Tβ4 and Tβ10



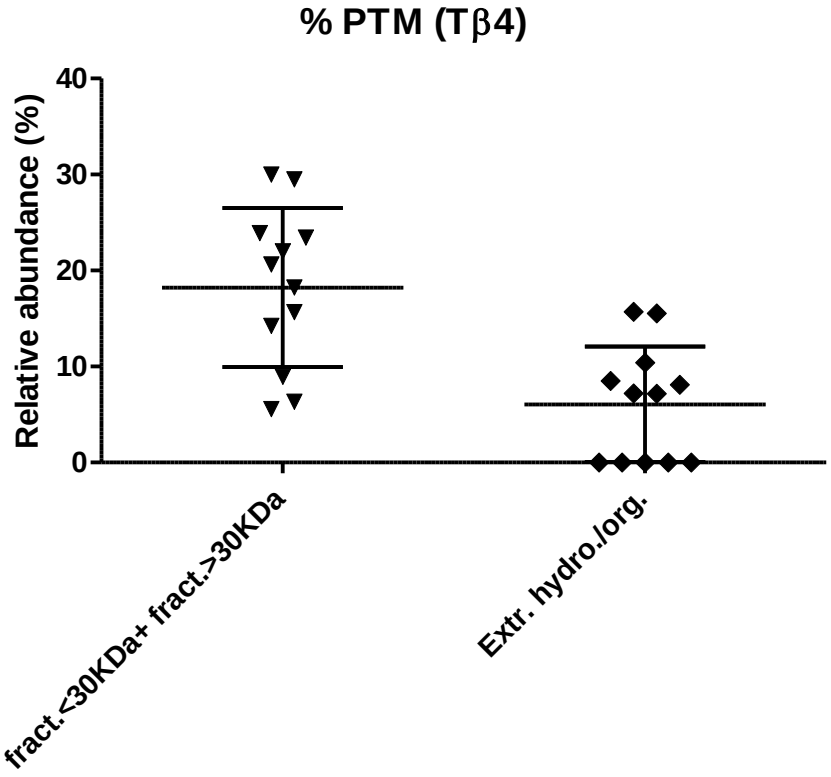
Statistical analysis highlighted a more abundance of thymosins β4 and β10 in the protein extracts obtained with ultrafiltration than in those obtained by TFA 0.05%/20% CAN treatment (Fig. 6). p values were 0.02 for both peptides. Thus, the first extraction procedure was applied for all the samples collected in this study.

Fig. 6. Comparison between the <30KDa and >30KDa fractions and hydro-organic extract for Tβ4 and Tβ10



The relative abundance of the total modified species, calculated as percentage (% PTM sum/ PTM + Tβ4(10)) resulted higher in samples treated with ultrafiltration, as shown in Fig. 7, which reports the comparison between Tβ4 modified species % obtained with ultrafiltration and with hydro/organic extraction. p value was 0.0005.

Fig. 7. Relative abundance (%) of total PTM of Tβ4 in the protein extracts obtained by the two extraction procedures.



4.10 *Comparison between superficial, deep tumor and healthy mucosa.*

From each of the patients, tissue pieces of superficial and deep tumor, as well as healthy colonic mucosa have been provided, with the exception of the case #1, for which the healthy mucosa has not been provided. Four of the 18 patients were affected by not malignant tumor (Table 1) and thus not included in the statistical analysis.

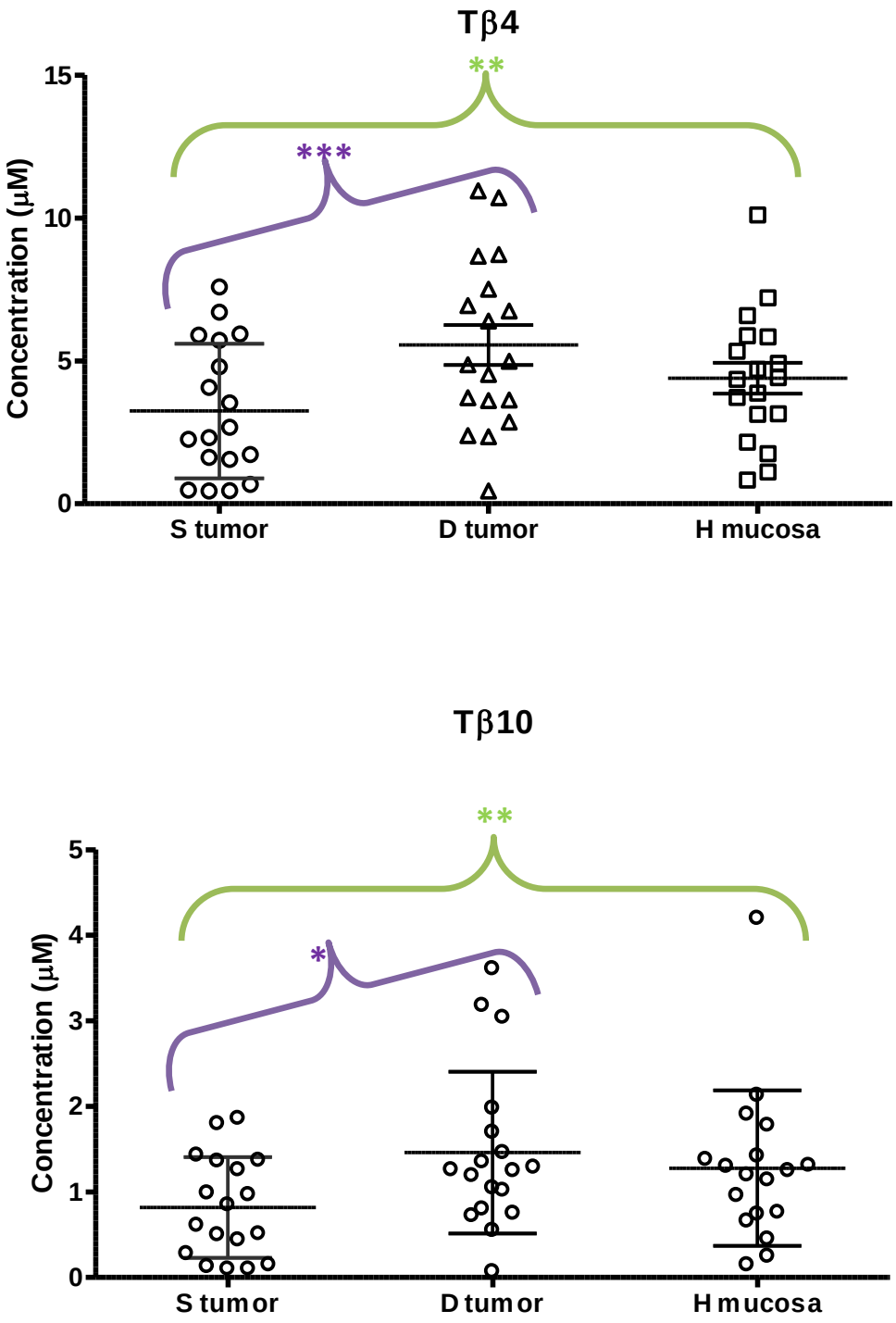
As shown in Table 4 and in Fig. 8 the lower concentration of Tβ4 is measured in *S* tumor, both with respect to *D* tumor ($p = 0.0004$), and to *H* mucosa ($p = 0.03$).

The concentration of Tβ4 is slightly higher in *D* than the *H* tissue, although not significantly ($p = 0.08$). Similarity, the lower concentration of Tβ10 is measured in *S* tissue both respect to *D* ($p = 0.02$) and to *H* tissue ($p = 0.01$). None differences are found between *D* and *H* tissues. The difference in Tβ10 concentration between superficial tumor and healthy mucosa remained significant ($p = 0.013$) also after exclusion of the highest point present in the H group (Fig. 8).

Table 4. UniProt-KB code, concentration (mean and standard deviation) of Tβ4 and Tβ10, and p value obtained by comparing S, D and H tissues.

Peptide (UniProt-KB code)	**Concentration (μM) Mean ± SD**			**p value**		
	S	**D**	**H**	**D vs S**	**D vs H**	**S vs H**
Tβ4 (P62328)	3.25 ± 2.35	5.55 ± 2.97	4.39 ± 2.30	0.0004↑D	NS	0.03↑H
Tβ10 (P63313)	0.82 ± 0.59	1.46 ± 0.95	1.28 ± 0.91	0.02	NS	0.01

Fig. 8. Distribution of concentrations (μM)) of Tβ4 and Tβ10 peptides in S, D and H tissues.



4.11 *Tβ4 and Tβ10 modified proteoforms*

During our analysis, four Tβ4 modified proteoforms are found:

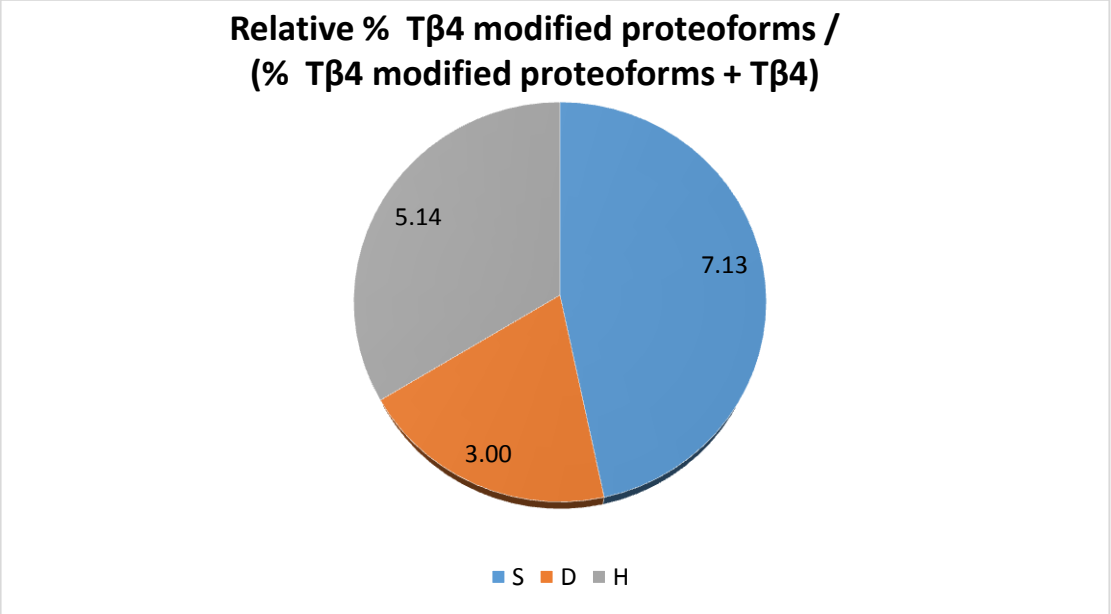
- Tβ4 Fragment 1-41
- Tβ4 M6-sulfoxide (Tβ4 M6-sulfox)
- Tβ4 K16/K25-diacetylated (Tβ4 2K-acet)

Among these, only Tβ4 Met6-sulfox shows a significant variation in concentration between S tumor and H mucosa ($p = 0.03$) according to Tβ4 trend.

Considering the total concentration of all Tβ4 modified proteoforms not significant differences are observed among the three type of tissue.

Moreover, considering the relative percentage of all Tβ4 modified proteoforms respect to total (Tβ4 + Tβ4 modified proteoforms) the lowest percentage was observed in the deep tumor (S vs D $p= 0.004$ and H vs D $p= 0.02$) (Fig. 9), due to major concentration of unmodified Tβ4 in D tissue.

Fig. 9. Relative percentage of Tβ4 modified proteoforms. The percentage was calculated by considering all the modified species of Tβ4 with respect the total Tβ4.



Two Tβ10 modified proteoforms are found: Tβ10 fragment 1-41 and Tβ10 M6-sulfoxide (Tβ10 M6-sulfox). Among these, only Tβ10 fragment 1-41 shows a significant decreased concentration in S tumor with respect the deep one ($p = 0.04$), according to Tβ10 trend.

Considering the total concentration of all Tβ10 modified forms, a significant lower concentration of modified forms in the S tumor tissue was measured ($p = 0.03$ comparing S and H).

Differently to Tβ4, considering the relative percentage of all Tβ10 modified proteoforms respect to total (Tβ10 + Tβ10 modified proteoforms) not significant differences between various types of tissue were observed, since the concentration of Tβ10 modified proteoforms changes in the same way of the unmodified peptide.

4.12 *Other peptides and proteins detected*

Other peptides and proteins were detected in the samples analyzed, 59 components, of which we measured the mass value, among them, those observed in at least 40% of the samples were quantified by XIC procedure. The structural characterization was performed only on components with interesting variation in their level among the different kind of tissues compared. The 8 identified components were: Pro-Thymosin α, Tα11; Parathymosin; Ubiquitin; SH3BP1; FABP1 and its 94:T>A variant; Carbonic anhydrase 1 (Tables 2-3).

Fig. 10 shows the chromatographic TIC profile obtained by HPLC-low-resolution ESI-MS analysis of the <30KDa protein fraction from S tissue, the eluition time of the peptides/proteins of interest are indicated.

As β-thymosins, other proteins and peptides were detectable both in <30KDa fraction and in >30KDa fraction, instead, some proteins was distributed only in the <30KDa fraction whereas others only in the >30KDa fraction. Pro-thymosin α and parathymosin were detectable only in the >30KDa fraction (Fig. 11 panel A) and not in <30KDa fraction (Fig. 11, panel B). Tα11 peptide was observed only in the low molecular weight fraction. Ubiquitin and SH3BP1 were highly concentrated in the <30KDa fraction and only sporadically present in the high molecular weight fraction, CA1 and FABP1 were detected in both protein fractions.

In reason of that, for our statistical analysis we considered the sum of the XIC peak areas measured in both fractions (<30KDa and >30KDa) obtained from the same tissue, this value was corrected on the base of total protein concentration measured in the initial step of protein extraction.

Fig. 10. TIC profile of protein fraction <30KDa of H tissue with principal proteins/peptides detected

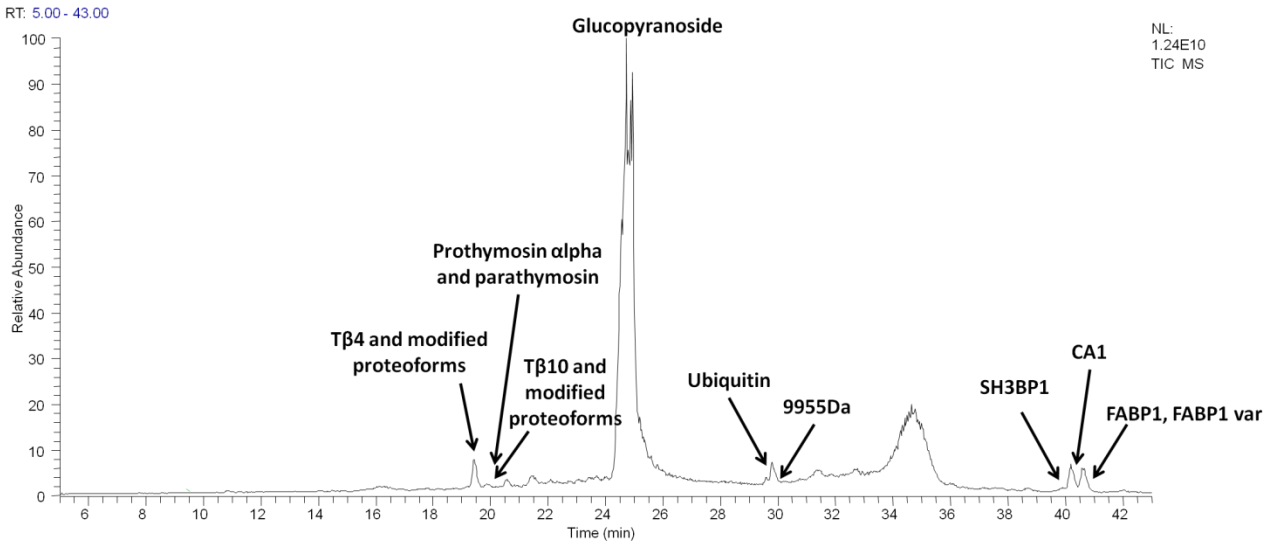
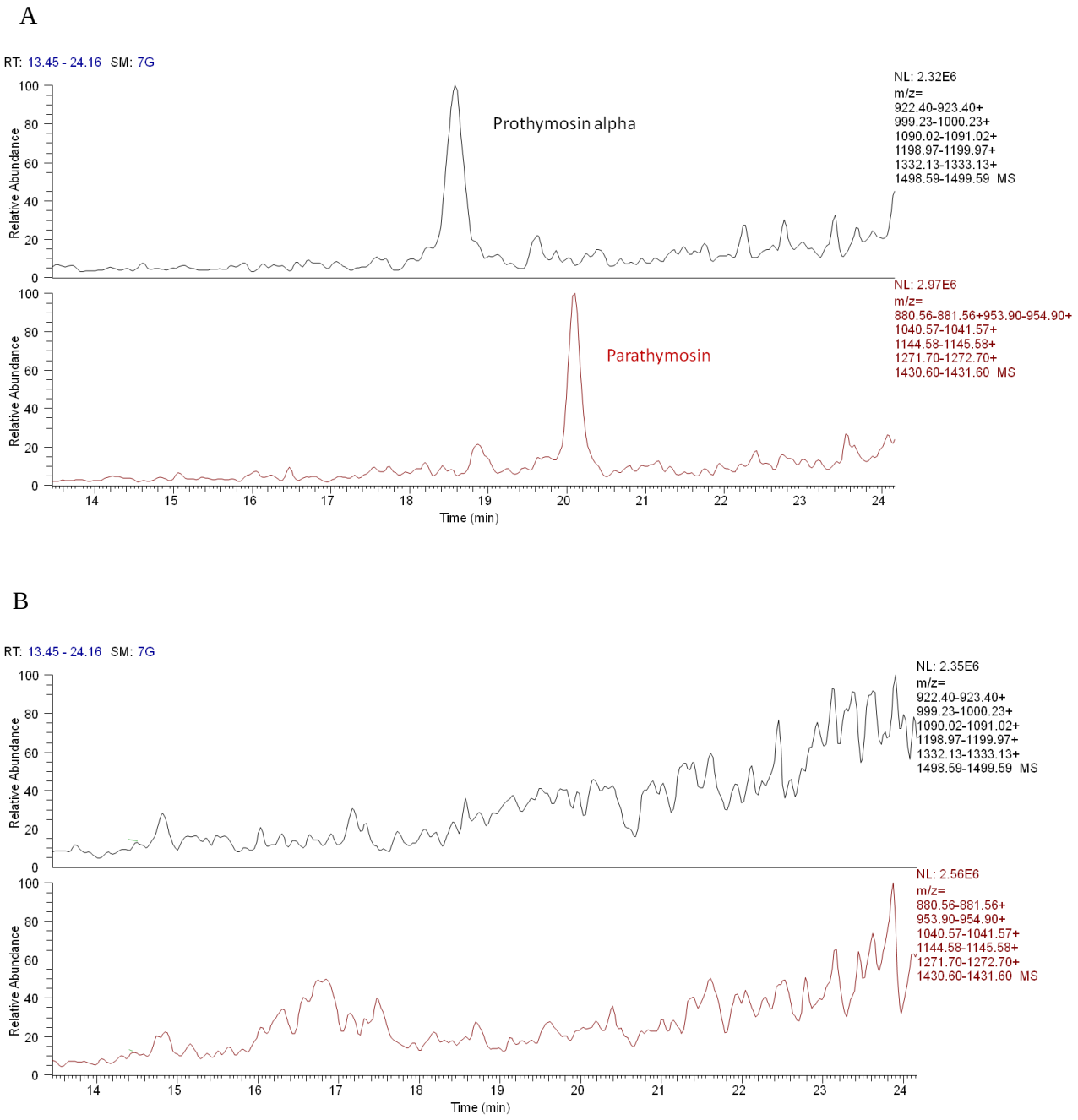


Fig. 11. XIC peaks of pro-thymosin α and parathymosin in >30KDa fraction (A), and in the <30 KDa protein fraction (B).



The results of the statistical analysis performed comparing between them the XIC peak area values of peptides/proteins measured in the different type of tissues, are reported in the Table 5.

Similarly with thymosins β4 and 10, Pro-Tα was detected with high abundance in the deep tumoral tissue (Table 5 and Fig. 12A), its fragment, Tα11, instead was significantly more concentrated in healthy mucosa than in the tumoral tissues, particularly respect to the superficial tumor (Table 5). Differently to the pro-Tα, the parathymosin, which was always detected together the first, did not show significant differences, even if its level in the S tumor was basically lower than in the other two kind of tissues (Table 5).

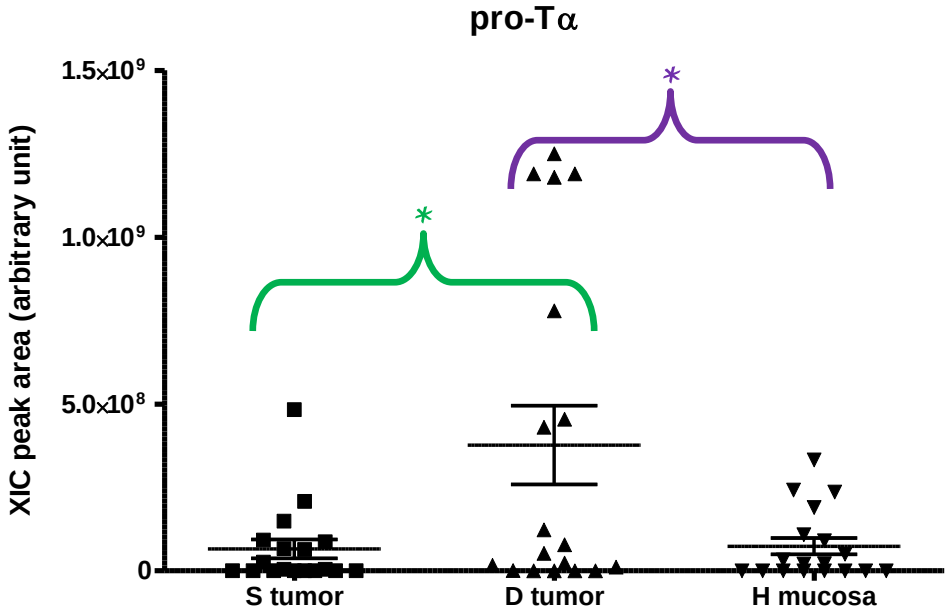
Ubiquitin concentration exhibited the same trend of thymosins β4 and 10, and pro-Tα being more abundant in *D* tissue respect to *S* ($p = 0.006$) tissue and *H* mucosa ($p = 0.007$) as shown in Table 5 and in Fig. 12B.

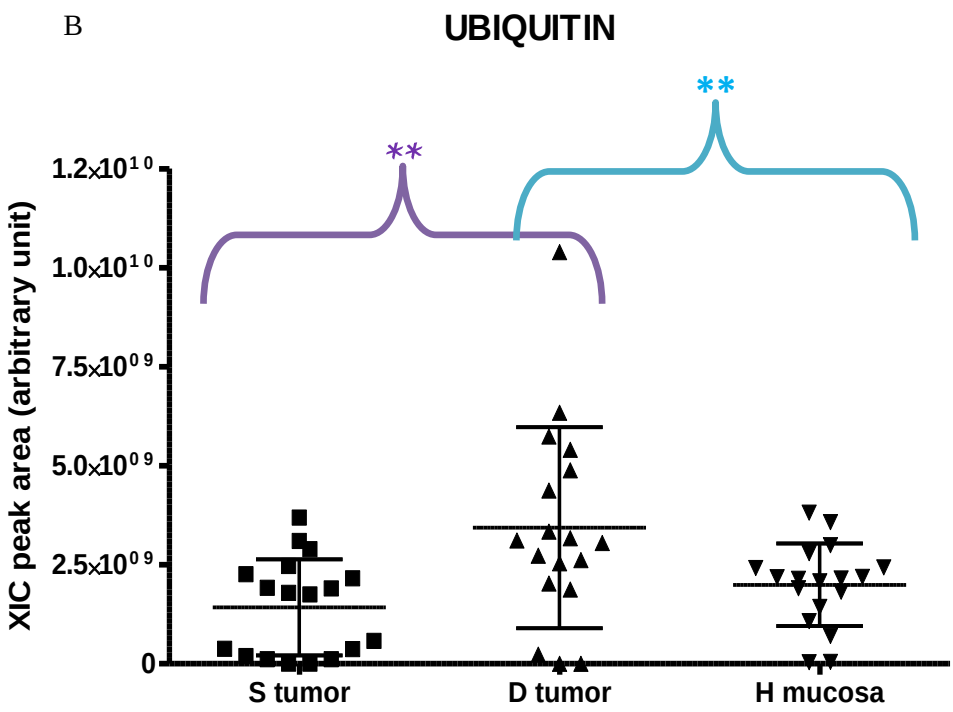
Table 5. XIC peak areas (mean and standard deviation) of proteins and peptides quantified, and p value obtained by statistical comparison between S, D and H tissues.

Peptide / Protein	XIC peak area (10^8) Mean ± SD			p value		
	S	D	H	D vs S	D vs H	S vs H
Pro-Tα, Isof. 2	0.65 ± 1.20	3.8 ± 5.0	0.73 ± 1.05	0.02↑D	0.02↑D	NS
Tα11	0.17 ± 0.41	0.28 ± 0.53	0.63 ± 1.14	NS	NS	0.04↑H
Parathymosin	0.06 ± 0.15	0.83 ± 2.35	0.88 ± 2.55	NS	NS	NS
Ubiquitin	14.2 ± 12.1	34.3 ± 25.4	19.6 ± 10.4	0.006↑D	0.007↑D	NS
SH3BP-1	0.50 ± 1.01	1.79 ± 1.41	2.30 ± 1.58	0.02↑D	NS	0.0001↑H
FABP1	11.0 ± 14.3	27.2 ± 69.2	56.7 ± 43.4	NS	0.0005↑H	0.0003↑H
FABP1 94:T>A variant	2.9 ± 11.8	1.4 ± 3.0	14.3 ± 33.8	NS	NS	NS
Carbonic anhydrase 1	16.8 ± 23.5	8.8 ± 14.5	41.0 ± 45.4	NS	0.01↑H	NS
9955 Da protein	1.4 ± 2.93	10.0 ± 23	4.5 ± 5.3	NS	NS	0.007↑H

Fig. 12. XIC peak area distribution of Pro-Tα and Ubiquitin in S, D and H mucosa

A





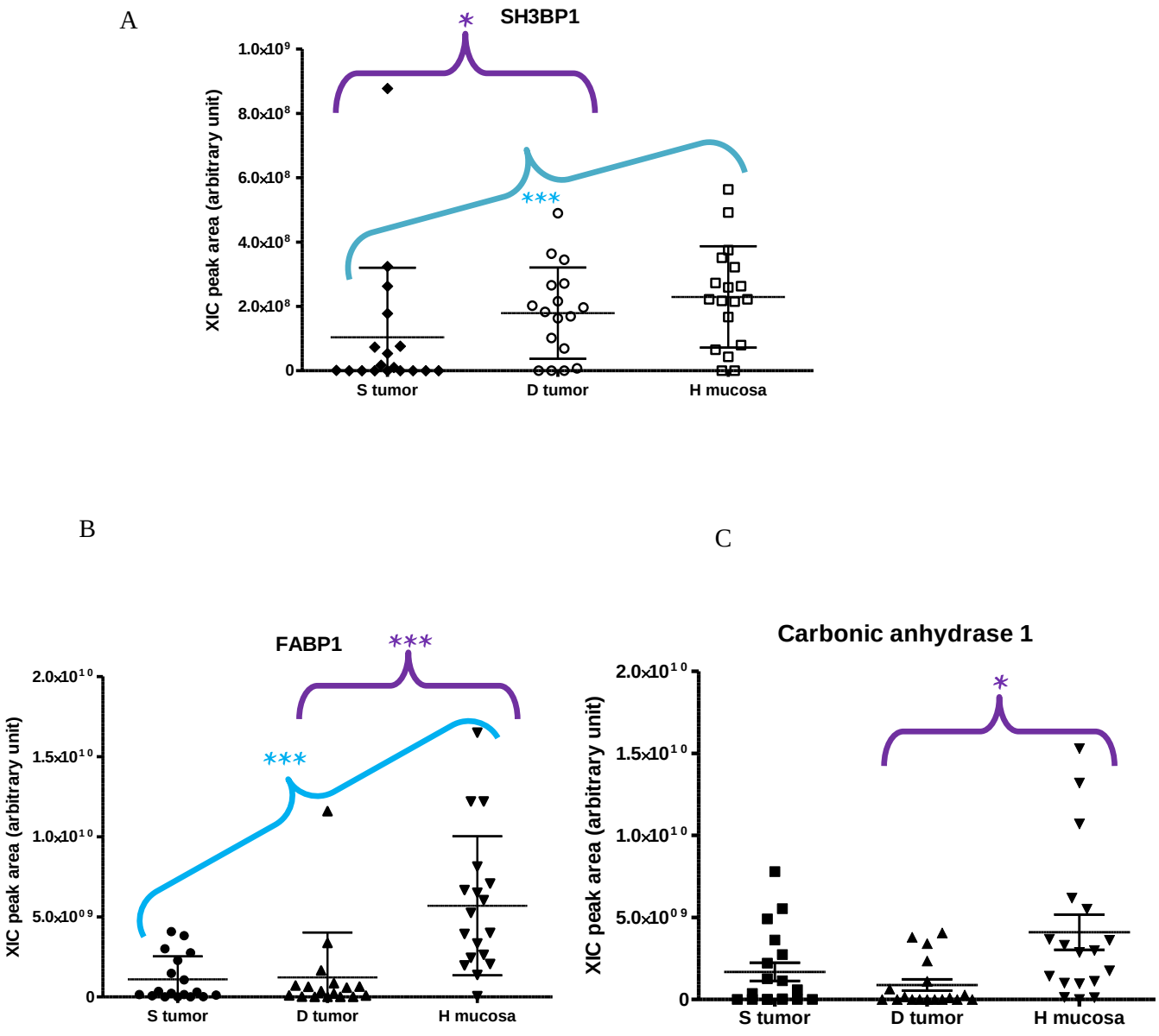
SH3BP1 showed the minimum level in the superficial tumor of CRC patients and the maximum level in the health mucosa. S tissue showed a significant lower concentration of SH3BP1 both with respect to D and H tissues (Table 5, Fig. 13A). Not significant difference was obtained by the comparison between healthy mucosa and deep tumor, although H tissue exhibited a higher level of SH3BP1 than D tumoral tissue (Table 5).

FABP1 and its variant were detected in healthy colonic mucosa with significant higher levels than in the tumoral tissues (Table 5). The p value of the comparison between H and D tissues remained significant both including and excluding the highest point present in the group D (Fig. 13B). The 94:T>A variant of FABP1 exhibited basically the same trend of the principal species although did not highlighted significant variations (Table 5).

CA-1, similarly to FABP1, is more abundant in H mucosa in particular with respect to the D tissue, where its concentration is minimal (Fig. 13C), lower also than that one measured in the superficial tumor, even if this difference did not result statistically significant (Table 5).

The 9955 Da uncharacterized protein (Table 2), exhibited a trend similar to this of SH3BP1, with the minimal levels in the superficial tumor tissue with respect to the deep one and to the healthy mucosa (Table 5). Significant difference was obtained by comparing H with S tissue.

Fig. 13. XIC peak area distribution of SH3BP1 (A), FABP1 (B), and CA1 (C) in S, D and H mucosa.



4.13 *Correlation with Dukes stadium*

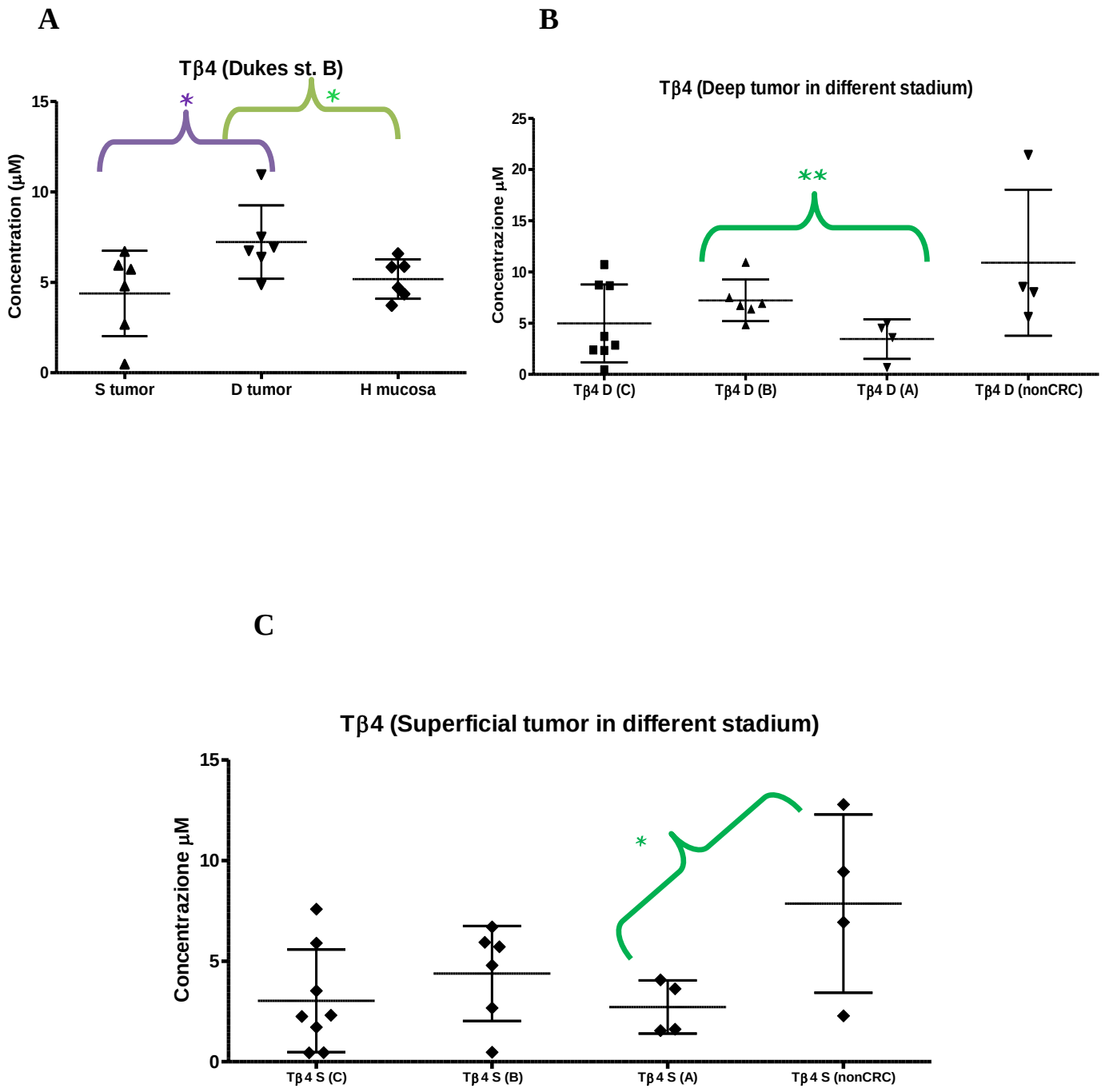
On the base of the degree of tumor infiltration (Dukes Stadium) from the selected CRC cases, 4 were classified as Dukes A, 6 cases were classified as Dukes B and 8 cases were classified as Dukes C. After grouping the patients in relation of the stadium, the comparison of peptides/proteins levels between deep, superficial tumor and healthy mucosa was repeated. Interesting results in relation with Dukes classification of patients were obtained for thymosins β4 and β10, ubiquitin, SH3BP1, CA1, FABP1, and the 9955 Da protein, but for pro-thymosin α, Tα11 and parathymosin.

4.14 Thymosins β4 and β10.

Only in the Dukes B patients was highlighted the significant higher concentration of the Tβ4 in the deep tumor with respect both the superficial one and the healthy mucosa ($p = 0.03$ for both cases) (Fig. 14, panel A). The level of Tβ4 in the D tumor tissue was compared in the different Dukes groups (Fig. 14, panel B), it may be noted that the peptide was more concentrated in the deep tumor of patients with stadium B than patients with stadium A ($p = 0.02$). The group of patients with stadium C was enough inhomogeneous regarding concentration of Tβ4 in the deep tumor. Interestingly, if the D tumor tissue in NON-CRC group is compared with the others (Fig. 14, apnel B), it may be noted that the concentration of Tβ4 appeared basically higher than in the stadium A, which can be individuate as the stadium where the Tβ4 level reached the minimum value in the deep part of tumor.

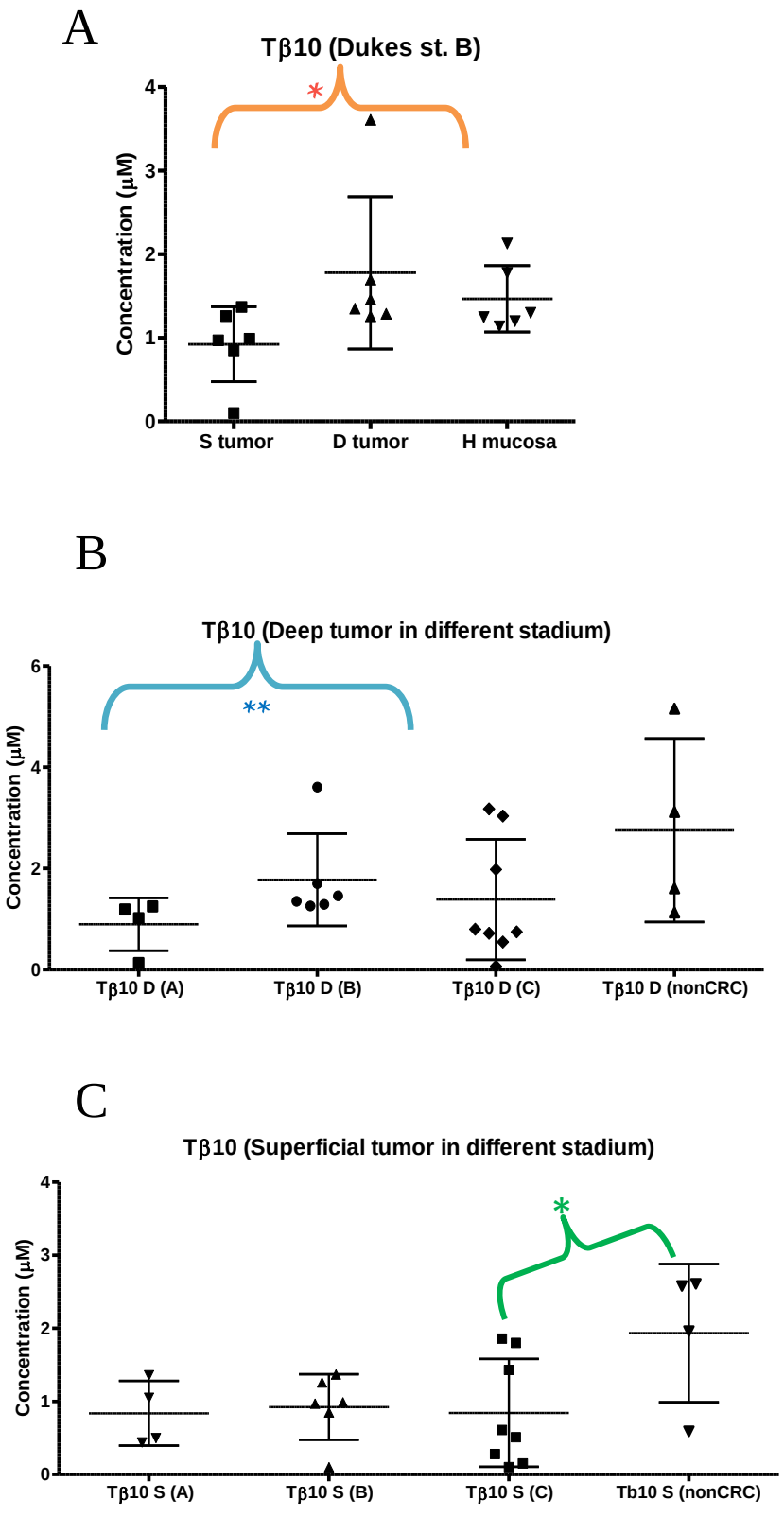
When considering the level of Tβ4 in the superficial tumor tissue (S) of patients according with the stadium, it may be noted that not significant differences were between the different Dukes groups of CRC patients (Fig. 14, panel C). Also in this case, Tβ4 was more concentrated in the S tissue of patients NON-CRC than of patients CRC. A significant difference was obtained from the comparison between NON-CRC group and the Dukes C stadium ($p = 0.03$), this last was individuate as the stadium where the Tβ4 reached the minimum level in the superficial part of tumor. Three of NON-CRC patients, having a low degree adenoma, presented a high Tβ4 concentration, instead the fourth patient with high degree adenoma presented a Tβ4 concentration comparable to that of CRC patients.

Fig. 14. Concentration distribution of Tβ4 in relation to Dukes Stadium. Dukes stadium B in S,D and H (A), Dukes stadium A,B,C comparison in D tissue (B), Dukes stadium A,B,C comparison in S tissue (C).



Patients classified as Dukes B presented a major concentration of Tβ10 in D tissue, significantly different with respect S tissue ($p = 0.02$). The H mucosa present a more high concentration of this peptide respect to S tissue, even if a significant difference was not obtainable. No significant difference between D tumor and H mucosa was found (Fig. 15, panel A). Comparing level of peptide in D tissue according to stadium, it can be observed that Tβ10 is more concentrated in B stadium patients than in A stadium patients ($p = 0.01$) (Fig. 15, panel B). The group of patients with C stadium is inhomogeneous regarding Tβ10 concentration in D tissue. Similarly to Tβ4, the stadium A was that where Tβ10 reached the minimum value also with respect to the NON-CRC patients. Considering the level of Tβ10 in S tumor of the patients according with the stadium, not significant differences were observed between the various stadiums in CRC patients (Fig. 15, panel C). Tβ10 was slightly more concentrated in S tissue of NON-CRC patients than in stadium C patients ($p = 0.048$). Also in this case, NON-CRC patients with high degree adenoma present the lower concentration of Tβ10 than patients with low degree adenoma in S tissue.

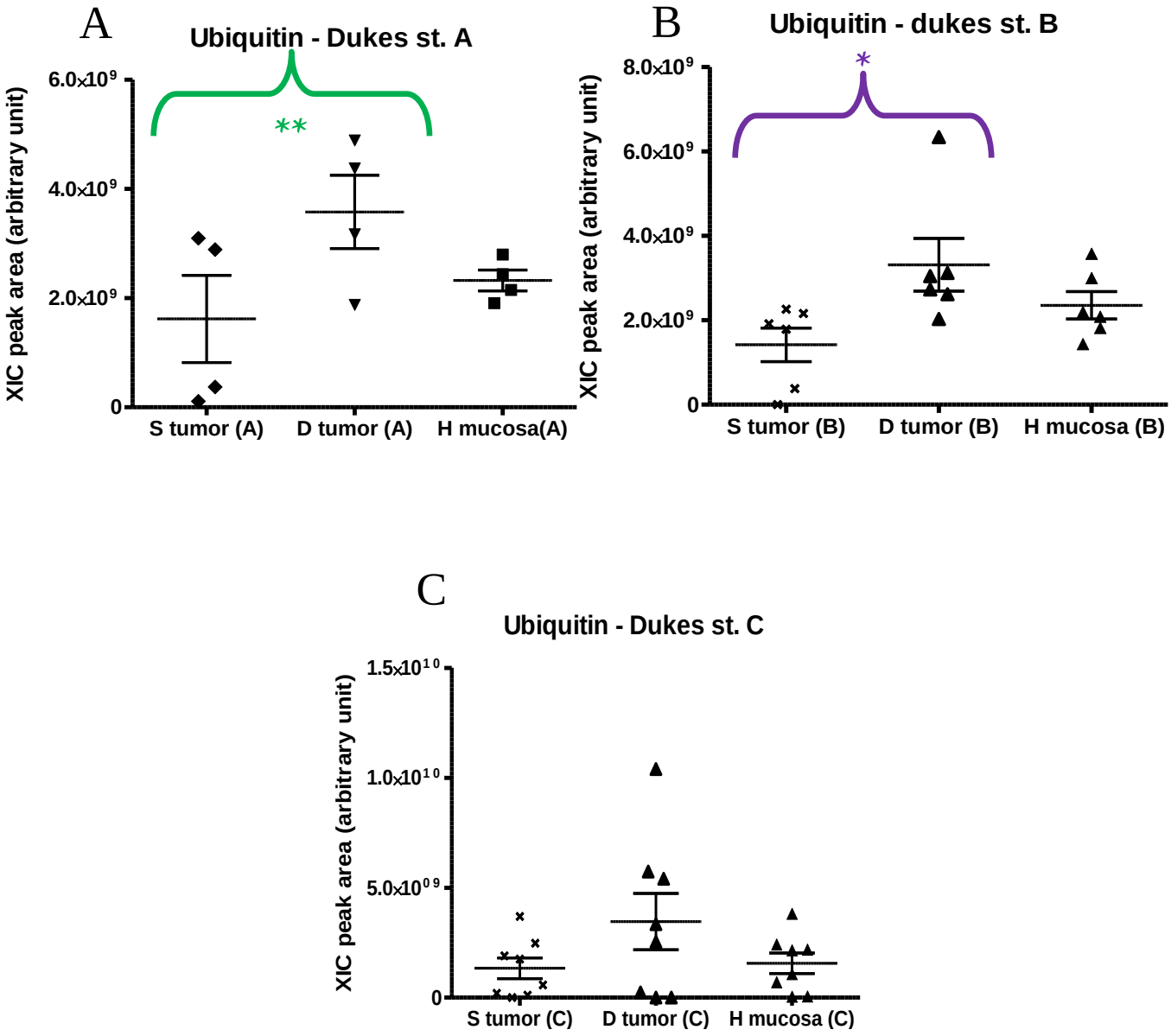
Fig. 15. Concentration distribution of Tβ10 in relation to Dukes Stadium. Dukes stadium B in S,D and H (A), D, Dukes stadium A,B,C comparison in D tissue (B), Dukes stadium A,B,C comparison in S tissue (C).



4.15 Ubiquitin.

The level variations of ubiquitin in relation to kind of tissue and in relation to the Dukes stadium were very similar to that of thymosins β4 and β10. Indeed, it was even more evident, since in all the Dukes stadium, A, B, and C, ubiquitin was more concentrated in deep tumor than in the superficial one (A stadium p = 0.006 D↑; B stadium p = 0.015 D↑) (Fig. 16, panel A and B) in C stadium this difference appeared but not in statistically significant way (Fig. 16, panel C).

Fig. 16. Concentration distribution of Ubiquitin in relation to Dukes Stadium. Dukes stadium A in S,D and H (A), Dukes stadium B in S,D and H (B), Dukes stadium C in S,D and H (C)

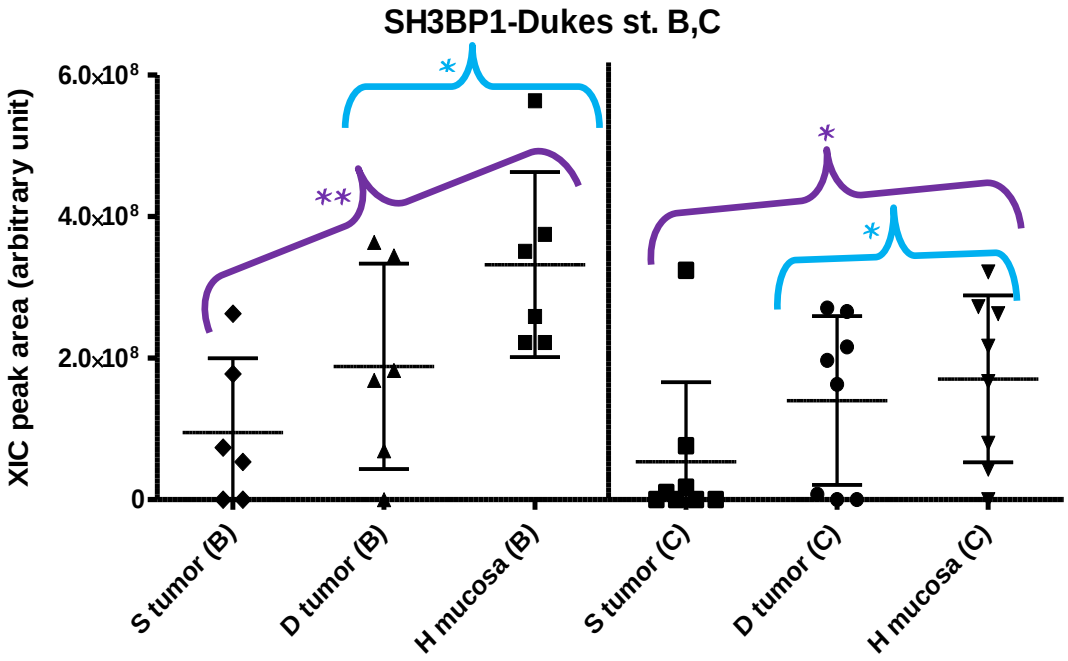


4.22 SH3BP1

The variation observed in SH3BP1 levels, reaching the smallest value in the superficial tumor, is primarily due to patients in B and C stadium. The level of SH3BP1 was higher in healthy mucosa of patients in B stadium than in their tumoral tissues (p = 0.003 H vs S; p = 0.01 H vs D) (Fig. 17, left panel). Analogously, patients in stadium C showed a higher level of the protein in H tissue with respect their S tumoral tissue (p = 0.04). If the highest point present in the S tumor group of stadium C (Fig. 17, right panel) was excluded, the difference between superficial tumor and healthy mucosa was even more evident (p = 0.02), as well as that with the deep tumor (p = 0.04).

The SH3BGRL3 level is higher in NON-CRC than CRC patients in B and C stadium.

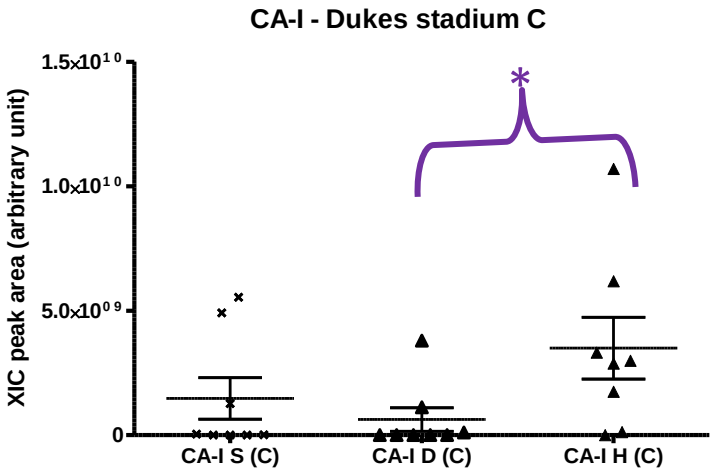
Fig. 17. Concentration distribution of SH3BP1 in relation to Dukes Stadium. Dukes stadium B in S,D and H (left panel), Dukes stadium C in S,D and H (right panel)



4.17 Carbonic Anhydrase 1

The level variation observed for CA1 was due primarily to patients in C stadium. The CA1 level in H tissue of the patients was significantly higher than in their deep tumoral tissue (p = 0.03). Tendentially also patients in B stadium showed the same variations.

Fig. 18. Concentration distribution of CA1 in S, D, and H tissues in patients classified as Dukes Stadium C.

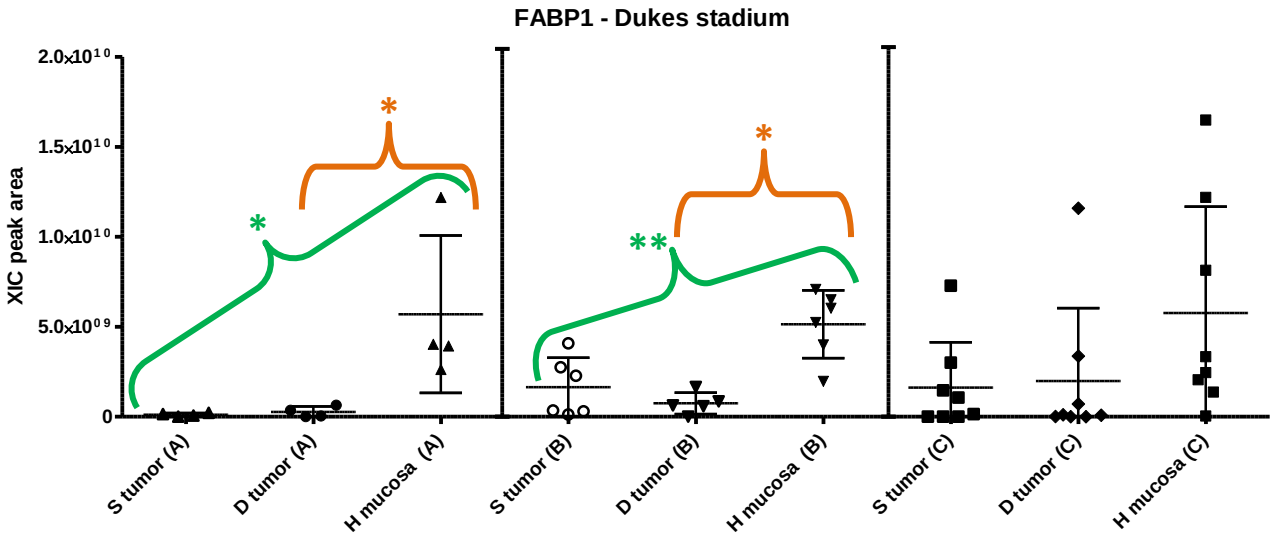


4.18 FABP1

The differences observed in FABP1 levels was due principally to the patients in A and B stadium.

FABP1 levels in H tissue of stadium A patients was more abundant than in their tumoral tissue: p = 0.03 for both H vs S and H vs D (Fig. 19, left panel). This difference has been highlighted also in stadium B patients: p = 0.03, H vs S; p = 0.004 H vs D (Fig. 19, central panel). similar variations, even if not significant can be observed in stadium C patients (Fig. 19, right panel). Moreover, it was observed that in patients in stadium A FABP1 was quite absent in the tumoral tissue, whereas appeared more concentrated in tumoral tissue of patients in stadium B and C. The difference was significant in the comparison of the FABP1 level in superficial tumor of stadium A patients against that one of stadium B patients (p = 0.04).

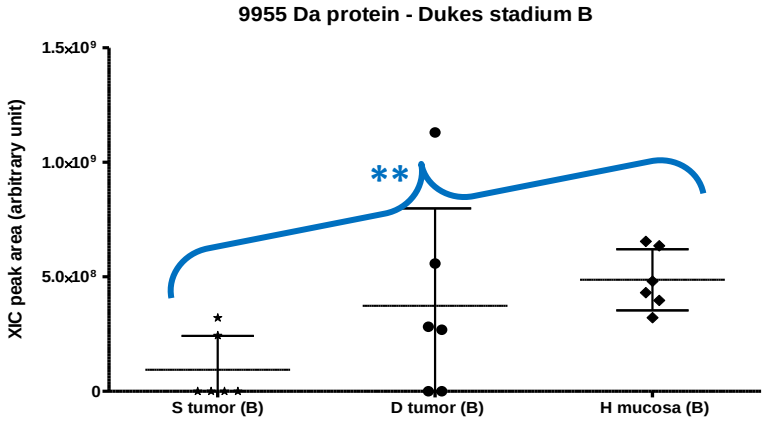
Fig. 19. Concentration distribution of FABP1 in S, D, and H tissues in patients in relation to Dukes Stadium A (left panel), B (central panel), and C (right panel).



4.19 9955 Da protein.

The patients classified as stadium B exhibited the higher levels of the 9955 Da protein in healthy mucosa with respect the superficial tumoral tissue (p = 0.004) (Fig. 20). Similar trend was observed in stadium C patients even if in not significantly way.

Fig. 20. Concentration distribution of CA1 in S, D, and H tissues in patients classified as Dukes Stadium B.



5.0 *Discussion*

The procedure used to prepare the protein extracts from colonic tissues revealed to be optimal to detect thymosins β4 and β10, but also useful to investigate quantitative variations of other peptides and small proteins. The use of ultrafiltration step rather the extraction with acid/organic treatment resulted advantageous in terms of yield, despite it required longer time and it caused a repartition of thymosins and other proteins in two protein fraction (low and high molecular weight). The fact that the thymosins are partly retained by the ultrafiltration membrane has suggested that Tβ4 and Tβ10 may interact in non-covalently mode with large proteins present in the raw protein extract, and form complexes that are blocked by the membrane filter. When the high molecular weight fraction was solubilized with 0.05%TFA/ 20%ACN the probable complexes β-thymosins/high molecular weight proteins are dissolved and these latest probably precipitate freeing the thymosins, as suggest the formation of a pellet following the sample centrifugation. Preliminary bottom-up proteomics analysis performed on these pellets seem to confirm the absence of thymosins (data not reported).

Indeed, it is konwn that Tβ4 is able to interact, as well as with G-actin, fibrin, and collagene, even with several other proteins to modulate their activity (Qiu P, et al., 2011). Several intracellular proteins have been identified as Tβ-partners, like proteins involved in the cellular adhesion (LIM, PINCH-1, ILK) (Bock-Marquette, I., et al., 2004) ; (Bednarek, R., et al., 2008) ; (Lee, S. et al., 2008), and the nuclear protein hMLH1 (Brieger A, et al., 2007). Similar information on Tβ10 are absent, it is desirable a future study based on a deep proteomic analysis of the pellet components to complete the picture of knowledge about the interaction between thymosins β4, and β10, and the high molecular weight components, which could be originated in the intestinal tissues investigated in this study.

The results obtained in the present study strongly suggested that the deep and superficial tumor tissue present different protein profiles, by quantitative point of view, and that the change in protein levels appears to be correlated to the severity of the tumor stage (from adenoma to Dukes' C stadium), especially as regards some peptides/proteins, that could have an important role in CRC carcinogenesis.

On the basis of statistical analysis we can individuate two groups of peptides/proteins:

a) Group 1. Proteins/peptides with the maximum level in *D* tumor and a level in *S* tumor ≤ to that of *H* mucosa:
    - Tβ4;

- Tβ10;
- Ubiquitin.

b) Group 2. Proteins/peptides with the maximum level in *H* mucosa and the minimum level in *S* and/or *D* tumor:
   - FABP1;
   - SH3BP1;
   - CA-I;
   - 9955 Da protein.

5.1 *Thymosin β4 and β10 in colorectal cancer*

In the protein extracts from intestinal mucosa analyzed in this studies it was possible characterized several protein species of Tβ4 and Tβ10, the characterization of the same species were reported in a previous study published by my research group (Cabras et al., 2015). The proteolytic fragments missing the two C-terminal residues amino acid residues, Tβ4 1-41 and Tβ10 1-41, are reported as the most common fragments detected for both peptides (Cabras et al., 2015), and we can consider them as naturally occurring fragments of Tβ4 and Tβ10, since they were observed in samples treated with inhibitor cocktails. Huff et al. speculated that these truncated proteoforms of Tβ4 and Tβ10 might be generated in vivo by a carboxydipeptidase (T Huff, Muller, & Hannappel, 1997). Many other N- and C-terminal proteolytic fragments of Tβ4 have been detected in several biological matrices (Plavina, Hincapie, Wakshull, Subramanyam, & Hancock, 2008), and the production of thymosin β fragments might change according to immunological and tumor diseases. Indeed, the levels of Tβ4 and Tβ10 proteolytic peptides were increased in several malignant tumor histotypes (Hardesty, Kelley, Mi, Low, & Caprioli, 2011), including pediatric brain tumors such as medulloblastoma (high-grade tumor) and pilocytic astrocytoma (low-grade tumor) (Martelli et al., 2015). The majority of thymosins are N-terminally acetylated. Acetylation can also occur also on lysine residues 3, 11, 14, 16, 25, 31 and 38 (Hannappel, 2010). In protein extracts from human intestinal mucosa biopsies, it was possible to identify Nε-lysine acetylated derivative of Tβ4 diacetylated on Lys16 and Lys25. The acetylation on these lysine residues was first detected by Choudhary et al. (Choudhary et al., 2009) in protein extracts from cell cultures. The Met6 oxidation on Tβ10 and Tβ4 was already observed in other protein extracts from cells and tissues, as well as biological fluids (Choudhary et al., 2009).

Modified species of thymosins β4/β10 were detected in very low concentration and their level variation were negligible with respect to that of the unmodified peptides. Quantitative analysis revealed that β-thymosins are more abundant in deep tumor corresponding to the front of tumor invasion, than in the surface of tumor, often defined "old" tumor, because constituted by initial transformed cells. This suggests a potential involvement in metastatization process for these peptides. Furthermore, by performing an analysis on the base of Dukes classification it was possible to note that the B Dukes stadium, corresponding to the intermediate level of tumor, mainly influences the β-thymosins trend that we observe in our patients. According to our results, Nemolato et al. (Sonia Nemolato et al., 2012) have seen that Tβ4 is expressed in the majority of colon cancers, with preferential immunoreactivity in deep tumor regions. The preferential expression of the peptide and the increase in intensity of the immunostaining at the invasion front suggests a possible link between the peptide and the process of epithelial mesenchymal transition, suggesting a role for Tβ4 in colorectal cancer invasion and metastasis.

Epithelial-mesenchymal transition (EMT) is a complex process characterized by the loss of original epithelial features in embryonic and in tumor cells, associated with the gain of a mesenchymal phenotype and producing non-polarized isolated cells embedded in the extracellular matrix (Levayer & Lecuit, 2008). EMT is a key step in cancer progression and permits to tumor cells to acquire an invasive behavior and disseminate (Garber, 2008). Moreover, is a major mechanism by which cancer cells become invasive, penetrating vessel endothelium and entering circulation thus forming metastases (Tse & Kalluri, 2007);(Guarino, Rubino, & Ballabio, 2007). Tβ4 regulate actin polymerization by binding and sequestering monomeric G-actin (Goldstein et al., 2005). Moreover, has been hypothesized that Tβ4 triggers EMT in colorectal carcinoma by upregulating integrin-linked kinase (ILK) (H.-C. Huang et al., 2007). Overexpression of Tβ4 has been shown to upregulate ILK (Bock-Marquette, Saxena, White, Dimaio, & Srivastava, 2004), and consequently to cause the suppression of E-cadherin expression, resulting in disruption of adherents junctions and induction of EMT (Wang et al., 2010).

Tβ4 overexpression is known to be associated with increased invasion and distant metastasis of human colorectal cancer cells (W.-S. et al., 2004). This concept is based on the observation that Tβ4 may facilitate tumor cell motility and induce intra- and peritumoral angiogenesis (Larsson & Holck, 2007). Moreover, the localization of Tβ4 changed during cancer progression, moving from the cell membrane to the Golgi apparatus (Sonia Nemolato et al., 2012).

In a study on β-thymosins expression in hepatocellular carcinoma, both Tβ4 and Tβ10 were detected in tumor cells. Moreover, Tβ10 showed a strong expression in cells undergoing stromal invasion, in contrast with the absence of reactivity for Tβ4 (Theunissen et al., 2014). Tβ10 has recently been recognized as being an important player in the metastatic cascade including tumor angiogenesis, invasion, and metastasis and it also has been described in several cancer: 1) Tβ10 was detected in the majority of the goiters, hyperproliferative cancer tissue, and thyroid adenoma, but not in normal thyroid (Chiappetta et al., 2004); 2) The Tβ10 expression levels correlated significantly with the stage of lung cancer, distant metastases, lymph node metastases, and degree of differentiation of lung cancer (McDoniels-Silvers et al., 2002) ; (Y. J. Lee et al., 2011) ; 3) In breast cancer, Tβ10 staining was detected mainly in the malignant tissues (Verghese-Nikolakaki et al., 1996); 4) The expression of Tβ10 was down-regulated in ovarian cancer compared with normal ovary tissues (S. H. Lee et al., 2001)(Lee et al., 2001). About colorectal cancer still poorly is known about this peptide. Taken together, these data indicate different roles for the two thymosins in different tumor cells, regarding their activity in favoring tumor cell invasion and metastasis.

5.2 *Pro-thymosin α and Parathymosin*

In our study we identified two derivatives of Isoform II of pro-thymosin α (pro-Tα, 109 amino acid residues), corresponding to the N-terminally truncated form missing the first two residues, and the peptide Tα11, which derives from proteolysis of pro-Tα by a lysosomal asparaginyl endopeptidase (Sarandeses, Covelo, Diaz-Jullien, & Freire, 2003). Moreover we characterized the parathymosin (102 amino acid residues) N-terminally acetylated after removal of the Met1 residue. Noticeable amounts of these components were found in bioptic tissues from pediatric brain tumors (Martelli et al., 2015), and were already characterized in intestinal tissues extracts by my group in a previous study (Cabras et al., 2015). The peptide Tα1, the major naturally occurring fragment of pro-thymosin α, was never detected in the protein extracts analyzed.

Similarity to Tβ4 also pro-Tα is more abundant on the front of tumor invasion (deep tumor) and in S tissue, probably having a role in the progress of the cancer. In contrast, Tα11 appears more abundant in H mucosa with respect to S tumor and D tumor, having probably a different mechanism in carcinogenesis respect pro-Tα1. Parathymosin shows no significant change in the three type of tissues.

Pro-thymosin α and parathymosin are two ubiquitous small acidic nuclear proteins that are thought to be involved in cell cycle progression, proliferation, and cell differentiation (Vareli et al., 2000). These proteins were initially isolated from rat thymus and share 40% sequence homology (Haritos, Goodall, & Horecker, 1984) ; (Haritos, Tsolas, & Horecker, 1984). Pro-Tα is a 12.4- kDa protein, is coded by a gene that is located on human chromosome 2 and consists of five exons interrupted by four introns (Szabo et al., 1993). Pro-Tα is a nuclear oncoprotein-transcription factor that possesses multiple functions for cell robustness (Ueda, Matsunaga, & Halder, 2012), it plays its roles both inside and outside the cell, modulating cell cycle progression, stimulating cell proliferation and differentiation, regulating the mechanism of defense and preventing apoptosis. It is reported that it also correlates with c-myc expression (C. G. Wu, Boers, Reitsma, van Deventer, & Chamuleau, 1997). It was overexpressed in several human malignancies (Vareli, Frangou-Lazaridis, van der Kraan, Tsolas, & van Driel, 2000). Elevated pro-thymosin α levels were measured in liver regeneration, colon cancer, and breast cancer as well as hepatocellular carcinoma tissue (Tsitsiloni et al., 1993) ; (S. Q. Wu et al., 1996) ; (Heidecke, Eckert, Schulze-Forster, & Maurer, 1997).

Little information is available for Tα11 functions, but it seems to exert a biological activity very similar to Tα1, which displays immunoregulatory properties (Caldarella et al., 1983).

Parathymosin is an 11.5-kDa acidic protein coded by a gene consisting of five exons, located on human chromosome 17 (Szabo, Clinton, Macera, & Horecker, 1989) ; (Trompeter & Soling, 1992). Compared with pro-Tα, much less is known about parathymosin. Because of their strong sequence homology it is supposed that these proteins may play similar roles.

5.3 *Carbonic Anhydrase*

Carbonic anhydrase (CA) is a zinc enzyme that reversibly catalyzes hydration of CO_2. This reaction is crucial for maintenance of pH homeostasis of the body (Viikila et al., 2016). Sequence studies of human erythrocyte carbonic anhydrase C show that it is a single chain of 259 amino acid residues (Lin & Deutsch, 1974).

Sixteen CA isoenzymes have been described in mammals (discovered isozymes were assigned numerical names in the order of their discovery) and play physiological roles in erythrocytes, including CO_2 transport, ion secretion and pH regulation (Supuran, 2008). Through these chemical reactions they are involved in several downstream physiological

processes, such as bone resorption, vision, and production of saliva, bile, pancreatic juice and gastric juice (Supuran, 2007). Under hypoxic conditions cells produce acidic metabolic products via anaerobic glycolysis. This pathway is inhibited in the presence of enough oxygen. In particular, tumor cells have a tendency to upregulate glucose intake and increase the rate of anaerobic glycolysis even when the amount of oxygen is sufficient (Gatenby & Gillies, 2004). For these reason tumor cells need CA enzymes and many other proteins, such as ion transporters, to maintain physiological intracellular pH (Parks, Chiche, & Pouyssegur, 2013). During this process extracellular pH decreases, which in turn, disturbs physiological processes of the surrounding normal tissue and thus promotes cancer growth (Neri & Supuran, 2011). During the last 20 years, CA proteins have been studied as potential markers for various cancers. Indeed, the changes in pH regulation and altered cellular metabolism are key features of solid tumors (Kazokaité, Ames, Becker, Deitmer, & Matulis, 2016) and CA play an important role in tumor acid-base homeostasis under hypoxic conditions (Wykoff et al., 2000).

Carbonic anhydrase 1 (CA1) is a cytosolic isozymes and it is the most abundant protein in erythrocytes (Sly & Hu, 1995). Uzozie et al. have observed that, measuring expression of CA1 in colon adenomas and normal mucosa, this protein was increased in normal mucosa respect to tumor areas (Uzozie et al., 2014). According with Uzozie and colleagues our results show a major expression of CA1 in H mucosa with respect to tumor tissues. Our results demonstrated that this difference existed with respect both superficial and deep tumor, and that was evident mostly in the patients in Dukes stadium C. Probably, variations in the expression of this protein may be linked to the pH dysregulation in the tumor tissues, especially in the more advanced progression stadium.

5.4 *SH3 domain Binding Protein 1*

SH3BP1 is a small ubiquitous protein encoded by the *SH3BGRL* gene mapping to chromosome 1p36.11. It belongs to SH3BGR family characterized to have a N-terminal region contains a proline-rich sequence (PLPPQIF), which conforms to the SH3 binding motif (PXXP), in addition, SH3BP1 seems to present high similarity with glutaredoxin 1 (Mazzocco M., et al. 2001).

The SH3BP1 presents a C-terminal region highly enriched in glutamic acid residues, which is predicted to assume a largely α-helical conformation associated with loss of heterozygosity in several tumors, as neuroblastoma (Mora et al., 2000) or prostate cancer

(Suarez et al., 2000). Among this functions SH3BP1 is involved in regulation of blood vessel endothelial cell migration and in regulation of actin cytoskeleton organization and actin filament depolymerization (Tata et al., 2014). In addition, SH3BP1 was characterized as a novel downstream effector of Sema3E-PlexinD1 providing an explanation for how extracellular signals are translated into cytoskeletal changes and unique cell behaviour (Tata et al 2014). Majid and colleagues have demonstrated that the expression of SH3BP1, is downregulated in v-Rel-expressing fibroblasts, lymphoid cells, and splenic tumor cells. The v-rel oncogene is the most efficient transforming member of the Rel/NF-kappaB family of transcription factors. v-Rel induces avian and mammalian lymphoid cell tumors and transforms chicken embryo fibroblasts in culture by the aberrant regulation of genes under the control of Rel/NF-kappaB proteins (Majid, Liss, You, & Bose, 2006). Data on SH3BP1 expression in CRC have not been reported so far, the downregulation of SH3BP1 in CRC tissues observed in our study is in according with the study of Majid and colleagues. Moreover, results obtained in this study demonstrated that the levels of SH3BP1 were different between superficial and deep tumor and the minimum concentration of this protein was observed in the surface of the tumor. These variations have been highlighted especially in patients in B and C stadium suggesting that the protein could have a role in the invasion processes of the CRC. Interestingly SH3BP1 is able to modulate polymerization of G-actin like Tβ4, the two components could have complementary in the EMT in colorectal carcinoma.

5.5 *Ubiquitin*

Ubiquitin (Ub) is a small eukaryotic protein consisting of 76 amino acid, that is covalently attached to proteins by the consecutive actions of three distinct enzymes (Hershko & Ciechanover, 1998). Ub is first activated by an Ub-activating enzyme (E1), transferred to an Ub-conjugating enzyme (E2), and then attached to a target protein under the control of an Ub ligase (E3) (Weissman, 2001). The Ub system is extremely versatile and can play multiple essential roles in various cellular processes by regulating of protein stability, protein interactions, trafficking, and activation. Alterations in the Ub system have been observed in many types of human cancers and many of its components, when deregulated, have been found to play key roles in cellular processes relevant to tumorigenesis (Ciechanover & Schwartz, 2004). Moreover, elevated level of Ub has been observed in most cancer cells (Y Ishibashi et al., 1991); (Kanayama et al., 1991); (Yoshio Ishibashi et al.,

2004); (De Méndez Morelva & Antonio, 2009). Because degradation of proteins such as cyclin, p53 and p27 (tumor suppressor proteins) is Ub-dependent (Glotzer, Murray, & Kirschner, 1991); (Maki, Huibregtse, & Howley, 1996) an involvement of Ub in cellular carcinogenesis has been suggested. Cancers exhibit various stress phenotypes, including proteotoxic stress (Luo, Solimini, & Elledge, 2009) ; (Nickolay Neznanov, 2011). Ub itself is a stress-inducible protein (Finley, Özkaynak, & Varshavsky, 1987) and increased of this protein is likely to support the ability of cancer cells to overcome escalating cellular stresses. So, elevated level of Ub during tumorigenesis becomes essential and is maintained for the survival and proliferation of the cancer cells. In our results Ub is more concentrated in D tumor (in the invasion front of the tumor) with respect the surface of tumor and the healthy mucosa. The variation in the levels of ubiquitin is very similar to that of thymosins β4 and β10, suggesting an analogous role in trigging the EMT in colorectal carcinoma. The level variation of ubiquitin between the different kind of colonic tissues was very strong in all patients, especially in these in B and A stadium, which is the less advanced, differently to thymosins that showed a correlation mostly with the Dukes stadium B. Ubiquitin appears, thus, to be an early biomarker more efficient than thymosins as regards the stadium classification and according to other previously studies (Oh C.et al, 2013).

5.6 *Fatty acid binding protein*

The fatty acid binding proteins (FABPs) are a group of low molecular weight proteins involved in the intracellular transport of long-chain bioactive fatty acids including linoleic acid and its derivates such as arachidonic acid (Glatz & van der Vusse, 1996). It is an abundant cytosolic protein that regulates lipid transport and metabolism and is required for cholesterol synthesis and metabolism (Chan et al., 1985). The designation of each of these proteins has been derived from the tissue from which it was originally isolated and key members of this group of proteins include liver fatty acid binding protein, intestinal fatty acid binding protein and epidermal fatty acid binding protein (Veerkamp, van Moerkerk, Prinsen, & van Kuppevelt, 1999); (Richieri, Ogata, Zimmerman, Veerkamp, & Kleinfeld, 2000). FABP contains amino acids that are known to possess antioxidant function (Wang et al., 2005), plays an active part in fatty acid-mediated signal transduction pathways and regulation of gene expression (Hertzel & Bernlohr, 1998) and is involved in modulating cell division (Sorof, 1994), cell growth and differentiation (Schroeder et al., 2001) and also by preventing high intracellular fatty acid concentrations, protect cells against the cytotoxic

effects of fatty acids (Glatz, van Nieuwenhoven, Luiken, Schaap, & van der Vusse, 1997). Alterations that occur in individual FABP expression during tumour development and progression (Celis et al., 1996); (Jing et al., 2000) may contribute to tumorigenesis. Additionally, it has been suggested that the expression of individual fatty acid binding proteins in tumours may also serve as useful diagnostic markers and novel therapeutic targets (Das, Hammamieh, Neill, Melhem, & Jett, 2001). Lawrie et colleagues demonstrated that FABP1 expression is down-regulated in the cancerous tissue of human colon with respect the normal mucosa (Lawrie, et al.,2004). Results obtained in this study are in accordance with these, and highlighted that FABP1 is quite absent in tumoral tissue of patients classified as Dukes stadium A and its concentration increased in the tumoral tissue of patients in stadium B and C, although was significantly lower than in normal mucosa.

We have characterized also the FABP1 94:T>A variant in colonic tissues examined, its level did not change in the different kind of tissue, normal and tumoral, even if appeared to be more abundant in normal mucosa than in tumor tissue. This variant was investigated for its capacity to alter serum lipoprotein cholesterol levels in human subjects, nothing is known whereby the variant elicits these effects, and alterations in levels of this protein in the tumor (such as a reduction in the levels) may contribute to tumorigenesis (H. Huang et al., 2015).

6.0 *Conclusions*

The results reported in this thesis have the advantage of coming from the analysis of three tissue samples taken from different colon areas of the same patient, superficial and deep tumor and health mucosa as an internal control. We have optimized the procedure of protein tissue extraction in order to have the highest yield of β-thymosins. The technique optimization has been conducted on both three types tissues provided for each case, ensuring a certain confidence acceptability of the results. From the data collected in this thesis can be deduced that the key point for achieving an optimal extraction method for β-thymosins is to use fresh tissue and to proceed with an ultrafiltration step.

The proteomic characterization of the tissue samples showed that, away from the presence of β-thymosins, also various peptides/proteins were measured (such as ubiquitin, FABP1, SH3BP1, CA1, prothymosin and parathymosin). Thus, the method here applied appears to be ideal for extracting the protein fraction with low molecular weight from colon tissue samples. By applying this procedure, it was possible to observe how various peptides and proteins behaved differently depending on the type of tissue (surface or deep) and when

compared to healthy mucosa. In fact, the work, here described, highlighted that Tβ4 and β10, Ubiquitin and pro-Tα localizes preferentially in the deep tumor, which is the front of tumor invasion, whereas CA1, FABP1 and SH3BP1 are down-regulated in the tumoral tissue, and reached the minimum level in the surface of the tumor that corresponds to the old transformed tissue. The variation levels of these peptides and proteins in the different kind of tissues examined appeared to be linked to the Dukes stadium and thus to the degree of tumor infiltration, especially as regards thymosins β, ubiquitin and SH3BP1. These data are promising in order to individuate novel candidate biomarkers useful to typify colon-rectal cancer.

Biblography

Álvarez-Chaver, P., Otero-Estévez, O., de la Cadena, M. P., Rodríguez-Berrocal, F. J., & Martínez-Zorzano, V. S. (2014). Proteomics for discovery of candidate colorectal cancer biomarkers. *World Journal of Gastroenterology*, *20*(14), 3804–3824.

Astler, V. B., & Coller, F. A. (1954). The Prognostic Significance of Direct Extension of Carcinoma of the Colon and Rectum. *Annals of Surgery*, *139*(6), 846–51.

Barderas, R., Babel, I., & Casal, J. I. (2010). Colorectal cancer proteomics, molecular characterization and biomarker discovery. *Proteomics - Clinical Applications*.

Bednarek, R., Boncela, J., Smolarczyk, K., Cierniewska-Cieslak, A., Wyroba, E., and Cierniewski. C. S. (2008). Ku80 as novel receptor for thymosin β4 that mediates its intracellular activity different from G-actin sequestering. *J. Biol. Chem*. 283, 1534–1544.

Bingham, S., & Riboli, E. (2004). Diet and cancer [mdash] the European Prospective Investigation into Cancer and Nutrition. *Nat Rev Cancer*, *4*(3), 206–215.

Bock-Marquette, I., Saxena, A., White, M. D., Dimaio, J. M., & Srivastava, D. (2004). Thymosin beta4 activates integrin-linked kinase and promotes cardiac cell migration, survival and cardiac repair. *Nature*, *432*(7016), 466–472.

Brieger A, Plotz G , Zeuzam S, Trojan J. (2007). Thymosin beta 4 expression and nuclear transport are regulated by hMLH1. *Biochem Biophys Res Commun* 364:731-736.

Cabras, T., Iavarone, F., Martelli, C., Delfino, D., Rossetti, D. V., Inserra, I., … Castagnola, M. (2015). High-resolution mass spectrometry for thymosins detection and characterization. *Expert Opinion on Biological Therapy*, 0.

Caldarella, J., Goodall, G. J., Felix, A. M., Heimer, E. P., Salvin, S. B., & Horecker, B. L. (1983). Thymosin alpha 11: a peptide related to thymosin alpha 1 isolated from calf thymosin fraction 5. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(24), 7424–7427.

Calon, A., Espinet, E., Palomo-Ponce, S., Tauriello, D. V. F., Iglesias, M., Cespedes, M. V., … Batlle, E. (2012). Dependency of colorectal cancer on a TGF-beta-driven program in stromal cells for metastasis initiation. *Cancer Cell*, *22*(5), 571–584.

Carvalho, B., Sillars-Hardebol, A. H., Postma, C., Mongera, S., Droste, J. T. S., Obulkasim, A., … Meijer, G. A. (2012). Colorectal adenoma to carcinoma progression is accompanied by changes in gene expression

associated with ageing, chromosomal instability, and fatty acid metabolism. *Cellular Oncology*, *35*(1), 53–63.

Celis, J. E., Ostergaard, M., Basse, B., Celis, A., Lauridsen, J. B., Ratz, G. P., … Rasmussen, H. H. (1996). Loss of adipocyte-type fatty acid binding protein and other protein biomarkers is associated with progression of human bladder transitional cell carcinomas. *Cancer Res.*, *56*(20), 4782–4790.

Cha, H.-J., Jeong, M.-J., & Kleinman, H. K. (2003). Role of thymosin beta4 in tumor metastasis and angiogenesis. *Journal of the National Cancer Institute*, *95*(22), 1674–1680.

Chan, L., Wei, C. F., Li, W. H., Yang, C. Y., Ratner, P., Pownall, H., … Smith, L. C. (1985). Human liver fatty acid binding protein cDNA and amino acid sequence. Functional and evolutionary implications. *The Journal of Biological Chemistry*, *260*(5), 2629–32.

Chiappetta, G., Pentimalli, F., Monaco, M., Fedele, M., Pasquinelli, R., Pierantoni, G. M., … Fusco, A. (2004). Thymosin beta-10 gene expression as a possible tool in diagnosis of thyroid neoplasias. *Oncology Reports*, *12*(2), 239–243.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., … Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science (New York, N.Y.)*, *325*(5942), 834–40.

Ciechanover, A., & Schwartz, A. L. (2004). The ubiquitin system: Pathogenesis of human diseases and drug targeting. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1695*(1), 3–17.
Concepts, B. (2012). AJCC Cancer Staging Atlas. *AJCC Cancer Staging Atlas*, 23–32.

Das, R., Hammamieh, R., Neill, R., Melhem, M., & Jett, M. (2001). Expression pattern of fatty acid-binding proteins in human normal and cancer prostate cells and tissues. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *7*(6), 1706–1715.

De Méndez Morelva, T., & Antonio, L. B. (2009). Immunohistochemical expression of ubiquitin and telomerase in cervical cancer. *Virchows Archiv*, *455*(3), 235–243.

De Wit, M., Fijneman, R. J. A., Verheul, H. M. W., Meijer, G. A., & Jimenez, C. R. (2013). Proteomics in colorectal cancer translational research: biomarker discovery for clinical applications. *Clinical Biochemistry*, *46*(6), 466–479.

Dukes, C. E. (1980). The classification of cancer of the rectum. *Diseases of the Colon & Rectum*, *23*(8), 605–611.

Faa, G., Nemolato, S., Cabras, T., Fanni, D., Gerosa, C., Fanari, M., … Castagnola, M. (2012). Thymosin beta4 expression reveals intriguing similarities between fetal and cancer cells. *Annals of the New York*

Academy of Sciences, *1269*, 53–60.

Finley, D., Özkaynak, E., & Varshavsky, A. (1987). The yeast polyubiquitin gene is essential for resistance to high temperatures, starvation, and other stresses. *Cell*, *48*(6), 1035–1046.

Fu, X., Cui, P., Chen, F., Xu, J., Gong, L., Jiang, L., … Xiao, Y. (2015). Thymosin beta4 promotes hepatoblastoma metastasis via the induction of epithelial-mesenchymal transition. *Molecular Medicine Reports*, *12*(1), 127–132.

Garber, K. (2008, February). Epithelial-to-mesenchymal transition is important to metastasis, but questions remain. *Journal of the National Cancer Institute*. United States.

Gatenby, R. A., & Gillies, R. J. (2004). Why do cancers have high aerobic glycolysis? *Nature Reviews Cancer*, *4*(11), 891–899.

Ginsberg RJ, Vokes EE, R. A. (1997). DeVita, Cancer: principles and practice of oncology. 5th Edn., (July), 858–911.

Glatz, J. F., & van der Vusse, G. J. (1996). Cellular fatty acid-binding proteins: their function and physiological significance. *Progress in Lipid Research*, *35*(3), 243–282.

Glatz, J. F., van Nieuwenhoven, F. a, Luiken, J. J., Schaap, F. G., & van der Vusse, G. J. (1997). Role of membrane-associated and cytoplasmic fatty acid-binding proteins in cellular fatty acid metabolism. *Prostaglandins, Leukotrienes, and Essential Fatty Acids*, *57*(4–5), 373–8.

Glotzer, M., Murray, A. W., & Kirschner, M. W. (1991). Cyclin is degraded by the ubiquitin pathway. *Nature*, *349*(6305), 132–138.

Goldstein, A. L. (2003, November). Thymosin beta4: a new molecular target for antitumor strategies. *Journal of the National Cancer Institute*. United States.

Goldstein, A. L., Hannappel, E., & Kleinman, H. K. (2005). Thymosin beta4: actin-sequestering protein moonlights to repair injured tissues. *Trends in Molecular Medicine*, *11*(9), 421–429.

Goldstein, A. L., Hannappel, E., Sosne, G., & Kleinman, H. K. (2012). Thymosin beta4: a multi-functional regenerative peptide. Basic properties and clinical applications. *Expert Opinion on Biological Therapy*, *12*(1), 37–51.

Guarino, M., Rubino, B., & Ballabio, G. (2007). The role of epithelial-mesenchymal transition in cancer pathology. *Pathology*, *39*(3), 305–318.

Haggar, F. A., & Boushey, R. P. (2009). Colorectal Cancer Epidemiology: Incidence, Mortality, Survival, and Risk Factors. *Clinics in Colon and Rectal Surgery*, *22*(4), 191–197.

Hannappel, E. (2007). Beta-Thymosins. *Annals of the New York Academy of Sciences*, *1112*, 21–37.

Hannappel, E. (2010). Thymosin beta4 and its posttranslational modifications. *Annals of the New York Academy of Sciences*, *1194*, 27–35.

Hardesty, W. M., Kelley, M. C., Mi, D., Low, R. L., & Caprioli, R. M. (2011). Protein signatures for survival and recurrence in metastatic melanoma. *Journal of Proteomics*, *74*(7), 1002–1014.

Haritos, A. A., Goodall, G. J., & Horecker, B. L. (1984). Prothymosin alpha: isolation and properties of the major immunoreactive form of thymosin alpha 1 in rat thymus. *Proceedings of the National Academy of Sciences of the United States of America*, *81*(4), 1008–1011.

Haritos, A. A., Tsolas, O., & Horecker, B. L. (1984). Distribution of prothymosin alpha in rat tissues. *Proceedings of the National Academy of Sciences of the United States of America*, *81*(5), 1391–1393.

Heidecke, H., Eckert, K., Schulze-Forster, K., & Maurer, H. R. (1997). Prothymosin alpha 1 effects in vitro on chemotaxis, cytotoxicity and oxidative response of neutrophils from melanoma, colorectal and breast tumor patients. *International Journal of Immunopharmacology*, *19*(8), 413–420.

Hermsen, M., Postma, C., Baak, J., Weiss, M., Rapallo, A., Sciutto, A., … Meijer, G. (2002). Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. *Gastroenterology*, *123*(4), 1109–1119.

Hershko, A., & Ciechanover, A. (1998). The ubiquitin system. *Annual Review of Biochemistry*, *67*, 425–479.

Hertzel, A. V., & Bernlohr, D. A. (1998). Regulation of adipocyte gene expression by polyunsaturated fatty acids. In *Molecular and Cellular Biochemistry* (Vol. 188, pp. 33–39).

Hoff, G., & Dominitz, J. a. (2010). Contrasting US and European approaches to colorectal cancer screening: which is best? *Gut*, *59*(3), 407–14.

Hong, K.-O., Lee, J.-I., Hong, S.-P., & Hong, S.-D. (2016). Thymosin beta4 induces proliferation, invasion, and epithelial-to-mesenchymal transition of oral squamous cell carcinoma. *Amino Acids*, *48*(1), 117–127.

Huang, H.-C., Hu, C.-H., Tang, M.-C., Wang, W.-S., Chen, P.-M., & Su, Y. (2007). Thymosin beta4 triggers an epithelial-mesenchymal transition in colorectal carcinoma by upregulating integrin-linked kinase. *Oncogene*, *26*(19), 2781–2790.

Huang, H., McIntosh, A. L., Landrock, K. K., Landrock, D., Storey, S. M., Martin, G. G., … Schroeder, F. (2015). Human FABP1 T94A variant enhances cholesterol uptake. *Biochimica et Biophysica Acta*, *1851*(7), 946–955.

Huff, T., Müller, C. S. G., Otto, A. M., Netzker, R., & Hannappel, E. (2001). β-thymosins, small acidic peptides with multiple functions. *International Journal of Biochemistry and Cell Biology*.

Huff, T., Muller, C. S., & Hannappel, E. (1997). C-terminal truncation of thymosin beta10 by an intracellular protease and its influence on the interaction with G-actin studied by ultrafiltration. *FEBS Letters*, *414*(1), 39–44.

Huff, T., Otto, A. M., Muller, C. S. G., Meier, M., & Hannappel, E. (2002). Thymosin beta4 is released from human blood platelets and attached by factor XIIIa (transglutaminase) to fibrin and collagen. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *16*(7), 691–696.

Huff, T., Rosorius, O., Otto, A. M., Muller, C. S. G., Ballweber, E., Hannappel, E., & Mannherz, H. G. (2004). Nuclear localisation of the G-actin sequestering peptide thymosin beta4. *Journal of Cell Science*, *117*(Pt 22), 5333–5341.

Hurlstone, D. P., Cross, S. S., Adam, I., Shorthouse, A. J., Brown, S., Sanders, D. S., & Lobo, A. J. (2004). Efficacy of high magnification chromoscopic colonoscopy for the diagnosis of neoplasia in flat and depressed lesions of the colorectum: a prospective analysis. *Gut*, *53*(2), 284–90.

Inzitari, R., Cabras, T., Pisano, E., Fanali, C., Manconi, B., Scarano, E., … Messana, I. (2009). HPLC-ESI-MS analysis of oral human fluids reveals that gingival crevicular fluid is the main source of oral thymosins β4 and β10. *Journal of Separation Science*, *32*(1), 57–63.

Ishibashi, Y., Hanyu, N., Suzuki, Y., Yanai, S., Tashiro, K., Usuba, T., … Yanaga, K. (2004). Quantitative analysis of free ubiquitin and multi-ubiquitin chain in colorectal cancer. *Cancer Letters*, *211*(1), 111–117.

Ishibashi, Y., Takada, K., Joh, K., Ohkawa, K., Aoki, T., & Matsuda, M. (1991). Ubiquitin immunoreactivity in human malignant tumours. *British Journal of Cancer*, *63*(2), 320–322.

Jing, C., Beesley, C., Foster, C. S., Rudland, P. S., Fujii, H., Ono, T., … Ke, Y. (2000). Identification of the messenger RNA for human cutaneous fatty acid-binding protein as a metastasis inducer. *Cancer Research*, *60*(9), 2390–2398.

Jones, S., Chen, W.-D., Parmigiani, G., Diehl, F., Beerenwinkel, N., Antal, T., … Markowitz, S. D. (2008). Comparative lesion sequencing provides insights into tumor evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(11), 4283–8.

Kanayama, H. omi, Tanaka, K., Aki, M., Kagawa, S., Miyaji, H., Satoh, M., … Ichihara, A. (1991). Changes in Expressions of Proteasome and Ubiquitin Genes in Human Renal Cancer Cells. *Cancer Research*, *51*(24), 6677–6685.

Kannagi, R., Sakuma, K., Cai, B.-H., & Yu, S.-Y. (2015). Tumor-Associated Glycans and Their Functional Roles in the Multistep Process of Human Cancer Progression. In T. Suzuki, K. Ohtsubo, & N. Taniguchi (Eds.), *Sugar Chains: Decoding the Functions of Glycans* (pp. 139–158). Tokyo: Springer Japan.

Kazokaité, J., Ames, S., Becker, H. M., Deitmer, J. W., & Matulis, D. (2016). Selective inhibition of human carbonic anhydrase IX in Xenopus oocytes and MDA-MB-231 breast cancer cells. *Journal of Enzyme Inhibition and Medicinal Chemistry*.

Kirklin, J. W., Dockerty, M. B., & Waugh, J. M. (1949). The role of the peritoneal reflection in the prognosis of carcinoma of the rectum and sigmoid colon. *Surgery, Gynecology & Obstetrics*, *88*(3), 326–331.

Lamlum, H., Al Tassan, N., Jaeger, E., Frayling, I., Sieber, O., Reza, F. B., … Tomlinson, I. (2000). Germline APC variants in patients with multiple colorectal adenomas, with evidence for the particular importance of E1317Q. *Human Molecular Genetics*, *9*(15), 2215–2221.

Larsson, L.-I., & Holck, S. (2007). Localization of thymosin beta-4 in tumors. *Annals of the New York Academy of Sciences*, *1112*, 317–325.

L C Lawrie, S R Dundas, S Curran, and G I Murray (2004). Liver fatty acid binding protein expression in colorectal neoplasia. *British Journal of cancer*, 90(10): 1955–1960

Lee, S. H., Zhang, W., Choi, J. J., Cho, Y. S., Oh, S. H., Kim, J. W., … Lee, J. H. (2001). Overexpression of the thymosin beta-10 gene in human ovarian cancer cells disrupts F-actin stress fiber and leads to apoptosis. *Oncogene*, *20*(46), 6700–6706.

Lee, S. J., So, I. S., Park, S. Y., and Kim, I. S. (2008) Thymosin β4 is involved in stabilin-2-mediated apoptotic cell engulfment. *FEBS Lett.* 582, 2161–2166

Lee, Y. J., Kim, J. H., Kim, S. K., Ha, S. J., Mok, T. S., Mitsudomi, T., & Cho, B. C. (2011). Lung cancer in never smokers: Change of a mindset in the molecular era. *Lung Cancer*, *72*(1), 9–15.

Levayer, R., & Lecuit, T. (2008). Breaking down EMT. *Nature Cell Biology*, *10*(7), 757–759.

Lin, K. T., & Deutsch, H. F. (1974). Human carbonic anhydrases. XII. The complete primary structure of the C isozyme. *The Journal of Biological Chemistry*, *249*(8), 2329–2337.

Luo, J., Solimini, N. L., & Elledge, S. J. (2009). Principles of Cancer Therapy: Oncogene and Non-oncogene

Addiction. *Cell*.

Majid, S. M., Liss, a S., You, M., & Bose, H. R. (2006). The suppression of SH3BGRL is important for v-Rel-mediated transformation. *Oncogene*, *25*(5), 756–68.

Maki, C. G., Huibregtse, J. M., & Howley, P. M. (1996). In vivo ubiquitination and proteasome-mediated degradation of p53. *Cancer Research*, *56*(11), 2649–2654.

Malinda, K. M., Sidhu, G. S., Mani, H., Banaudha, K., Maheshwari, R. K., Goldstein, A. L., & Kleinman, H. K. (1999). Thymosin beta4 accelerates wound healing. *The Journal of Investigative Dermatology*, *113*(3), 364–368.

Mannherz, H. G., & Hannappel, E. (2009). The beta-thymosins: intracellular and extracellular activities of a versatile actin binding protein family. *Cell Motility and the Cytoskeleton*, *66*(10), 839–851.

Markowitz, S. D., & Bertagnolli, M. M. (2009). Molecular Origins of Cancer: Molecular Basis of Colorectal Cancer. *The New England Journal of Medicine*, *361*(25), 2449–2460.

Martelli, C., Iavarone, F., D'Angelo, L., Arba, M., Vincenzoni, F., Inserra, I., … Desiderio, C. (2015). Integrated proteomic platforms for the comparative characterization of medulloblastoma and pilocytic astrocytoma pediatric brain tumors: a preliminary study. *Molecular bioSystems*, *11*(6), 1668–83.

Mazzocco M., Arrigo P., Egeo A., Maffei M., Vergano A., Di Lisi R., Ghiotto F., Ciccone E., Cinti R., Ravazzolo R., Scartezzini P. (2001). "A novel human homologue of the SH3BGR gene encodes a small protein similar to glutaredoxin 1 of Escherichia coli." *Biochem. Biophys. Res. Commun*., 285:540-545

McDoniels-Silvers, A. L., Nimri, C. F., Stoner, G. D., Lubet, R. A., & You, M. (2002). Differential gene expression in human lung adenocarcinomas and squamous cell carcinomas. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *8*(4), 1127–1138.

Mora, J., Cheung, N. K., Kushner, B. H., LaQuaglia, M. P., Kramer, K., Fazzari, M., … Gerald, W. L. (2000). Clinical categories of neuroblastoma are associated with different patterns of loss of heterozygosity on chromosome arm 1p. *J Mol Diagn*, *2*(1), 37–46.

Nelson, J. E., Meier, D. E., Oei, E. J., Nierman, D. M., Senzel, R. S., Manfredi, P. L., … Morrison, R. S. (2001). Self-reported symptom experience of critically ill cancer patients receiving intensive care. *Critical Care Medicine*, *29*(2), 277–282.

Nemolato, S., Cabras, T., Cau, F., Fanari, M. U., Fanni, D., Manconi, B., … Faa, G. (2010). Different thymosin Beta 4 immunoreactivity in foetal and adult gastrointestinal tract. *PloS One*, *5*(2), e9111.

Nemolato, S., Cabras, T., Fanari, M. U., Cau, F., Fanni, D., Gerosa, C., … Faa, G. (2010). Immunoreactivity of thymosin beta 4 in human foetal and adult genitourinary tract. *European Journal of Histochemistry : EJH*, *54*(4), e43.

Nemolato, S., Cabras, T., Fanari, M. U., Cau, F., Fraschini, M., Manconi, B., … Faa, G. (2010). Thymosin beta 4 expression in normal skin, colon mucosa and in tumor infiltrating mast cells. *European Journal of Histochemistry : EJH*, *54*(1), e3.

Nemolato, S., Messana, I., Cabras, T., Manconi, B., Inzitari, R., Fanali, C., … Castagnola, M. (2009). Thymosin beta(4) and beta(10) levels in pre-term newborn oral cavity and foetal salivary glands evidence a switch of secretion during foetal development. *PloS One*, *4*(4), e5109.

Nemolato, S., Restivo, A., Cabras, T., Coni, P., Zorcolo, L., Orru, G., … Faa, G. (2012). Thymosin beta 4 in colorectal cancer is localized predominantly at the invasion front in tumor cells undergoing epithelial mesenchymal transition. *Cancer Biology & Therapy*, *13*(4), 191–197.

Nemolato, S., Van Eyken, P., Cabras, T., Cau, F., Fanari, M. U., Locci, A., … Faa, G. (2011). Expression pattern of thymosin beta 4 in the adult human liver. *European Journal of Histochemistry : EJH*, *55*(3), e25.

Neri, D., & Supuran, C. T. (2011). Interfering with pH regulation in tumours as a therapeutic strategy. *Nature Reviews. Drug Discovery*, *10*(10), 767–77.

Nickolay Neznanov, A. P. K. L. N. P. S.-B. A. V. G. (2011). Proteotoxic stress targeted therapy (PSTT): induction of protein misfolding enhances the antitumor effect of the proteasome inhibitor bortezomib. *Oncotarget*, *2*(3), 209.

Oh C, Park S, Lee EK, Yoo YJ (2013). Downregulation of ubiquitin level via knockdown of polyubiquitin gene Ubb as potential cancer therapeutic intervention. *Scientific Reports*, doi: 10.1038

Parks, S. K., Chiche, J., & Pouyssegur, J. (2013). Disrupting proton dynamics and energy metabolism for cancer therapy. *Nature Reviews Cancer*, *13*(9), 611–623.

Plavina, T., Hincapie, M., Wakshull, E., Subramanyam, M., & Hancock, W. S. (2008). Increased plasma concentrations of cytoskeletal and Ca2+-binding proteins and their peptides in psoriasis patients. *Clinical Chemistry*, *54*(11), 1805–14.

Qiu P, Wheater MK, Qiu Y, Sosne G. (2011) Thymosin β4 inhibits TNF-α-induced NF-κB activation, IL-8 expression, and the sensitizing effects by its partners PINCH-1 and ILK. *FASEB J*, 25:1815-1826

Richieri, G. V, Ogata, R. T., Zimmerman, a W., Veerkamp, J. H., & Kleinfeld, a M. (2000). Fatty acid binding proteins from different tissues show distinct patterns of fatty acid interactions. *Biochemistry*, *39*(24),

7197–204.

Safer, D., Sosnick, T. R., & Elzinga, M. (1997). Thymosin beta 4 binds actin in an extended conformation and contacts both the barbed and pointed ends. *Biochemistry*, *36*(19), 5806–5816.

Sagynaliev, E., Steinert, R., Nestler, G., Lippert, H., Knoch, M., & Reymond, M. A. (2005). Web-based data warehouse on gene expression in human colorectal cancer. In *Proteomics* (Vol. 5, pp. 3066–3078).

Sarandeses, C. S., Covelo, G., Diaz-Jullien, C., & Freire, M. (2003). Prothymosin alpha is processed to thymosin alpha 1 and thymosin alpha 11 by a lysosomal asparaginyl endopeptidase. *The Journal of Biological Chemistry*, *278*(15), 13286–13293.

Scartezzini, P., Egeo, A., Colella, S., Fumagalli, P., Arrigo, P., Nizetic, D., … Rasore-Quartino, A. (1997). Cloning a new human gene from chromosome 21q22.3 encoding a glutamic acid-rich protein expressed in heart and skeletal muscle. *Human Genetics*, *99*(3), 387–392.

Schroeder, F., Atshaves, B. P., Starodub, O., Boedeker, A. L., Smith 3rd, R. R., Roths, J. B., … Kier, A. B. (2001). Expression of liver fatty acid binding protein alters growth and differentiation of embryonic stem cells. *Mol Cell Biochem*, *219*(1–2), 127–138.

Sethi, M. K., Hancock, W. S., & Fanayan, S. (2016). Identifying N - Glycan Biomarkers in Colorectal Cancer by Mass Spectrometry.

Sillars-Hardebol, A. H., Carvalho, B., Van Engeland, M., Fijneman, R. J., & Meijer, G. A. (2012). The adenoma hunt in colorectal cancer screening: Defining the target. *Journal of Pathology*.

Sly, W. S., & Hu, P. Y. (1995). Human Carbonic Anhydrases and Carbonic Anhydrase Deficiencies. *Annual Review of Biochemistry*, *64*(1), 375–401.

Sorof, S. (1994). Modulation of mitogenesis by liver fatty acid binding protein. *Cancer Metastasis Reviews*, *13*(3–4), 317–336.

Sribenja, S., Wongkham, S., Wongkham, C., Yao, Q., Sribenja, S., Wongkham, S., … Chen, C. (2016). Roles and Mechanisms of β -Thymosins in Cell Migration and Cancer Metastasis : An Update Roles and Mechanisms of β -Thymosins in Cell Migration and Cancer Metastasis : An Update, *7907*(October).

Suarez, B. K., Lin, J., Burmester, J. K., Broman, K. W., Weber, J. L., Banerjee, T. K., … Catalona, W. J. (2000). A genome screen of multiplex sibships with prostate cancer. *American Journal of Human Genetics*, *66*, 933–944.

Sung, J. J. Y., Lau, J. Y. W., Goh, K. L., & Leung, W. K. (2005). Increasing incidence of colorectal cancer in

Asia: implications for screening. *The Lancet. Oncology*, *6*(11), 871–876.

Supuran, C. T. (2007). Carbonic anhydrases as drug targets--an overview. *Current Topics in Medicinal Chemistry*, *7*(9), 825–833.

Supuran, C. T. (2008). Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. *Nature Reviews. Drug Discovery*, *7*(2), 168–181.

Szabo, P., Clinton, M., Macera, M., & Horecker, B. L. (1989). Localization of the gene coding for parathymosin to chromosome 17 in humans. *Cytogenetics and Cell Genetics*, *50*(2–3), 91–92.

Szabo, P., Panneerselvam, C., Clinton, M., Frangou-Lazaridis, M., Weksler, D., Whittington, E., … Horecker, B. L. (1993). Prothymosin alpha gene in humans: organization of its promoter region and localization to chromosome 2. *Human Genetics*, *90*(6), 629–634.

Tanaka, T., Tanaka, M., Tanaka, T., & Ishigamori, R. (2010). Biomarkers for colorectal cancer. *International Journal of Molecular Sciences*.

Tata, A., Stoppel, D. C., Hong, S., Ben-Zvi, A., Xie, T., & Gu, C. (2014). An image-based RNAi screen identifies SH3BP1 as a key effector of Semaphorin 3E-PlexinD1 signaling. *The Journal of Cell Biology*, *205*(4), 573–590.

Theunissen, W., Fanni, D., Nemolato, S., Di Felice, E., Cabras, T., Gerosa, C., … Faa, G. (2014). Thymosin beta 4 and thymosin beta 10 expression in hepatocellular carcinoma. *European Journal of Histochemistry : EJH*, *58*(1), 2242.

Trompeter, H. I., & Soling, H. D. (1992). Cloning and characterisation of a gene encoding the 11.5 kDa zinc-binding protein (parathymosin-alpha). *FEBS Letters*, *298*(2–3), 245–248.

Tse, J. C., & Kalluri, R. (2007). Mechanisms of metastasis: epithelial-to-mesenchymal transition and contribution of tumor microenvironment. *Journal of Cellular Biochemistry*, *101*(4), 816–829.

Tsitsiloni, O. E., Stiakakis, J., Koutselinis, A., Gogas, J., Markopoulos, C., Yialouris, P., … Voelter, W. (1993). Expression of alpha-thymosins in human tissues in normal and abnormal growth. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(20), 9504–9507.

Ueda, H., Matsunaga, H., & Halder, S. K. (2012). Prothymosin ?? plays multifunctional cell robustness roles in genomic, epigenetic, and nongenomic mechanisms. *Annals of the New York Academy of Sciences*, *1269*(1), 34–43.

Uzozie, A., Nanni, P., Staiano, T., Grossmann, J., Barkow-Oesterreicher, S., Shay, J. W., … Marra, G. (2014).

Sorbitol dehydrogenase overexpression and other aspects of dysregulated protein expression in human precancerous colorectal neoplasms: a quantitative proteomics study. *Molecular & Cellular Proteomics : MCP*, *13*(5), 1198–1218.

Vareli, K., Frangou-Lazaridis, M., van der Kraan, I., Tsolas, O., & van Driel, R. (2000). Nuclear distribution of prothymosin alpha and parathymosin: evidence that prothymosin alpha is associated with RNA synthesis processing and parathymosin with early DNA replication. *Experimental Cell Research*, *257*(1), 152–61.

Veerkamp, J. H., van Moerkerk, H. T., Prinsen, C. F., & van Kuppevelt, T. H. (1999). Structural and functional studies on different human FABP types. *Molecular and Cellular Biochemistry*, *192*(1–2), 137–42.

Verghese-Nikolakaki, S., Apostolikas, N., Livaniou, E., Ithakissios, D. S., & Evangelatos, G. P. (1996). Preliminary findings on the expression of thymosin beta-10 in human breast cancer. *British Journal of Cancer*, *74*(9), 1441–1444.

Viikila, P., Kivela, A. J., Mustonen, H., Koskensalo, S., Waheed, A., Sly, W. S., … Haglund, C. (2016). Carbonic anhydrase enzymes II, VII, IX and XII in colorectal carcinomas. *World Journal of Gastroenterology*, *22*(36), 8168–8177.

Wang Wei-Shu, Chen Po-Min, Hsiao Hung-Liang, Wang Huann-Sheng, Liang Wen-Yih, Su Yeu (2004). Overexpression of the thymosin β-4 gene is associated with increased invasion of SW480 colon carcinoma cells and the distant metastasis of human colorectal carcinoma. *Oncogene*, *23*(39), 6666–6671.

Wang, G., Gong, Y., Anderson, J., Sun, D., Minuk, G., Roberts, M. S., & Burczynski, F. J. (2005). Antioxidative function of L-FABP in L-FABP stably transfected Chang liver cells. *Hepatology (Baltimore, Md.)*, *42*(4), 871–879.

Weissman, a M. (2001). Themes and variations on ubiquitylation. *Nature Reviews. Molecular Cell Biology*, *2*(3), 169–178.

Wisniewski, J. R., Ostasiewicz, P., Dus, K., Zielinska, D. F., Gnad, F., & Mann, M. (2012). Extensive quantitative remodeling of the proteome between normal colon tissue and adenocarcinoma. *Molecular Systems Biology*, *8*, 611.

Wu, C. G., Boers, W., Reitsma, P. R., van Deventer, S. J., & Chamuleau, R. A. (1997). Overexpression of prothymosin alpha, concomitant with c-myc, during rat hepatic carcinogenesis. *Biochem Biophys Res Commun*, *232*(3), 817–821.

Wu, S. Q., Hafez, G. R., Xing, W., Newton, M., Chen, X. R., & Messing, E. (1996). The correlation between the loss of chromosome 14q with histologic tumor grade, pathologic stage, and outcome of patients with nonpapillary renal cell carcinoma. *Cancer*, *77*(6), 1154–1160.

Wykoff, C. C., Beasley, N. J. P., Watson, P. H., Turner, K. J., Pastorek, J., Sibtain, A., … Ratcliffe, P. J. (2000). Hypoxia-inducible expression of tumor-associated carbonic anhydrases. *Cancer Research*, *60*(24), 7075–7083.

Yin, X., Zhang, Y., Guo, S., Jin, H., Wang, W., & Yang, P. (2015). Large scale systematic proteomic quantification from non-metastatic to metastatic colorectal cancer. *Scientific Reports*, *5*, 12120.

Yoon, S. Y., Lee, H. R., Park, Y., Kim, J. H., Kim, S. Y., Yoon, S. R., … Cho, D. (2011). Thymosin beta4 expression correlates with lymph node metastasis through hypoxia inducible factor-alpha induction in breast cancer. *Oncology Reports*, *25*(1), 23–31.

Young, J. D., Lawrence, A. J., MacLean, A. G., Leung, B. P., McInnes, I. B., Canas, B., … Stevenson, R. D. (1999). Thymosin beta 4 sulfoxide is an anti-inflammatory agent generated by monocytes in the presence of glucocorticoids. *Nature Medicine*, *5*(12), 1424–1427.

Yu, S.-L., Chen, H.-Y., Chang, G.-C., Chen, C.-Y., Chen, H.-W., Singh, S., … Yang, P.-C. (2008). MicroRNA signature predicts survival and relapse in lung cancer. *Cancer Cell*, *13*(1), 48–57.

Zhang, Z., & Marshall, A. G. (1998). A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. *Journal of the American Society for Mass Spectrometry*, *9*(3), 225–233.

Supporting Information.

Table S1.

m/z values of peptides and proteins investigated in this study and utilized for their XIC quantification by HPLC-low-resolution-ESI-MS analysis.

Peptide / Protein (SwissProt code)	Post-translational modifications	Time elution (min)	Exp. (Theor.) Average Mass value (Da)	m/z and charge
Tβ4 (P62328)	Nα-acetylated	19.3-19.8	4962.8 ± 0.6 (4963.5)	*993.80 (+5), 1241.90, (+4), 1655.50(+3)*
	Nα-acetylated; Fragm. 1-41	19.6-20.1	4746.7 ± 0.6 (4747.31)	*950.46(+5), 1187.83,(+4), 1583.44(+3)*
	Nα-acetylated M6-sulfoxide	17.4-18.0	4978.8 ± 0.6 (4979.46)	*996.90(+5) 1245.90(+4), 1660.80(+3)*
	Nα-acetylated; Nε-acetylated K16/K25-diacetylated	21.2-21.6	5046.93 ± 0.6 (5047.58)	*1010.52(+5), 1262.90,(+4), 1683.53(+3)*
Tβ10 (P63313)	Nα-acetylated	20.3-20.6	4935.9 ± 0.6 (4936.48)	*988.30(+5), 1235.10,(+4), 1646.50(+3)*
	Nα-acetylated; Fragm. 1-41	20.0-20.7	4735.8 ± 0.6 (4736.28)	*948.26,(+5) 1185.07(+4), 1579.76(+3)*
	Nα-acetylated; M6-sulfoxide	20.3-20.9	4952.12 ± 0.6 (4952.48)	*1005.12,(+5) 1256.15,(+4), 1674.53 (+3)*
Pro-Tα, Isof. 2 (P06454)	M1 missing; Nα-acetylated;	20.7-21.2	11984 ± 1 (11984.7)	*922.90(+13), 999.73(+12), 1090.52(+11), 1199.47(+10), 1332.63(+9), 1499.09(+8)*
Tα11	Nα-acetylated; Frag. 1-36 of pro-Tα	19.5-20.2	3789.59 ± 0.5 (3790.02)	*759.00(+5), 948.51(+4), 1264.34(+3), 1896.01(+2)*
Parathymosin (P20962)	Nα-acetylated; Fragm. 2-102	20.5-20.7	11440 ± 1 (11440.8)	*881.06(+13), 954.40(+12), 1041.07(+11), 1145.08(+10), 1272.20(+9), 1431.10(+8)*
Ubiquitin (P0CG48)		29.5-30.2	8564.2 ± 1 (8564.8)	*779.62(+11), 857.48(+10), 952.64(+9), 1071.60(+8), 1224.54(+7), 1428.47(+6), 1713.96(+5)*
SH3BP-1 (Q9H299)	M1 missing; Nα-acetylated	38.5-38.9	10347.5 ± 1 (10348.5)	*797.04(+13), 863.38(+12), 941.77(+11), 1035.85(+10), 1150.83(+9), 1294.56(+8), 1479.36(+7), 1725.75(+6)*
FABP1 (P07148)	M1 missing; Nα-acetylated	40.5-40.9	14118 ± 2 (14119.2)	*785.30(+18), 831.44(+17), 883.34(+16), 942.16(+15), 1009.39(+14), 1086.95(+13), 1177.45(+12), 1284.40(+11), 1412.74(+10), 1569.60(+9), 1765.68(+8)*

Peptide / Protein (SwissProt code)	Post-translational modifications	Time elution (min)	Exp. (Theor.) Average Mass value (Da)	m/z and charge
FABP1 94:T>A variant, (P07148)	M1 missing; Nα-acetylated	40.5-40.9	14088 ± 2 (14089.2)	*783.64(+18), 829.68(+17), 881.48(+16), 940.17(+15), 1007.26(+14), 1084.66(+13), 1174.97(+12), 1281.69(+11), 1409.76(+10), 1566.29(+9), 1761.95(+8)*
Carbonic anhydrase 1, (P00915)	M1 missing; Nα-acetylated	40.3-40.7	28778 ± 3 (28781)	*929.26(+31), 960.20(+30), 993.28(+29), 1028.71(+28), 1066.78(+27), 1107.77(+26), 1152.04(+25), 1200.00(+24), 1252.13(+23), 1309.00(+22), 1371.29(+21), 1439.80(+20), 1515.53(+19), 1599.67(+18), 1693.71(+17)*
9955 Da protein		29.6-29.9	9955 ± 1	*905.89(+11), 996.38(+10), 1106.97(+9), 1245.22(+8), 1422.97(+7), 1659.96(+6), 1991.76(+5)*

Supporting Information Section I

Results of the top-down proteomics analysis performed on proteins of interest present in the >30 KDa and in the <30 KDa fractions.

Fig. S1. Thymosin β4.

$[M+H]^+$ 4961.49951 Da (Theor. 4961.4935). MS/MS analysis performed on different m/z ions in different samples.

a) Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+6H]^{6+}$ monoisotopic m/z: 827.75598 Da (+0.95 mmu/+1.14 ppm). RT: 19.07 min. Analysis performed by Proteome discoverer software. XCorr:4.61.

b) Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+4H]^{4+}$ monoisotopic m/z: 1241.12878 Da (-0.13 mmu/-0.11 ppm). RT: 19.24 min. Analysis performed by Proteome discoverer software. XCorr:6.05.

c) Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+5H]^{5+}$ monoisotopic m/z: 993.10529 Da (+0.7 mmu/+0.7 ppm). RT: 19.25 min. Analysis performed by Proteome discoverer software. XCorr:5.29

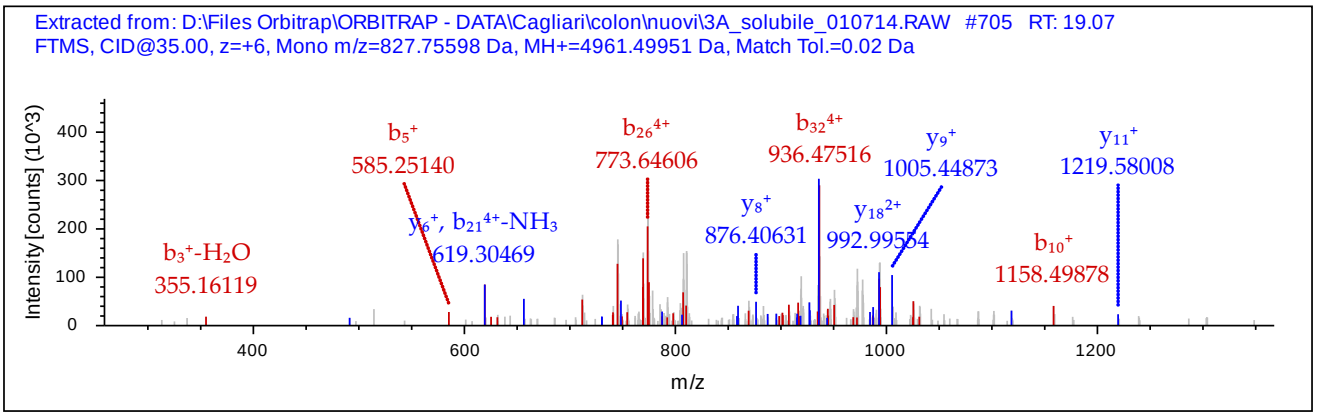
S1-Acetyl (42.01057 Da). Identified with: Sequest HT (v1.3). Fragment match tolerance used for search: 0.8 Da

Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3

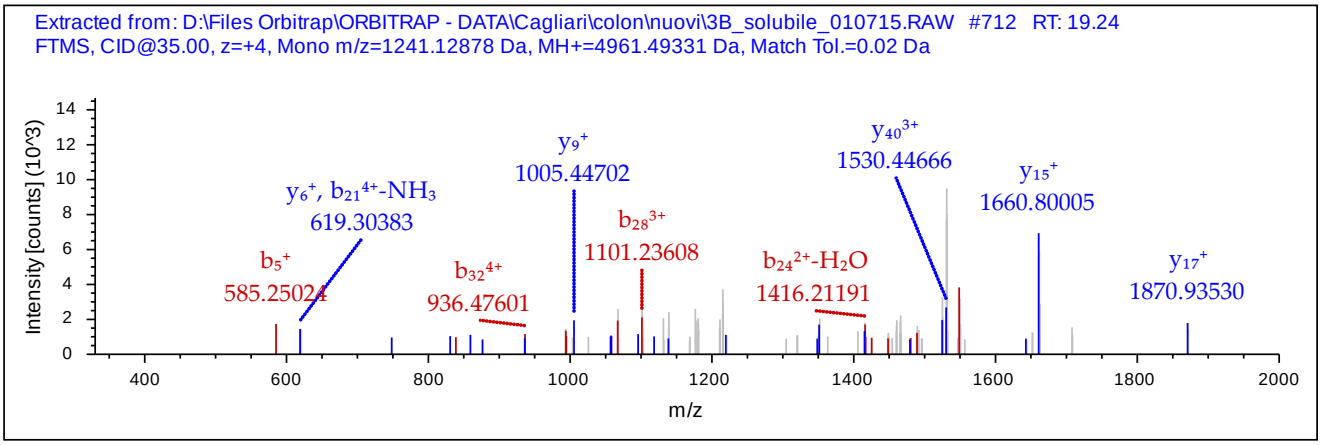
Protein references (1):

- Thymosin beta-4 OS=Homo sapiens GN=TMSB4X PE=1 SV=2 - [TYB4_HUMAN]

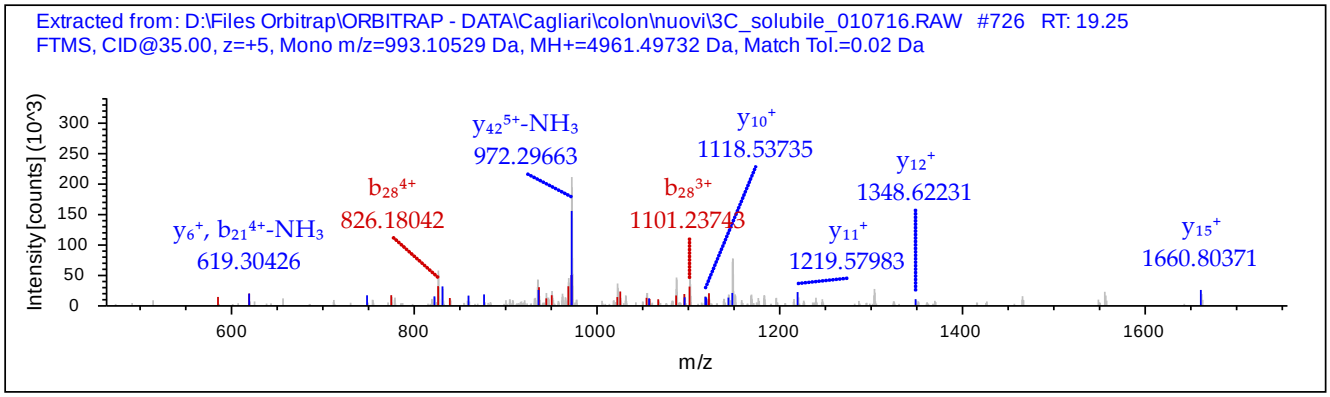
a)



b)



c)

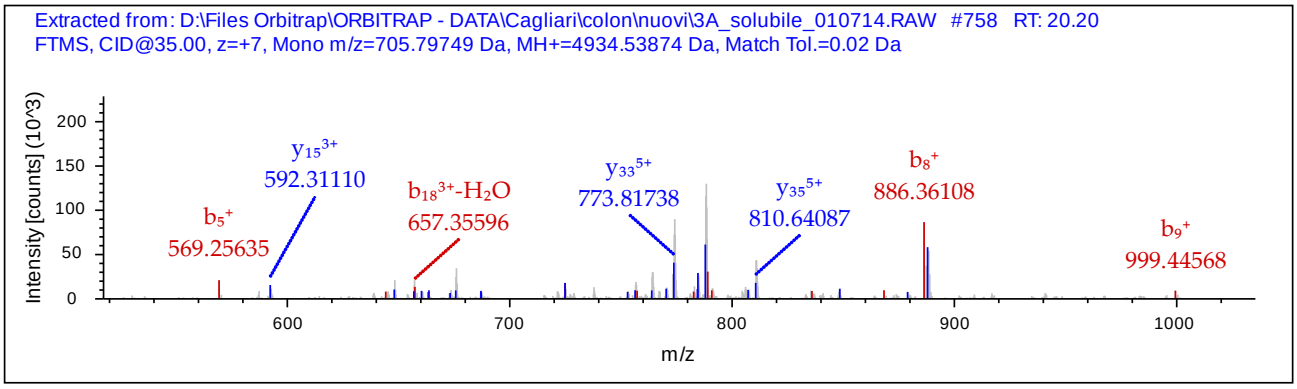


m/z ion match on Tβ4 sequence: b ions matched in the three MS/MS spectra are in red; y ions in blue

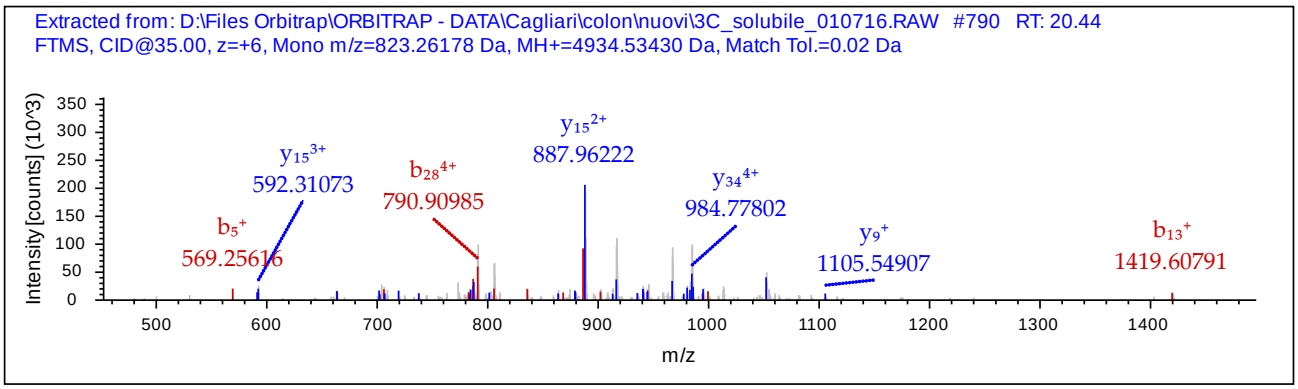
	b^{+}	b^{2+}	b^{3+}	b^{4+}	b^{5+}	b^{6+}	Seq.	y^{+}	y^{2+}	y^{3+}	y^{4+}	y^{5+}	y^{6+}	#2
#1	130.0498	65.5285	44.0214	33.2679	26.8158	22.5143	S-Acetyl							
2	245.0768	123.0421	82.3638	62.0247	49.8212	41.6855	D	4832.4512	2416.7293	1611.4886	1208.8683	967.2961	806.2479	42
3	373.1718	187.0895	125.0621	94.0484	75.4402	63.0347	K	4717.4243	2359.2158	1573.1463	1180.1115	944.2907	787.0768	41
4	470.2246	235.6159	157.4130	118.3116	94.8507	79.2102	P	4589.3293	2295.1683	1530.4480	1148.0878	918.6717	765.7276	40
5	585.2515	293.1294	195.7554	147.0683	117.8561	98.3813	D	4492.2766	2246.6419	1498.0970	1123.8246	899.2611	749.5522	39
6	716.2920	358.6496	239.4355	179.8285	144.0642	120.2214	M	4377.2496	2189.1284	1459.7547	1095.0679	876.2557	730.3810	38
7	787.3291	394.1682	263.1146	197.5877	158.2717	132.0609	A	4246.2091	2123.6082	1416.0746	1062.3077	850.0476	708.5409	37
8	916.3717	458.6895	306.1288	229.8484	184.0802	153.5680	E	4175.1720	2088.0896	1392.3955	1044.5485	835.8402	696.7014	36
9	1029.4558	515.2315	343.8235	258.1194	206.6970	172.4154	I	4046.1294	2023.5683	1349.3813	1012.2878	810.0317	675.1943	35
10	1158.4984	579.7528	386.8377	290.3801	232.5055	193.9225	E	3933.0453	1967.0263	1311.6866	984.0168	787.4149	656.3470	34
11	1286.5934	643.8003	429.5360	322.4038	258.1245	215.2716	K	3804.0027	1902.5050	1268.6724	951.7561	761.6064	634.8399	33
12	1433.6618	717.3345	478.5588	359.1709	287.5382	239.7830	F	3675.9077	1838.4575	1225.9741	919.7324	735.9874	613.4907	32
13	1548.6887	774.8480	516.9011	387.9276	310.5436	258.9542	D	3528.8393	1764.9233	1176.9513	882.9653	706.5737	588.9793	31
14	1676.7837	838.8955	559.5994	419.9514	336.1626	280.3034	K	3413.8124	1707.4098	1138.6090	854.2086	683.5683	569.8081	30
15	1763.8157	882.4115	588.6101	441.7094	353.5690	294.8087	S	3285.7174	1643.3623	1095.9107	822.1848	657.9493	548.4590	29
16	1891.9107	946.4590	631.3084	473.7331	379.1880	316.1579	K	3198.6854	1599.8463	1066.9000	800.4268	640.5429	533.9536	28
17	2004.9948	1003.0010	669.0031	502.0042	401.8048	335.0052	L	3070.5904	1535.7988	1024.2017	768.4031	614.9239	512.6045	27
18	2133.0898	1067.0485	711.7014	534.0279	427.4238	356.3544	K	2957.5063	1479.2568	986.5070	740.1320	592.3071	493.7571	26
19	2261.1847	1131.0960	754.3998	566.0516	453.0428	377.7035	K	2829.4114	1415.2093	943.8086	708.1083	566.6881	472.4080	25
20	2362.2324	1181.6198	788.0823	591.3136	473.2523	394.5448	T	2701.3164	1351.1618	901.1103	676.0846	541.0691	451.0588	24
21	2491.2750	1246.1411	831.0965	623.5742	499.0608	416.0519	E	2600.2687	1300.6380	867.4278	650.8226	520.8596	434.2175	23
22	2592.3227	1296.6650	864.7791	648.8361	519.2704	432.8932	T	2471.2261	1236.1167	824.4136	618.5620	495.0510	412.7104	22
23	2720.3813	1360.6943	907.4653	680.8508	544.8821	454.2363	Q	2370.1784	1185.5929	790.7310	593.3001	474.8415	395.8691	21
24	2849.4239	1425.2156	950.4795	713.1114	570.6906	475.7434	E	2242.1199	1121.5636	748.0448	561.2854	449.2298	374.5260	20
25	2977.5188	1489.2631	993.1778	745.1352	596.3096	497.0925	K	2113.0773	1057.0423	705.0306	529.0248	423.4213	353.0189	19
26	3091.5618	1546.2845	1031.1921	773.6459	619.1182	516.0997	N	1984.9823	992.9948	662.3323	497.0010	397.8023	331.6698	18
27	3188.6145	1594.8109	1063.5430	797.9091	638.5287	532.2752	P	1870.9394	935.9733	624.3180	468.4903	374.9937	312.6626	17
28	3301.6986	1651.3529	1101.2377	826.1801	661.1455	551.1225	L	1773.8866	887.4469	591.9670	444.2271	355.5831	296.4872	16
29	3398.7514	1699.8793	1133.5886	850.4433	680.5561	567.2980	P	1660.8025	830.9049	554.2724	415.9561	332.9663	277.6398	15
30	3485.7834	1743.3953	1162.5993	872.2013	697.9625	581.8033	S	1563.7497	782.3785	521.9214	391.6929	313.5558	261.4644	14
31	3613.8784	1807.4428	1205.2976	904.2251	723.5815	603.1525	K	1476.7177	738.8625	492.9108	369.9349	296.1494	246.9590	13
32	3742.9210	1871.9641	1248.3118	936.4857	749.3900	624.6596	E	1348.6227	674.8150	450.2124	337.9111	270.5304	225.6099	12
33	3843.9687	1922.4880	1281.9944	961.7476	769.5996	641.5008	T	1219.5801	610.2937	407.1982	305.6505	244.7219	204.1028	11
34	3957.0527	1979.0300	1319.6891	990.0186	792.2164	660.3482	I	1118.5325	559.7699	373.5157	280.3886	224.5123	187.2615	10
35	4086.0953	2043.5513	1362.7033	1022.2793	818.0249	681.8553	E	1005.4484	503.2278	335.8210	252.1176	201.8955	168.4141	9
36	4214.1539	2107.5806	1405.3895	1054.2939	843.6366	703.1984	Q	876.4058	438.7065	292.8068	219.8569	176.0870	146.9070	8
37	4343.1965	2172.1019	1448.4037	1086.5546	869.4451	724.7055	E	748.3472	374.6772	250.1206	187.8423	150.4753	125.5639	7
38	4471.2915	2236.1494	1491.1020	1118.5783	895.0641	746.0546	K	619.3046	310.1559	207.1064	155.5816	124.6667	104.0568	6
39	4599.3501	2300.1787	1533.7882	1150.5930	920.6758	767.3977	Q	491.2096	246.1085	164.4081	123.5579	99.0478	82.7077	5
40	4670.3872	2335.6972	1557.4672	1168.3523	934.8833	779.2373	A	363.1511	182.0792	121.7219	91.5432	73.4360	61.3646	4
41	4727.4086	2364.2080	1576.4744	1182.6076	946.2876	788.7408	G	292.1139	146.5606	98.0428	73.7839	59.2286	49.5251	3
42	4856.4512	2428.7293	1619.4886	1214.8683	972.0961	810.2479	E	235.0925	118.0499	79.0357	59.5286	47.8243	40.0215	2
43							S	106.0499	53.5286	36.0215	27.2679	22.0158	18.5144	1

Fig. S2. Thymosin β10.
$[M+H]^+$ 4934.53874Da (Theor. 4934.5302). MS/MS analysis performed on different m/z ions in different samples.
a) Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+7H]^{7+}$ monoisotopic m/z: 705.79749 Da (+1.17 mmu/+1.66 ppm). RT: 20.20 min. Analysis performed by Proteome discoverer software. XCorr:4.81.
b) Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+6H]^{6+}$ monoisotopic m/z: 823.26178 Da (+0.62 mmu/+0.76 ppm). RT: 20.44 min. Analysis performed by Proteome discoverer software. XCorr:6.30
S1-Acetyl (42.01057 Da). Identified with: Sequest HT (v1.3). Fragment match tolerance used for search: 0.8 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Thymosin beta-10 OS=Homo sapiens GN=TMSB10 PE=1 SV=2 - [TYB10_HUMAN]

a)



b)



m/z ion match on Tβ10 sequence: b ions matched in the two MS/MS spectra are in red; y ions in blue

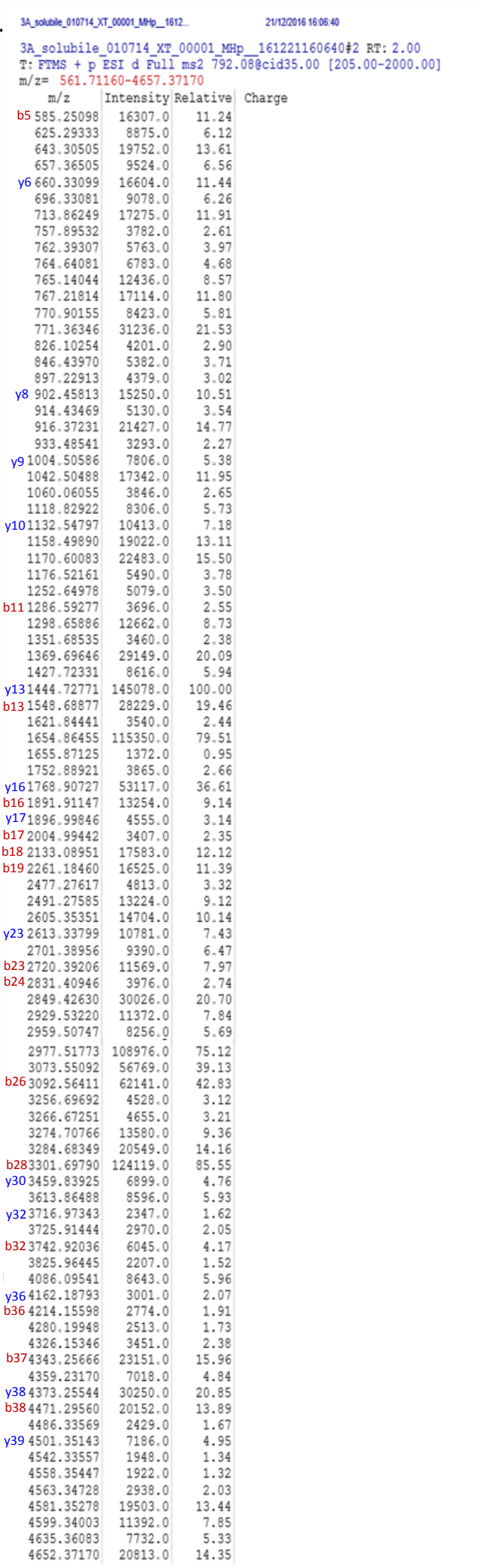
#1	b^+	b^{2+}	b^{3+}	b^{4+}	b^{5+}	b^{6+}	Seq.	y^+	y^{2+}	y^{3+}	y^{4+}	y^{5+}	y^{6+}	y^{7+}	#2
1	114.0550	57.5311	38.6898	29.2692	23.6168	19.8486	A-Acetyl								43
2	229.0819	115.0446	77.0322	58.0259	46.6222	39.0197	D	4821.4829	2411.2451	1607.8325	1206.1262	965.1024	804.4199	689.6467	42
3	357.1769	179.0921	119.7305	90.0497	72.2412	60.3689	K	4706.4559	2353.7316	1569.4902	1177.3694	942.0970	785.2487	673.2142	41
4	454.2297	227.6185	152.0814	114.3129	91.6518	76.5443	P	4578.3610	2289.6841	1526.7918	1145.3457	916.4780	763.8996	654.9149	40
5	569.2566	285.1319	190.4237	143.0696	114.6571	95.7155	D	4481.3082	2241.1577	1494.4409	1121.0825	897.0675	747.7241	641.0503	39
6	700.2971	350.6522	234.1039	175.8297	140.8652	117.5556	M	4366.2812	2183.6443	1456.0986	1092.3258	874.0621	728.5529	624.6178	38
7	757.3186	379.1629	253.1110	190.0851	152.2695	127.0592	G	4235.2407	2118.1240	1412.4184	1059.5656	847.8540	706.7129	605.8978	37
8	886.3612	443.6842	296.1252	222.3458	178.0781	148.5663	E	4178.2193	2089.6133	1393.4113	1045.3103	836.4497	697.2093	597.7519	36
9	999.4452	500.2263	333.8199	250.6168	200.6949	167.4136	I	4049.1767	2025.0920	1350.3971	1013.0496	810.6412	675.7022	579.3172	35
10	1070.4824	535.7448	357.4990	268.3761	214.9023	179.2531	A	3936.0926	1968.5499	1312.7024	984.7786	788.0243	656.8548	563.1623	34
11	1157.5144	579.2608	386.5097	290.1341	232.3087	193.7585	S	3865.0555	1933.0314	1289.0233	967.0193	773.8169	645.0153	553.0142	33
12	1304.5828	652.7950	435.5325	326.9012	261.7224	218.2699	F	3778.0234	1889.5154	1260.0127	945.2613	756.4105	630.5100	540.5810	32
13	1419.6098	710.3085	473.8748	355.6579	284.7278	237.4410	D	3630.9550	1815.9812	1210.9899	908.4942	726.9968	605.9986	519.5712	31
14	1547.7047	774.3560	516.5731	387.6816	310.3468	258.7902	K	3515.9281	1758.4677	1172.6475	879.7375	703.9914	586.8274	503.1388	30
15	1618.7419	809.8746	540.2521	405.4409	324.5542	270.6297	A	3387.8331	1694.4202	1129.9492	847.7137	678.3724	565.4783	484.8395	29
16	1746.8368	873.9221	582.9505	437.4647	350.1732	291.9789	K	3316.7960	1658.9016	1106.2702	829.9545	664.1650	553.6387	474.6914	28
17	1859.9209	930.4641	620.6452	465.7357	372.7900	310.8262	L	3188.7010	1594.8541	1063.5719	797.9307	638.5460	532.2896	456.3921	27
18	1988.0159	994.5116	663.3435	497.7594	398.4090	332.1754	K	3075.6169	1538.3121	1025.8772	769.6597	615.9292	513.4422	440.2372	26
19	2116.1108	1058.5591	706.0418	529.7832	424.0280	353.5245	K	2947.5220	1474.2646	983.1788	737.6360	590.3102	492.0931	421.9380	25
20	2217.1585	1109.0829	739.7244	555.0451	444.2375	370.3658	T	2819.4270	1410.2171	940.4805	705.6122	564.6912	470.7439	403.6387	24

21	2346.2011	1173.6042	782.7386	587.3057	470.0460	391.8729	**E**	2718.3793	1359.6933	906.7980	680.3503	544.4817	453.9026	389.2033	**23**
22	2447.2488	1224.1280	816.4211	612.5677	490.2556	408.7142	**T**	2589.3367	1295.1720	863.7838	648.0896	518.6732	432.3955	370.7686	**22**
23	2575.3074	1288.1573	859.1073	644.5823	515.8673	430.0573	**Q**	2488.2890	1244.6482	830.1012	622.8277	498.4636	415.5542	356.3332	**21**
24	2704.3500	1352.6786	902.1215	676.8430	541.6758	451.5644	**E**	2360.2305	1180.6189	787.4150	590.8131	472.8519	394.2111	338.0392	**20**
25	2832.4449	1416.7261	944.8198	708.8667	567.2948	472.9136	**K**	2231.1879	1116.0976	744.4008	558.5524	447.0434	372.7040	319.6045	**19**
26	2946.4879	1473.7476	982.8341	737.3774	590.1034	491.9207	**N**	2103.0929	1052.0501	701.7025	526.5287	421.4244	351.3549	301.3052	**18**
27	3047.5356	1524.2714	1016.5167	762.6394	610.3129	508.7620	**T**	1989.0500	995.0286	663.6882	498.0180	398.6158	332.3477	285.0134	**17**
28	3160.6196	1580.8135	1054.2114	790.9104	632.9298	527.6093	**L**	1888.0023	944.5048	630.0056	472.7560	378.4063	315.5064	270.5780	**16**
29	3257.6724	1629.3398	1086.5623	815.1736	652.3403	543.7848	**P**	1774.9182	887.9627	592.3109	444.4850	355.7895	296.6591	254.4231	**15**
30	3358.7201	1679.8637	1120.2449	840.4355	672.5498	560.6261	**T**	1677.8654	839.4364	559.9600	420.2218	336.3789	280.4836	240.5584	**14**
31	3486.8150	1743.9112	1162.9432	872.4592	698.1688	581.9752	**K**	1576.8178	788.9125	526.2774	394.9599	316.1694	263.6424	226.1231	**13**
32	3615.8576	1808.4325	1205.9574	904.7199	723.9774	603.4823	**E**	1448.7228	724.8650	483.5791	362.9362	290.5504	242.2932	207.8238	**12**
33	3716.9053	1858.9563	1239.6400	929.9818	744.1869	620.3236	**T**	1319.6802	660.3437	440.5649	330.6755	264.7419	220.7861	189.3891	**11**
34	3829.9894	1915.4983	1277.3347	958.2528	766.8037	639.1710	**I**	1218.6325	609.8199	406.8824	305.4136	244.5323	203.9448	174.9537	**10**
35	3959.0320	1980.0196	1320.3489	990.5135	792.6122	660.6781	**E**	1105.5484	553.2779	369.1877	277.1426	221.9155	185.0975	158.7989	**9**
36	4087.0906	2044.0489	1363.0350	1022.5281	818.2239	682.0212	**Q**	976.5058	488.7566	326.1735	244.8819	196.1070	163.5904	140.3642	**8**
37	4216.1332	2108.5702	1406.0492	1054.7888	844.0325	703.5283	**E**	848.4473	424.7273	283.4873	212.8673	170.4953	142.2473	122.0701	**7**
38	4344.2281	2172.6177	1448.7476	1086.8125	869.6515	724.8774	**K**	719.4047	360.2060	240.4731	180.6066	144.6868	120.7402	103.6355	**6**
39	4500.3293	2250.6683	1500.7813	1125.8378	900.8717	750.8943	**R**	591.3097	296.1585	197.7748	148.5829	119.0678	99.3910	85.3362	**5**
40	4587.3613	2294.1843	1529.7920	1147.5958	918.2781	765.3996	**S**	435.2086	218.1079	145.7410	109.5576	87.8475	73.3742	63.0360	**4**
41	4716.4039	2358.7056	1572.8062	1179.8564	944.0866	786.9067	**E**	348.1765	174.5919	116.7304	87.7996	70.4411	58.8688	50.6029	**3**
42	4829.4880	2415.2476	1610.5008	1208.1275	966.7034	805.7541	**I**	219.1339	110.0706	73.7162	55.5389	44.6326	37.3617	32.1682	**2**
43							**S**	106.0499	53.5286	36.0215	27.2679	22.0158	18.5144	16.0134	**1**

Fig. S3a. Thymosin β4, fragment 1-41.
$[M+H]^+$ 4745.42599 Da (Theor. 4745.4189). MS/MS analysis performed on different m/z ions in different samples. Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+6H]^{6+}$ monoisotopic m/z: 792.07786 Da. RT: 19.09 min. Analysis manually performed and validate by comparison with the high-resolution MS/MS simulation of the software MS-Product available on the ProteinProspector website (http://prospector.ucsf.edu/prospector/mshome.htm). In green are highlighted the b and y Ions matching with the experimental ones.

Fig. S3b. Thymosin β10, fragment 1-41.
$[M+H]^+$ 4745.42599 Da (Theor. 4734.4141). MS/MS analysis performed on different m/z ions in different samples. Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+6H]^{6+}$ monoisotopic m/z: 790.24310 Da. RT: 19.82 min. Analysis manually performed and validate by comparison with the high-resolution MS/MS simulation of the software MS-Product available on the ProteinProspector website (http://prospector.ucsf.edu/prospector/mshome.htm). In green are highlighted the b and y Ions matching with the experimental ones.

Fig. S3a.



3A_solubile_010714_XT_00001_MHp__1612... 21/12/2016 16:06:40

3A_solubile_010714_XT_00001_MHp__161221160640#2 RT: 2.00
T: FTMS + p ESI d Full ms2 792.08@cid35.00 [205.00-2000.00]
m/z= 561.71160-4657.37170

	m/z	Intensity	Relative	Charge
b5	585.25098	16307.0	11.24	
	625.29333	8875.0	6.12	
	643.30505	19752.0	13.61	
	657.36505	9524.0	6.56	
y6	660.33099	16604.0	11.44	
	696.33081	9078.0	6.26	
	713.86249	17275.0	11.91	
	757.89532	3782.0	2.61	
	762.39307	5763.0	3.97	
	764.64081	6783.0	4.68	
	765.14044	12436.0	8.57	
	767.21814	17114.0	11.80	
	770.90155	8423.0	5.81	
	771.36346	31236.0	21.53	
	826.10254	4201.0	2.90	
	846.43970	5382.0	3.71	
	897.22913	4379.0	3.02	
y8	902.45813	15250.0	10.51	
	914.43469	5130.0	3.54	
	916.37231	21427.0	14.77	
	933.48541	3293.0	2.27	
y9	1004.50586	7806.0	5.38	
	1042.50488	17342.0	11.95	
	1060.06055	3846.0	2.65	
	1118.82922	8306.0	5.73	
y10	1132.54797	10413.0	7.18	
	1158.49890	19022.0	13.11	
	1170.60083	22483.0	15.50	
	1176.52161	5490.0	3.78	
	1252.64978	5079.0	3.50	
b11	1286.59277	3696.0	2.55	
	1298.65886	12662.0	8.73	
	1351.68535	3460.0	2.38	
	1369.69646	29149.0	20.09	
	1427.72331	8616.0	5.94	
y13	1444.72771	145078.0	100.00	
b13	1548.68877	28229.0	19.46	
	1621.84441	3540.0	2.44	
	1654.86455	115350.0	79.51	
	1655.87125	1372.0	0.95	
	1752.88921	3865.0	2.66	
y16	1768.90727	53117.0	36.61	
b16	1891.91147	13254.0	9.14	
y17	1896.99846	4555.0	3.14	
b17	2004.99442	3407.0	2.35	
b18	2133.08951	17583.0	12.12	
b19	2261.18460	16525.0	11.39	
	2477.27617	4813.0	3.32	
	2491.27585	13224.0	9.12	
	2605.35351	14704.0	10.14	
y23	2613.33799	10781.0	7.43	
	2701.38956	9390.0	6.47	
b23	2720.39206	11569.0	7.97	
b24	2831.40946	3976.0	2.74	
	2849.42630	30026.0	20.70	
	2929.53220	11372.0	7.84	
	2959.50747	8256.0	5.69	
	2977.51773	108976.0	75.12	
	3073.55092	56769.0	39.13	
b26	3092.56411	62141.0	42.83	
	3256.69692	4528.0	3.12	
	3266.67251	4655.0	3.21	
	3274.70766	13580.0	9.36	
	3284.68349	20549.0	14.16	
b28	3301.69790	124119.0	85.55	
y30	3459.83925	6899.0	4.76	
	3613.86488	8596.0	5.93	
y32	3716.97343	2347.0	1.62	
	3725.91444	2970.0	2.05	
b32	3742.92036	6045.0	4.17	
	3825.96445	2207.0	1.52	
	4086.09541	8643.0	5.96	
y36	4162.18793	3001.0	2.07	
b36	4214.15598	2774.0	1.91	
	4280.19948	2513.0	1.73	
	4326.15346	3451.0	2.38	
b37	4343.25666	23151.0	15.96	
	4359.23170	7018.0	4.84	
y38	4373.25544	30250.0	20.85	
b38	4471.29560	20152.0	13.89	
	4486.33569	2429.0	1.67	
y39	4501.35143	7186.0	4.95	
	4542.33557	1948.0	1.34	
	4558.35447	1922.0	1.32	
	4563.34728	2938.0	2.03	
	4581.35278	19503.0	13.44	
	4599.34003	11392.0	7.85	
	4635.36083	7732.0	5.33	
	4652.37170	20813.0	14.35	

b		Acetyl		y
130.0499	1	**S**	41	4703.4084
245.0768	2	**D**	40	4616.3764
373.1718	3	**K**	39	4501.3494
470.2245	4	**P**	38	4373.2545
585.2515	5	**D**	37	4276.2017
716.2920	6	**M**	36	4161.1748
787.3291	7	**A**	35	4030.1343
916.3717	8	**E**	34	3959.0972
1029.4557	9	**I**	33	3830.0546
1158.4983	10	**E**	32	3716.9705
1286.5933	11	**K**	31	3587.9279
1433.6617	12	**F**	30	3459.8330
1548.6887	13	**D**	29	3312.7645
1676.7836	14	**K**	28	3197.7376
1763.8156	15	**S**	27	3069.6426
1891.9106	16	**K**	26	2982.6106
2004.9947	17	**L**	25	2854.5156
2133.0896	18	**K**	24	2741.4316
2261.1846	19	**K**	23	2613.3366
2362.2323	20	**T**	22	2485.2417
2491.2749	21	**E**	21	2384.1940
2592.3225	22	**T**	20	2255.1514
2720.3811	23	**Q**	19	2154.1037
2849.4237	24	**E**	18	2026.0451
2977.5187	25	**K**	17	1897.0025
3091.5616	26	**N**	16	1768.9076
3188.6144	27	**P**	15	1654.8646
3301.6984	28	**L**	14	1557.8119
3398.7512	29	**P**	13	1444.7278
3485.7832	30	**S**	12	1347.6750
3613.8782	31	**K**	11	1260.6430
3742.9208	32	**E**	10	1132.5481
3843.9685	33	**T**	9	1003.5055
3957.0525	34	**I**	8	902.4578
4086.0951	35	**E**	7	789.3737
4214.1537	36	**Q**	6	660.3311
4343.1963	37	**E**	5	532.2726
4471.2913	38	**K**	4	403.2300
4599.3498	39	**Q**	3	275.1350
4670.3869	40	**A**	2	147.0764
---	41	**G**	1	76.0393

MH **4745.4190**
MH-H_2O **4727.4084**
MH-NH_3 **4728.3924**

Fig. S3b.

3A_solubile_010714_XT_00001_MHp__1612... 21/12/2016 16:58:24

3A_solubile_010714_XT_00001_MHp__161221165824#2 RT: 2.00
T: FTMS + p ESI d Full ms2 790.24@cid35.00 [205.00-2000.00]
m/z= 514.28766-4625.31610

	m/z	Intensity	Relative	Charge
y4	519.28766	8274.0	33.58	
b5	569.25543	3683.0	14.95	
y5	648.33014	2610.0	10.59	
	733.37329	1967.0	7.98	
b7	757.31689	1540.0	6.25	
	759.36108	2669.0	10.83	
	767.41461	1708.0	6.93	
	773.31244	8983.0	36.46	
	776.38837	3684.0	14.95	
	826.42950	1981.0	8.04	
	868.34570	2343.0	9.51	
	873.05072	3000.0	12.18	
b8	886.35999	24636.0	100.00	
y7	905.43048	6546.0	26.57	
	926.49225	1395.0	5.66	
	993.54510	996.0	4.04	
b9	999.44302	4578.0	18.58	
y8	1018.51410	2451.0	9.95	
	1056.51709	3790.0	15.38	
y9	1119.56212	7459.0	30.28	
y10	1248.60633	4973.0	20.19	
	1322.70452	799.0	3.24	
	1339.74468	1959.0	7.95	
y11	1376.69939	2762.0	11.21	
b13	1419.60718	2526.0	10.25	
	1445.76152	923.0	3.75	
	1469.72099	627.0	2.55	
y12	1477.74480	1915.0	7.77	
	1488.73454	581.0	2.36	
	1538.77727	2210.0	8.97	
	1556.78935	3641.0	14.78	
y13	1574.80094	1474.0	5.98	
y14	1687.88664	3295.0	13.37	
	1759.93755	941.0	3.82	
y16	1902.97366	10742.0	43.60	
	1904.98064	1765.0	7.16	
	1990.02275	1680.0	6.82	
	2091.14897	1519.0	6.17	
	2108.07866	801.0	3.25	
b19	2116.10611	3188.0	12.94	
	2162.18486	1641.0	6.66	
y22	2618.34317	2141.0	8.69	
	2620.31064	1768.0	7.18	
	2687.31296	485.0	1.97	
b24	2704.34049	2643.0	10.73	
y23	2747.39518	2725.0	11.06	
b25	2832.44344	4290.0	17.41	
b26	2946.48354	626.0	2.54	
b27	3047.53017	1194.0	4.85	
y28	3315.80488	3010.0	12.22	
Y30	3577.91054	2943.0	11.95	
Y31	3664.93374	10112.0	41.05	
b33	3717.95644	3384.0	13.74	
y32	3735.97158	15704.0	63.74	
Y33	3849.04922	5500.0	22.33	
Y38	4378.23225	9303.0	37.76	
y39	4506.35021	1767.0	7.17	

b			Acetyl		y
114.0550	1		**A**	41	4692.4037
229.0819	2		**D**	40	4621.3665
357.1769	3		**K**	39	4506.3396
454.2296	4		**P**	38	4378.2446
569.2566	5		**D**	37	4281.1919
700.2971	6		**M**	36	4166.1649
757.3185	7		**G**	35	4035.1244
886.3611	8		**E**	34	3978.1030
999.4452	9		**I**	33	3849.0604
1070.4823	10		**A**	32	3735.9763
1157.5143	11		**S**	31	3664.9392
1304.5827	12		**F**	30	3577.9072
1419.6097	13		**D**	29	3430.8388
1547.7046	14		**K**	28	3315.8118
1618.7417	15		**A**	27	3187.7169
1746.8367	16		**K**	26	3116.6797
1859.9208	17		**L**	25	2988.5848
1988.0157	18		**K**	24	2875.5007
2116.1107	19		**K**	23	2747.4058
2217.1584	20		**T**	22	2619.3108
2346.2010	21		**E**	21	2518.2631
2447.2487	22		**T**	20	2389.2205
2575.3072	23		**Q**	19	2288.1728
2704.3498	24		**E**	18	2160.1143
2832.4448	25		**K**	17	2031.0717
2946.4877	26		**N**	16	1902.9767
3047.5354	27		**T**	15	1788.9338
3160.6195	28		**L**	14	1687.8861
3257.6722	29		**P**	13	1574.8020
3358.7199	30		**T**	12	1477.7493
3486.8149	31		**K**	11	1376.7016
3615.8575	32		**E**	10	1248.6066
3716.9051	33		**T**	9	1119.5640
3829.9892	34		**I**	8	1018.5164
3959.0318	35		**E**	7	905.4323
4087.0904	36		**Q**	6	776.3897
4216.1330	37		**E**	5	648.3311
4344.2279	38		**K**	4	519.2885
4500.3290	39		**R**	3	391.1936
4587.3611	40		**S**	2	235.0925
---	41		**E**	1	148.0604

MH	4734.4142
MH-H_2O	4716.4037
MH-NH_3	4717.3877

Fig. S4. Thymosin β4, diacetylated on Lys16 and Lys25.
$[M+H]^+$ 5045.5183 Da (Theor. 5045.5146). MS/MS analysis performed on different m/z ions in different samples.
a) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+5H]^{5+}$ monoisotopic m/z: 1010.51 Da. RT: 22.23 min.
b) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+6H]^{6+}$ monoisotopic m/z: 842.09 Da. RT: 21.36 min.
c) Analysis manually performed and validate by comparison with the high-resolution MS/MS simulation of the software MS-Product available on the ProteinProspector website (http://prospector.ucsf.edu/prospector/mshome.htm). In green are highlighted the b and y Ions matching with the experimental ones.

a)

Caso3A_180314_XT_00001_MHp_ 22/12/2016 09:16:22

Caso3A_180314_XT_00001_MHp_#2 RT: 2.00
T: FTMS + p ESI d Full ms2 1010.51@cid35.00 [265.00-2000.00]
m/z= 946.688-4979.467

	m/z	Intensity	Relative	Charge
	951.688	994.0	5.13	
	953.489	1496.0	7.72	
	958.694	1348.0	6.96	
	979.487	1167.0	6.02	
	982.298	2791.0	14.41	
	983.900	1390.0	7.18	
	987.438	2198.0	11.35	
	987.692	1494.0	7.71	
	992.302	1227.0	6.33	
	1102.581	1142.0	5.90	
	1103.052	1005.0	5.19	
	1160.083	1044.0	5.39	
	1185.606	1259.0	6.50	
b13	1548.685	704.0	3.63	
y16	1871.942	782.0	4.04	
	1963.985	1239.0	6.40	
y18	1984.988	2066.0	10.67	
	3302.694	922.0	4.76	
	3385.709	2486.0	12.83	
b35	4170.126	1664.0	8.59	
b36	4298.164	1137.0	5.87	
b37	4427.199	2625.0	13.55	
b38	4554.302	1777.0	9.17	
	4570.310	804.0	4.15	
	4589.319	3370.0	17.40	
	4655.338	2185.0	11.28	
	4665.342	1449.0	7.48	
y40	4673.356	3417.0	17.64	
	4682.365	1652.0	8.53	
	4736.392	934.0	4.82	
	4754.395	5240.0	27.05	
	4784.427	587.0	3.03	
y41	4801.434	651.0	3.36	
	4812.431	2240.0	11.56	
	4903.438	1132.0	5.84	
	4911.461	590.0	3.05	
y42	4916.457	599.0	3.09	
	4923.452	5107.0	26.36	
	4935.463	977.0	5.04	
b42	4940.469	19371.0	100.00	
	4974.467	417.0	2.15	

b)

Caso3A_180314_XT_00001_MHp__161222105903 22/12/2016 10:59:03

Caso3A_180314_XT_00001_MHp__161222105903#2 RT: 2.00
F: FTMS + p ESI d Full ms2 842.09@cid35.00 [220.00-2000.00]
m/z= 614.305-4976.481

	m/z	Intensity	Relative	Charge
	619.305	3693.0	22.26	
	709.350	1464.0	8.82	
	748.346	2705.0	16.30	
	810.909	1578.0	9.51	
	816.918	1667.0	10.05	
	820.252	1893.0	11.41	
	827.174	995.0	6.00	
y8	876.405	2452.0	14.78	
	892.455	931.0	5.61	
b8	916.371	2094.0	12.62	
	930.282	1037.0	6.25	
	933.474	1074.0	6.47	
	956.685	1231.0	7.42	
	982.298	1875.0	11.30	
	982.491	1356.0	8.17	
	984.490	1826.0	11.00	
	987.296	1111.0	6.70	
y9	1005.447	3471.0	20.92	
	1005.525	1990.0	11.99	
	1548.690	2080.0	12.54	
	1555.759	4164.0	25.09	
	1572.826	1862.0	11.22	
	1586.808	714.0	4.30	
	1604.817	706.0	4.25	
	1651.851	1234.0	7.44	
	1655.841	2024.0	12.20	
y15	1660.801	16593.0	100.00	
	1694.366	968.0	5.83	
	1765.895	796.0	4.80	
	1852.924	1782.0	10.74	
	1869.941	6499.0	39.17	
y16	1871.942	6020.0	36.28	
y18	1984.983	3807.0	22.94	
	2440.238	869.0	5.24	
	3044.530	2927.0	17.64	
b25	3061.539	5822.0	35.09	
	3139.561	1890.0	11.39	
	3157.572	5099.0	30.73	
b26	3175.580	8932.0	53.83	
b27	3272.635	2331.0	14.05	
b28	3385.722	4212.0	25.38	
b33	3928.987	471.0	2.84	
b34	4041.072	699.0	4.21	
	4410.215	553.0	3.33	
	4427.198	1668.0	10.05	
	4655.329	1646.0	9.92	
b39	4683.380	468.0	2.82	
	4738.399	955.0	5.76	
b40	4754.416	1885.0	11.36	
	4793.410	572.0	3.45	
b41	4811.409	1018.0	6.14	
	4887.430	614.0	3.70	
	4897.466	512.0	3.09	
	4905.434	1721.0	10.37	
	4911.459	981.0	5.91	
	4918.484	696.0	4.19	
	4923.453	6685.0	40.29	

c)

b		Acetyl		y
130.0499	1	**S**	43	5003.5042
245.0768	2	**D**	42	4916.4721
373.1718	3	**K**	41	4801.4452
470.2245	4	**P**	40	4673.3502
585.2515	5	**D**	39	4576.2975
716.2920	6	**M**	38	4461.2705
787.3291	7	**A**	37	4330.2300
916.3717	8	**E**	36	4259.1929
1029.4557	9	**I**	35	4130.1503
1158.4983	10	**E**	34	4017.0663
1286.5933	11	**K**	33	3888.0237
1433.6617	12	**F**	32	3759.9287
1548.6887	13	**D**	31	3612.8603
1676.7836	14	**K**	30	3497.8333
1763.8156	15	**S**	29	3369.7384
1933.9212	16	**K(Acetyl)**	28	3282.7064
2047.0052	17	**L**	27	3112.6008
2175.1002	18	**K**	26	2999.5168
2303.1952	19	**K**	25	2871.4218
2404.2428	20	**T**	24	2743.3268
2533.2854	21	**E**	23	2642.2792
2634.3331	22	**T**	22	2513.2366
2762.3917	23	**Q**	21	2412.1889
2891.4343	24	**E**	20	2284.1303
3061.5398	25	**K(Acetyl)**	19	2155.0877
3175.5827	26	**N**	18	1984.9822
3272.6355	27	**P**	17	1870.9393
3385.7196	28	**L**	16	1773.8865
3482.7723	29	**P**	15	1660.8024
3569.8044	30	**S**	14	1563.7497
3697.8993	31	**K**	13	1476.7176
3826.9419	32	**E**	12	1348.6227
3927.9896	33	**T**	11	1219.5801
4041.0737	34	**I**	10	1118.5324
4170.1162	35	**E**	9	1005.4483
4298.1748	36	**Q**	8	876.4058
4427.2174	37	**E**	7	748.3472
4555.3124	38	**K**	6	619.3046
4683.3710	39	**Q**	5	491.2096
4754.4081	40	**A**	4	363.1510
4811.4295	41	**G**	3	292.1139
4940.4721	42	**E**	2	235.0925

MH	5045.5147
MH-H_2O	5027.5042
MH-NH_3	5028.4882

Fig. S5a. Thymosin β4, Met6-ox.
$[M+H]^+$ 4977.4934 Da (Theor. 4977.4885). MS/MS analysis performed on different m/z ions in different samples.
- Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+6H]^{6+}$ monoisotopic m/z: 830.42163 Da (+0.78 mmu/+0.93 ppm). RT: 18.41 min. Analysis performed by Proteome discoverer software. XCorr:4.81.
S1-Acetyl (42.01057 Da), M6-Oxidation (15.9949 Da)
Identified with: Sequest HT (v1.3); XCorr:4.89, Fragment match tolerance used for search: 0.6 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Thymosin beta-4 OS=Homo sapiens GN=TMSB4X PE=1 SV=2 - [TYB4_HUMAN]

#1	b	b^2	b^3	b	b	b	Seq.	y	y^2	y^3	y	y	y	#2
1	130.0499	65.5286	44.0215	33.2679	26.8158	22.5144	S-Acetyl							43
2	245.0768	123.0421	82.3638	62.0247	49.8212	41.6855	D	4848.4462	2424.7267	1616.8202	1212.8670	970.4951	808.9138	42
3	373.1718	187.0895	125.0621	94.0484	75.4402	63.0347	K	4733.4192	2367.2132	1578.4779	1184.1103	947.4897	789.7426	41
4	470.2246	235.6159	157.4130	118.3116	94.8507	79.2102	P	4605.3242	2303.1658	1535.7796	1152.0865	921.8707	768.3934	40
5	585.2515	293.1294	195.7554	147.0683	117.8561	98.3813	D	4508.2715	2254.6394	1503.4287	1127.8233	902.4601	752.2180	39
6	732.2869	366.6471	244.7672	183.8272	147.2632	122.8872	M-Oxidation	4393.2445	2197.1259	1465.0864	1099.0666	879.4547	733.0468	38
7	803.3241	402.1657	268.4462	201.5865	161.4706	134.7267	A	4246.2091	2123.6082	1416.0746	1062.3077	850.0476	708.5409	37
8	932.3667	466.6870	311.4604	233.8471	187.2792	156.2338	E	4175.1720	2088.0896	1392.3955	1044.5485	835.8402	696.7014	36
9	1045.4507	523.2290	349.1551	262.1181	209.8960	175.0812	I	4046.1294	2023.5683	1349.3813	1012.2878	810.0317	675.1943	35
10	1174.4933	587.7503	392.1693	294.3788	235.7045	196.5883	E	3933.0453	1967.0263	1311.6866	984.0168	787.4149	656.3470	34
11	1302.5883	651.7978	434.8676	326.4025	261.3235	217.9374	K	3804.0027	1902.5050	1268.6724	951.7561	761.6064	634.8399	33
12	1449.6567	725.3320	483.8904	363.1696	290.7372	242.4489	F	3675.9077	1838.4575	1225.9741	919.7324	735.9874	613.4907	32
13	1564.6837	782.8455	522.2327	391.9264	313.7426	261.6200	D	3528.8393	1764.9233	1176.9513	882.9653	706.5737	588.9793	31
14	1692.7786	846.8930	564.9311	423.9501	339.3616	282.9692	K	3413.8124	1707.4098	1138.6090	854.2086	683.5683	569.8081	30
15	1779.8107	890.4090	593.9417	445.7081	356.7680	297.4745	S	3285.7174	1643.3623	1095.9107	822.1848	657.9493	548.4590	29
16	1907.9056	954.4565	636.6401	477.7319	382.3870	318.8237	K	3198.6854	1599.8463	1066.9000	800.4268	640.5429	533.9536	28
17	2020.9897	1010.9985	674.3348	506.0029	405.0038	337.6710	L	3070.5904	1535.7988	1024.2017	768.4031	614.9239	512.6045	27
18	2149.0847	1075.0460	717.0331	538.0266	430.6228	359.0202	K	2957.5063	1479.2568	986.5070	740.1320	592.3071	493.7571	26
19	2277.1796	1139.0935	759.7314	570.0504	456.2418	380.3693	K	2829.4114	1415.2093	943.8086	708.1083	566.6881	472.4080	25
20	2378.2273	1189.6173	793.4140	595.3123	476.4513	397.2106	T	2701.3164	1351.1618	901.1103	676.0846	541.0691	451.0588	24
21	2507.2699	1254.1386	836.4282	627.5729	502.2598	418.7177	E	2600.2687	1300.6380	867.4278	650.8226	520.8596	434.2175	23
22	2608.3176	1304.6624	870.1107	652.8349	522.4693	435.5590	T	2471.2261	1236.1167	824.4136	618.5620	495.0510	412.7104	22
23	2736.3762	1368.6917	912.7969	684.8495	548.0811	456.9021	Q	2370.1784	1185.5929	790.7310	593.3001	474.8415	395.8691	21
24	2865.4188	1433.2130	955.8111	717.1102	573.8896	478.4092	E	2242.1199	1121.5636	748.0448	561.2854	449.2298	374.5260	20
25	2993.5138	1497.2605	998.5094	749.1339	599.5086	499.7584	K	2113.0773	1057.0423	705.0306	529.0248	423.4213	353.0189	19
26	3107.5567	1554.2820	1036.5237	777.6446	622.3172	518.7655	N	1984.9823	992.9948	662.3323	497.0010	397.8023	331.6698	18
27	3204.6095	1602.8084	1068.8747	801.9078	641.7277	534.9410	P	1870.9394	935.9733	624.3180	468.4903	374.9937	312.6626	17
28	3317.6935	1659.3504	1106.5694	830.1788	664.3445	553.7883	L	1773.8866	887.4469	591.9670	444.2271	355.5831	296.4872	16
29	3414.7463	1707.8768	1138.9203	854.4420	683.7551	569.9638	P	1660.8025	830.9049	554.2724	415.9561	332.9663	277.6398	15
30	3501.7783	1751.3928	1167.9310	876.2000	701.1615	584.4691	S	1563.7497	782.3785	521.9214	391.6929	313.5558	261.4644	14
31	3629.8733	1815.4403	1210.6293	908.2238	726.7805	605.8183	K	1476.7177	738.8625	492.9108	369.9349	296.1494	246.9590	13
32	3758.9159	1879.9616	1253.6435	940.4844	752.5890	627.3254	E	1348.6227	674.8150	450.2124	337.9111	270.5304	225.6099	12
33	3859.9636	1930.4854	1287.3260	965.7464	772.7985	644.1667	T	1219.5801	610.2937	407.1982	305.6505	244.7219	204.1028	11
34	3973.0476	1987.0275	1325.0207	994.0174	795.4154	663.0140	I	1118.5325	559.7699	373.5157	280.3886	224.5123	187.2615	10
35	4102.0902	2051.5488	1368.0349	1026.2780	821.2239	684.5211	E	1005.4484	503.2278	335.8210	252.1176	201.8955	168.4141	9
36	4230.1488	2115.5781	1410.7211	1058.2927	846.8356	705.8642	Q	876.4058	438.7065	292.8068	219.8569	176.0870	146.9070	8
37	4359.1914	2180.0994	1453.7353	1090.5533	872.6441	727.3713	E	748.3472	374.6772	250.1206	187.8423	150.4753	125.5639	7
38	4487.2864	2244.1468	1496.4337	1122.5771	898.2631	748.7205	K	619.3046	310.1559	207.1064	155.5816	124.6667	104.0568	6
39	4615.3450	2308.1761	1539.1198	1154.5917	923.8748	770.0636	Q	491.2096	246.1085	164.4081	123.5579	99.0478	82.7077	5
40	4686.3821	2343.6947	1562.7989	1172.3510	938.0822	781.9031	A	363.1511	182.0792	121.7219	91.5432	73.4360	61.3646	4
41	4743.4036	2372.2054	1581.8060	1186.6064	949.4865	791.4067	G	292.1139	146.5606	98.0428	73.7839	59.2286	49.5251	3
42	4872.4462	2436.7267	1624.8202	1218.8670	975.2951	812.9138	E	235.0925	118.0499	79.0357	59.5286	47.8243	40.0215	2
43							S	106.0499	53.5286	36.0215	27.2679	22.0158	18.5144	1

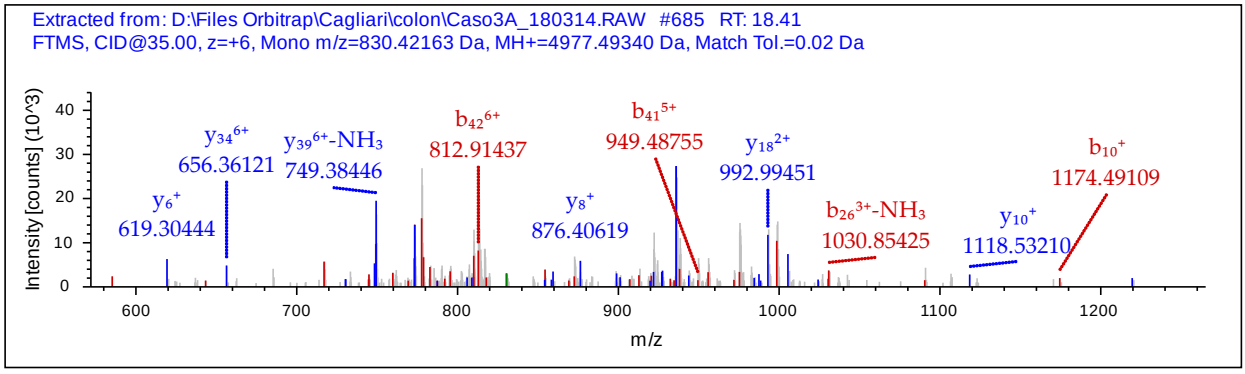


Fig. S5b. Thymosin β10, Met6-ox.
$[M+H]^+$ 4950.5282 Da (Theor. 4950.5252). MS/MS analysis performed on different m/z ions in different samples.
- Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+6H]^{6+}$ monoisotopic m/z: 825.92725 Da (+0.27 mmu/+0.33 ppm). RT: 19.53 min. Analysis performed by Proteome discoverer software. XCorr:3.88.
Fragment match tolerance used for search: 0.6 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Thymosin beta-10 OS=Homo sapiens GN=TMSB10 PE=1 SV=2 - [TYB10_HUMAN]

#1	b	b²	b³	b	b	b	Seq.	y	y²	y³	y	y	y	#2
1	114.0550	57.5311	38.6898	29.2692	23.6168	19.8486	A-Acetyl							43
2	229.0819	115.0446	77.0322	58.0259	46.6222	39.0197	D	4837.4778	2419.2425	1613.1641	1210.1249	968.3014	807.0857	42
3	357.1769	179.0921	119.7305	90.0497	72.2412	60.3689	K	4722.4508	2361.7291	1574.8218	1181.3682	945.2960	787.9145	41
4	454.2297	227.6185	152.0814	114.3129	91.6518	76.5443	P	4594.3559	2297.6816	1532.1235	1149.3444	919.6770	766.5654	40
5	569.2566	285.1319	190.4237	143.0696	114.6571	95.7155	D	4497.3031	2249.1552	1499.7726	1125.0812	900.2664	750.3899	39
6	716.2920	358.6497	239.4355	179.8285	144.0642	120.2214	M-Oxidation	4382.2762	2191.6417	1461.4302	1096.3245	877.2611	731.2188	38
7	773.3135	387.1604	258.4427	194.0838	155.4685	129.7250	G	4235.2407	2118.1240	1412.4184	1059.5656	847.8540	706.7129	37
8	902.3561	451.6817	301.4569	226.3445	181.2770	151.2321	E	4178.2193	2089.6133	1393.4113	1045.3103	836.4497	697.2093	36
9	1015.4402	508.2237	339.1516	254.6155	203.8939	170.0794	I	4049.1767	2025.0920	1350.3971	1013.0496	810.6412	675.7022	35
10	1086.4773	543.7423	362.8306	272.3748	218.1013	181.9189	A	3936.0926	1968.5499	1312.7024	984.7786	788.0243	656.8548	34
11	1173.5093	587.2583	391.8413	294.1328	235.5077	196.4243	S	3865.0555	1933.0314	1289.0233	967.0193	773.8169	645.0153	33
12	1320.5777	660.7925	440.8641	330.8999	264.9214	220.9357	F	3778.0234	1889.5154	1260.0127	945.2613	756.4105	630.5100	32
13	1435.6047	718.3060	479.2064	359.6566	287.9268	240.1068	D	3630.9550	1815.9812	1210.9899	908.4942	726.9968	605.9986	31
14	1563.6997	782.3535	521.9047	391.6804	313.5458	261.4560	K	3515.9281	1758.4677	1172.6475	879.7375	703.9914	586.8274	30
15	1634.7368	817.8720	545.5838	409.4397	327.7532	273.2955	A	3387.8331	1694.4202	1129.9492	847.7137	678.3724	565.4783	29
16	1762.8317	881.9195	588.2821	441.4634	353.3722	294.6447	K	3316.7960	1658.9016	1106.2702	829.9545	664.1650	553.6387	28
17	1875.9158	938.4615	625.9768	469.7344	375.9890	313.4920	L	3188.7010	1594.8541	1063.5719	797.9307	638.5460	532.2896	27
18	2004.0108	1002.5090	668.6751	501.7582	401.6080	334.8412	K	3075.6169	1538.3121	1025.8772	769.6597	615.9292	513.4422	26
19	2132.1058	1066.5565	711.3734	533.7819	427.2270	356.1904	K	2947.5220	1474.2646	983.1788	737.6360	590.3102	492.0931	25
20	2233.1534	1117.0804	745.0560	559.0438	447.4365	373.0316	T	2819.4270	1410.2171	940.4805	705.6122	564.6912	470.7439	24
21	2362.1960	1181.6017	788.0702	591.3045	473.2450	394.5387	E	2718.3793	1359.6933	906.7980	680.3503	544.4817	453.9026	23
22	2463.2437	1232.1255	821.7528	616.5664	493.4546	411.3800	T	2589.3367	1295.1720	863.7838	648.0896	518.6732	432.3955	22
23	2591.3023	1296.1548	864.4390	648.5810	519.0663	432.7231	Q	2488.2890	1244.6482	830.1012	622.8277	498.4636	415.5542	21
24	2720.3449	1360.6761	907.4532	680.8417	544.8748	454.2302	E	2360.2305	1180.6189	787.4150	590.8131	472.8519	394.2111	20
25	2848.4399	1424.7236	950.1515	712.8654	570.4938	475.5794	K	2231.1879	1116.0976	744.4008	558.5524	447.0434	372.7040	19
26	2962.4828	1481.7450	988.1658	741.3762	593.3024	494.5865	N	2103.0929	1052.0501	701.7025	526.5287	421.4244	351.3549	18
27	3063.5305	1532.2689	1021.8483	766.6381	613.5119	511.4278	T	1989.0500	995.0286	663.6882	498.0180	398.6158	332.3477	17
28	3176.6145	1588.8109	1059.5430	794.9091	636.1287	530.2752	L	1888.0023	944.5048	630.0056	472.7560	378.4063	315.5064	16
29	3273.6673	1637.3373	1091.8940	819.1723	655.5393	546.4506	P	1774.9182	887.9627	592.3109	444.4850	355.7895	296.6591	15
30	3374.7150	1687.8611	1125.5765	844.4342	675.7488	563.2919	T	1677.8654	839.4364	559.9600	420.2218	336.3789	280.4836	14
31	3502.8100	1751.9086	1168.2748	876.4580	701.3678	584.6411	K	1576.8178	788.9125	526.2774	394.9599	316.1694	263.6424	13
32	3631.8526	1816.4299	1211.2890	908.7186	727.1763	606.1482	E	1448.7228	724.8650	483.5791	362.9362	290.5504	242.2932	12
33	3732.9002	1866.9538	1244.9716	933.9805	747.3859	622.9894	T	1319.6802	660.3437	440.5649	330.6755	264.7419	220.7861	11
34	3845.9843	1923.4958	1282.6663	962.2515	770.0027	641.8368	I	1218.6325	609.8199	406.8824	305.4136	244.5323	203.9448	10
35	3975.0269	1988.0171	1325.6805	994.5122	795.8112	663.3439	E	1105.5484	553.2779	369.1877	277.1426	221.9155	185.0975	9
36	4103.0855	2052.0464	1368.3667	1026.5268	821.4229	684.6870	Q	976.5058	488.7566	326.1735	244.8819	196.1070	163.5904	8
37	4232.1281	2116.5677	1411.3809	1058.7875	847.2314	706.1941	E	848.4473	424.7273	283.4873	212.8673	170.4953	142.2473	7
38	4360.2231	2180.6152	1454.0792	1090.8112	872.8504	727.5432	K	719.4047	360.2060	240.4731	180.6066	144.6868	120.7402	6
39	4516.3242	2258.6657	1506.1129	1129.8365	904.0707	753.5601	R	591.3097	296.1585	197.7748	148.5829	119.0678	99.3910	5
40	4603.3562	2302.1817	1535.1236	1151.5945	921.4771	768.0654	S	435.2086	218.1079	145.7410	109.5576	87.8475	73.3742	4
41	4732.3988	2366.7030	1578.1378	1183.8552	947.2856	789.5725	E	348.1765	174.5919	116.7304	87.7996	70.4411	58.8688	3
42	4845.4829	2423.2451	1615.8325	1212.1262	969.9024	808.4199	I	219.1339	110.0706	73.7162	55.5389	44.6326	37.3617	2
43							S	106.0499	53.5286	36.0215	27.2679	22.0158	18.5144	1

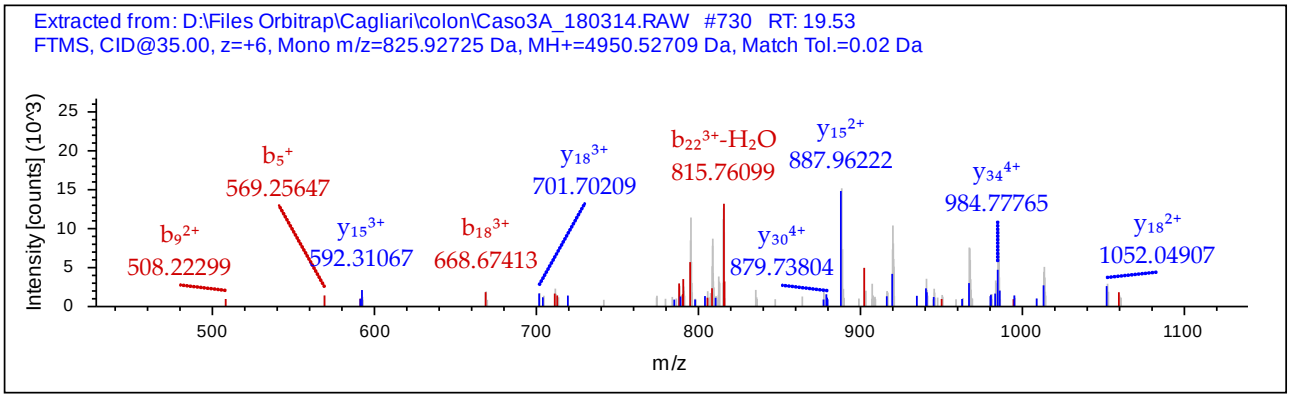


Fig. S6. Pro-thymosin α (3-111).
$[M+H]^+$ 11978.95 Da (Theor. 11978.90). MS/MS analysis performed on different m/z ions in different samples.
a) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+13H]^{13+}$ monoisotopic m/z: 922.85 Da. RT: 20.77 min.
b) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+12H]^{12+}$ monoisotopic m/z: 999.67 Da. RT: 20.88 min.
c) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+14H]^{14+}$ monoisotopic m/z: 857.00 Da. RT: 20.80 min.
d) Analysis manually performed and validate by comparison with the high-resolution MS/MS simulation of the software MS-Product available on the ProteinProspector website (http://prospector.ucsf.edu/prospector/mshome.htm). In green are highlighted the b and y Ions matching with the experimental ones.

a)

3A_solubile_010714_XT_00001_MHp__1612... 22/12/2016 11:55:43

3A_solubile_010714_XT_00001_MHp__161222115543#2 RT: 2.00
F: FTMS + p ESI d Full ms2 922.85@cid35.00 [240.00-2000.00]
m/z= 365.124-6293.723

	m/z	Intensity	Relative	Charge
	370.124	2444.0	14.59	
	631.220	3111.0	18.57	
	646.220	2831.0	16.90	
	760.262	6391.0	38.15	
	761.243	2191.0	13.08	
	775.261	2960.0	17.67	
	779.360	1142.0	6.82	
	822.214	1274.0	7.60	
	825.380	1092.0	6.52	
	848.801	2139.0	12.77	
	860.315	2588.0	15.45	
	870.813	2091.0	12.48	
	872.817	1344.0	8.02	
	875.290	2871.0	17.14	
	889.305	6368.0	38.01	
	896.240	2002.0	11.95	
	931.427	2133.0	12.73	
	943.467	1331.0	7.94	
	944.180	3278.0	19.57	
	958.216	2615.0	15.61	
	963.420	1655.0	9.88	
	986.318	1470.0	8.77	
	1004.332	2963.0	17.69	
	1018.347	6820.0	40.71	
	1068.477	2906.0	17.35	
	1090.368	1813.0	10.82	
	1101.815	1683.0	10.05	
	1102.436	1244.0	7.43	
	1133.376	2195.0	13.10	
	1147.391	8501.0	50.74	
	1262.416	1420.0	8.48	
	1276.433	2750.0	16.41	
	1319.437	1087.0	6.49	
	1736.497	664.0	3.96	
	2193.031	782.0	4.67	
b32	3488.682	6167.0	36.81	
y32	3570.458	1456.0	8.69	
y34	3814.568	1267.0	7.56	
y35	3871.590	12108.0	72.27	
b37	3942.948	2775.0	16.56	
	3968.581	1285.0	7.67	
y36	3986.619	16754.0	100.00	
	3996.886	580.0	3.46	
b38	4012.924	10268.0	61.29	
	4097.652	6129.0	36.58	
y37	4115.655	14394.0	85.91	
	4208.668	523.0	3.12	
	4226.691	6541.0	39.04	
y38	4244.700	13747.0	82.05	
b40	4257.008	5173.0	30.88	
y39	4374.730	856.0	5.11	
b41	4386.046	1150.0	6.86	
	4412.744	1217.0	7.26	
y40	4430.759	10210.0	60.94	
	4431.769	438.0	2.61	
	4501.078	765.0	4.57	
y42	4602.798	1274.0	7.60	
	4603.822	698.0	4.17	
	4668.128	1074.0	6.41	
	4685.151	840.0	5.01	
	4686.153	958.0	5.72	
y43	4731.839	9220.0	55.03	
	4842.891	1229.0	7.34	
y44	4860.891	5363.0	32.01	
	4908.199	570.0	3.40	
	4925.231	2452.0	14.64	
	4943.262	2574.0	15.36	
	4972.945	550.0	3.28	
Y45	4989.949	2677.0	15.98	
	5014.316	1670.0	9.97	
b48	5128.311	722.0	4.31	
	5354.389	6521.0	38.92	
b50	5372.404	5257.0	31.38	
	5453.446	1517.0	9.05	
b51	5471.459	1748.0	10.43	
	5693.170	1322.0	7.89	
	5715.547	1366.0	8.15	
	5750.192	2680.0	16.00	
Y53	5878.274	1059.0	6.32	
b55	5972.636	1113.0	6.64	
b56	6102.673	831.0	4.96	
	6213.684	649.0	3.87	
b57	6231.683	2544.0	15.18	
b58	6288.723	1230.0	7.34	

b)

3A_solubile_010714_XT_00001_MHp__1612... 22/12/2016 12:01:15

3A_solubile_010714_XT_00001_MHp__161222120115#2 RT: 2.00
F: FTMS + p ESI d Full ms2 999.67@cid35.00 [265.00-2000.00]
m/z= 626.220-7382.089

	m/z	Intensity	Relative	Charge
	631.220	2671.0	10.64	
	742.251	2060.0	8.20	
	746.247	1887.0	7.51	
	760.262	4262.0	16.97	
	818.265	1498.0	5.97	
	857.278	1296.0	5.16	
	875.288	2302.0	9.17	
	889.305	5437.0	21.65	
	896.573	2635.0	10.49	
	957.437	1246.0	4.96	
	958.215	2972.0	11.83	
	1014.775	1038.0	4.13	
	1018.348	4680.0	18.64	
	1027.266	1388.0	5.53	
	1027.468	2051.0	8.17	
	1090.368	1259.0	5.01	
	1098.016	2657.0	10.58	
	1134.379	1653.0	6.58	
	1136.436	1714.0	6.83	
	1139.442	1480.0	5.89	
	1147.394	8924.0	35.54	
	1150.642	1998.0	7.96	
	1192.131	1927.0	7.67	
	1258.419	1339.0	5.33	
	1277.436	1538.0	6.12	
	1405.470	1570.0	6.25	
y12	1421.671	1001.0	3.99	
b32	3488.689	10279.0	40.93	
b34	3656.755	2846.0	11.33	
y33	3699.541	1260.0	5.02	
b35	3771.819	2267.0	9.03	
y34	3814.551	1565.0	6.23	
y35	3871.585	9476.0	37.73	
b38	4012.919	5294.0	21.08	
	4098.650	1504.0	5.99	
y37	4115.657	9106.0	36.26	
	4226.692	3740.0	14.89	
y38	4244.696	12348.0	49.17	
b40	4257.009	6669.0	26.56	
b41	4385.041	2902.0	11.56	
y40	4430.764	10435.0	41.55	
	4500.084	1269.0	5.05	
y41	4545.787	4011.0	15.97	
	4546.793	1665.0	6.63	
	4556.110	863.0	3.44	
	4585.799	964.0	3.84	
y42	4602.791	25113.0	100.00	
	4668.139	2562.0	10.20	
	4685.152	3445.0	13.72	
	4713.835	685.0	2.73	
y43	4732.856	7837.0	31.21	
	4796.177	708.0	2.82	
	4842.882	1056.0	4.20	
Y44	4860.889	5509.0	21.94	
	4890.180	578.0	2.30	
y46	5118.982	2676.0	10.66	
	5230.019	1006.0	4.01	
y47	5249.028	2320.0	9.24	
y48	5337.352	1514.0	6.03	
	5354.391	6111.0	24.33	
b50	5372.409	7229.0	28.79	
	5377.099	4178.0	16.64	
	5453.446	2746.0	10.93	
b51	5470.694	3885.0	15.47	
	5471.440	1823.0	7.26	
	5505.032	3596.0	14.32	
	5691.167	1918.0	7.64	
	5715.532	1516.0	6.04	
y52	5749.183	6495.0	25.86	
b54	5843.576	1766.0	7.03	
y53	5878.234	1199.0	4.77	
b55	5972.618	1442.0	5.74	
b57	6230.706	2560.0	10.19	
b58	6287.725	1320.0	5.26	
b67	7377.089	2442.0	9.72	

c)

3A_solubile_010714_XT_00001_MHp__1612... 22/12/2016 11:59:27

3A_solubile_010714_XT_00001_MHp__161222115927#2 RT: 2.00
T: FTMS + p ESI d Full ms2 857.00@cid35.00 [225.00-2000.00]
m/z= 626.219-5754.203

	m/z	Intensity	Relative	Charge
	631.219	4839.0	14.90	
	687.257	1989.0	6.12	
	703.241	2568.0	7.91	
	742.249	3040.0	9.36	
	817.796	1712.0	5.27	
	875.288	3992.0	12.29	
	889.305	9638.0	29.68	
	904.828	2209.0	6.80	
	906.760	2165.0	6.67	
	907.831	1840.0	5.67	
	908.450	1940.0	5.97	
	911.839	2963.0	9.12	
	956.205	2179.0	6.71	
	971.978	3051.0	9.39	
	995.793	2776.0	8.55	
	1000.338	2746.0	8.46	
	1004.330	5251.0	16.17	
	1014.470	2051.0	6.32	
	1018.348	11942.0	36.77	
	1147.391	7934.0	24.43	
	1262.417	2661.0	8.19	
	1276.433	3006.0	9.26	
	1607.774	1128.0	3.47	
	2123.352	2022.0	6.23	
	2343.586	1726.0	5.31	
	2690.537	1131.0	3.48	
b32	3488.804	3465.0	10.67	
y32	3571.491	1010.0	3.11	
b34	3656.762	2284.0	7.03	
	3698.804	878.0	2.70	
Y33	3699.529	4460.0	13.73	
y34	3814.573	2033.0	6.26	
y35	3871.583	11074.0	34.10	
y36	3986.604	21078.0	64.90	
b38	4012.924	5724.0	17.63	
	4097.653	3375.0	10.39	
y37	4115.651	13160.0	40.52	
y38	4244.693	4509.0	13.88	
b40	4256.000	860.0	2.65	
	4430.769	9201.0	28.33	
	4545.773	6785.0	20.89	
	4584.812	4483.0	13.80	
	4602.800	32476.0	100.00	
	4713.840	2846.0	8.76	
	4731.842	9798.0	30.17	
	4842.889	1635.0	5.03	
	4860.885	8099.0	24.94	
	4971.918	1474.0	4.54	
	4989.927	2243.0	6.91	
	5353.393	1666.0	5.13	
	5356.338	713.0	2.20	
b50	5371.394	1680.0	5.17	
	5691.981	1986.0	6.12	
	5749.203	4520.0	13.92	

b		Acetyl		y	b				y	b				y
130.0499	1	**S**	109	11936.8931	5013.2952	47	**A**	63	7037.6529	10084.1121	93	**D**	17	2010.8258
245.0768	2	**D**	108	11849.8611	5128.3221	48	**D**	62	6966.6158	10213.1547	94	**E**	16	1895.7989
316.1139	3	**A**	107	11734.8342	5242.3650	49	**N**	61	6851.5889	10328.1817	95	**D**	15	1766.7563
387.1510	4	**A**	106	11663.7971	5371.4076	50	**E**	60	6737.5459	10443.2086	96	**D**	14	1651.7293
486.2195	5	**V**	105	11592.7599	5470.4760	51	**V**	59	6608.5034	10558.2355	97	**D**	13	1536.7024
601.2464	6	**D**	104	11493.6915	5585.5030	52	**D**	58	6509.4349	10657.3040	98	**V**	12	1421.6754
702.2941	7	**T**	103	11378.6646	5714.5456	53	**E**	57	6394.4080	10772.3309	99	**D**	11	1322.6070
789.3261	8	**S**	102	11277.6169	5843.5882	54	**E**	56	6265.3654	10873.3786	100	**T**	10	1207.5801
876.3581	9	**S**	101	11190.5849	5972.6308	55	**E**	55	6136.3228	11001.4735	101	**K**	9	1106.5324
1005.4007	10	**E**	100	11103.5528	6101.6734	56	**E**	54	6007.2802	11129.5685	102	**K**	8	978.4374
1118.4848	11	**I**	99	10974.5103	6230.7159	57	**E**	53	5878.2376	11257.6271	103	**Q**	7	850.3425
1219.5325	12	**T**	98	10861.4262	6287.7374	58	**G**	52	5749.1950	11385.7220	104	**K**	6	722.2839
1320.5801	13	**T**	97	10760.3785	6344.7589	59	**G**	51	5692.1736	11486.7697	105	**T**	5	594.1889
1448.6751	14	**K**	96	10659.3308	6473.8015	60	**E**	50	5635.1521	11601.7967	106	**D**	4	493.1413
1563.7021	15	**D**	95	10531.2359	6602.8441	61	**E**	49	5506.1095	11730.8393	107	**E**	3	378.1143
1676.7861	16	**L**	94	10416.2089	6731.8867	62	**E**	48	5377.0669	11845.8662	108	**D**	2	249.0717
1804.8811	17	**K**	93	10303.1249	6860.9292	63	**E**	47	5248.0243	---	109	**D**	1	134.0448
1933.9237	18	**E**	92	10175.0299	6989.9718	64	**E**	46	5118.9817					
2062.0186	19	**K**	91	10045.9873	7119.0144	65	**E**	45	4989.9391					
2190.1136	20	**K**	90	9917.8923	7248.0570	66	**E**	44	4860.8965					
2319.1562	21	**E**	89	9789.7974	7377.0996	67	**E**	43	4731.8540					
2418.2246	22	**V**	88	9660.7548	7434.1211	68	**G**	42	4602.8114					
2517.2930	23	**V**	87	9561.6864	7549.1480	69	**D**	41	4545.7899					
2646.3356	24	**E**	86	9462.6180	7606.1695	70	**G**	40	4430.7630					
2775.3782	25	**E**	85	9333.5754	7735.2121	71	**E**	39	4373.7415					
2846.4153	26	**A**	84	9204.5328	7864.2547	72	**E**	38	4244.6989					
2975.4579	27	**E**	83	9133.4957	7993.2973	73	**E**	37	4115.6563					
3089.5008	28	**N**	82	9004.4531	8108.3242	74	**D**	36	3986.6137					
3146.5223	29	**G**	81	8890.4101	8165.3457	75	**G**	35	3871.5868					
3302.6234	30	**R**	80	8833.3887	8280.3726	76	**D**	34	3814.5653					
3417.6504	31	**D**	79	8677.2876	8409.4152	77	**E**	33	3699.5384					
3488.6875	32	**A**	78	8562.2606	8524.4422	78	**D**	32	3570.4958					
3585.7402	33	**P**	77	8491.2235	8653.4848	79	**E**	31	3455.4688					
3656.7773	34	**A**	76	8394.1707	8782.5273	80	**E**	30	3326.4262					
3770.8203	35	**N**	75	8323.1336	8853.5645	81	**A**	29	3197.3836					
3827.8417	36	**G**	74	8209.0907	8982.6071	82	**E**	28	3126.3465					
3941.8847	37	**N**	73	8152.0692	9069.6391	83	**S**	27	2997.3039					
4012.9218	38	**A**	72	8038.0263	9140.6762	84	**A**	26	2910.2719					
4126.9647	39	**N**	71	7966.9892	9241.7239	85	**T**	25	2839.2348					
4256.0073	40	**E**	70	7852.9463	9298.7453	86	**G**	24	2738.1871					
4385.0499	41	**E**	69	7723.9037	9426.8403	87	**K**	23	2681.1656					
4499.0928	42	**N**	68	7594.8611	9582.9414	88	**R**	22	2553.0707					
4556.1143	43	**G**	67	7480.8182	9653.9785	89	**A**	21	2396.9696					
4685.1569	44	**E**	66	7423.7967	9725.0156	90	**A**	20	2325.9325					
4813.2155	45	**Q**	65	7294.7541	9854.0582	91	**E**	19	2254.8953					
4942.2581	46	**E**	64	7166.6955	9969.0852	92	**D**	18	2125.8528					

Fig. S7. Tα11.
$[M+H]^+$ 3788.83315 Da (Theor. 3788.8307 Da). MS/MS analysis performed on different m/z ions in different samples.
- Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+6H]^{6+}$ monoisotopic m/z: 947.96375 Da (+0.53 mmu/+0.56 ppm). RT: 21.63 min. Analysis performed by Proteome discoverer software. XCorr:3.12.
Fragment match tolerance used for search: 0.6 Da; Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (2):
- Prothymosin alpha OS=Homo sapiens GN=PTMA PE=1 SV=2 - [PTMA_HUMAN]
- Isoform 2 of Prothymosin alpha OS=Homo sapiens GN=PTMA - [PTMA_HUMAN]

#1	b^+	b^{2+}	b^{3+}	b^{4+}	Seq.	y^+	y^{2+}	y^{3+}	y^{4+}	#2
1	130.04987	65.52857	44.02147	33.26793	S-Acetyl					
2	245.07682	123.04205	82.36379	62.02466	D	3659.78844	1830.39786	1220.60100	915.70257	34
3	316.11394	158.56061	106.04283	79.78394	A	3544.76149	1772.88438	1182.25868	886.94583	33
4	387.15106	194.07917	129.72187	97.54322	A	3473.72437	1737.36582	1158.57964	869.18655	32
5	486.21948	243.61338	162.74468	122.31033	V	3402.68725	1701.84726	1134.90060	851.42727	31
6	601.24643	301.12685	201.08699	151.06707	D	3303.61883	1652.31305	1101.87779	826.66017	30
7	702.29411	351.65069	234.76955	176.32899	T	3188.59188	1594.79958	1063.53548	797.90343	29
8	789.32614	395.16671	263.78023	198.08699	S	3087.54420	1544.27574	1029.85292	772.64151	28
9	876.35817	438.68272	292.79091	219.84500	S	3000.51217	1500.75972	1000.84224	750.88350	27
10	1005.40077	503.20402	335.80511	252.10565	E	2913.48014	1457.24371	971.83156	729.12549	26
11	1118.48484	559.74606	373.49980	280.37667	I	2784.43754	1392.72241	928.81736	696.86484	25
12	1219.53252	610.26990	407.18236	305.63859	T	2671.35347	1336.18037	891.12267	668.59383	24
13	1320.58020	660.79374	440.86492	330.90051	T	2570.30579	1285.65653	857.44011	643.33191	23
14	1448.67517	724.84122	483.56324	362.92425	K	2469.25811	1235.13269	823.75755	618.06999	22
15	1563.70212	782.35470	521.90556	391.68099	D	2341.16314	1171.08521	781.05923	586.04624	21
16	1676.78619	838.89673	559.60025	419.95201	L	2226.13619	1113.57173	742.71691	557.28951	20
17	1804.88116	902.94422	602.29857	451.97575	K	2113.05212	1057.02970	705.02222	529.01849	19
18	1933.92376	967.46552	645.31277	484.23640	E	1984.95715	992.98221	662.32390	496.99475	18
19	2062.01873	1031.51300	688.01109	516.26014	K	1855.91455	928.46091	619.30970	464.73410	17
20	2190.11370	1095.56049	730.70942	548.28388	K	1727.81958	864.41343	576.61138	432.71035	16
21	2319.15630	1160.08179	773.72362	580.54453	E	1599.72461	800.36594	533.91305	400.68661	15
22	2418.22472	1209.61600	806.74642	605.31164	V	1470.68201	735.84464	490.89885	368.42596	14
23	2517.29314	1259.15021	839.76923	630.07874	V	1371.61359	686.31043	457.87605	343.65886	13
24	2646.33574	1323.67151	882.78343	662.33939	E	1272.54517	636.77622	424.85324	318.89175	12
25	2775.37834	1388.19281	925.79763	694.60004	E	1143.50257	572.25492	381.83904	286.63110	11
26	2846.41546	1423.71137	949.47667	712.35932	A	1014.45997	507.73362	338.82484	254.37045	10
27	2975.45806	1488.23267	992.49087	744.61997	E	943.42285	472.21506	315.14580	236.61117	9
28	3089.50099	1545.25413	1030.50518	773.13071	N	814.38025	407.69376	272.13160	204.35052	8
29	3146.52246	1573.76487	1049.51234	787.38607	G	700.33732	350.67230	234.11729	175.83979	7
30	3302.62358	1651.81543	1101.54604	826.41135	R	643.31585	322.16156	215.11013	161.58442	6
31	3417.65053	1709.32890	1139.88836	855.16809	D	487.21473	244.11100	163.07643	122.55914	5
32	3488.68765	1744.84746	1163.56740	872.92737	A	372.18778	186.59753	124.73411	93.80240	4
33	3585.74042	1793.37385	1195.91832	897.19056	P	301.15066	151.07897	101.05507	76.04312	3
34	3656.77754	1828.89241	1219.59736	914.94984	A	204.09789	102.55258	68.70415	51.77993	2
35					N	133.06077	67.03402	45.02511	34.02065	1

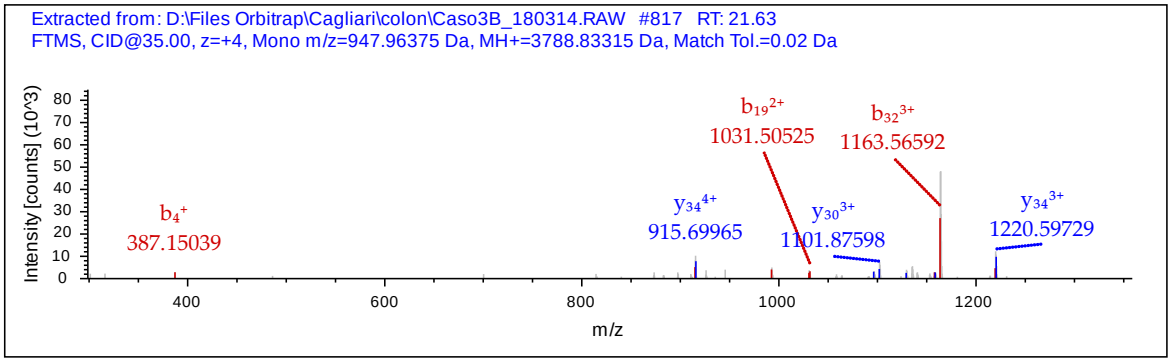


Fig. S8. Parathymosin.

$[M+H]^+$ 11435.1696 Da (Theor. 11435.1649). MS/MS analysis performed on different m/z ions in different samples.

a) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+11H]^{11+}$ monoisotopic m/z: 1041.12 Da. RT: 20.71 min.

b) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+9H]^{9+}$ monoisotopic m/z: 1272.14 Da. RT: 20.88 min.

c) Analysis manually performed and validate by comparison with the high-resolution MS/MS simulation of the software MS-Product available on the ProteinProspector website (http://prospector.ucsf.edu/prospector/mshome.htm). In green are highlighted the b and y Ions matching with the experimental ones.

a)

3A_solubile_010714_XT_00001_MHp__1612... 22/12/2016 12:55:12

3A_solubile_010714_XT_00001_MHp__161222125512#2 RT: 2.00
T: FTMS + p ESI d Full ms2 1041.12@cid35.00 [275.00-2000.00]
m/z= 972.489-6636.153

	m/z	Intensity	Relative	Charge
	977.489	1105.0	10.58	
	980.157	2955.0	28.30	
y12	1287.683	642.0	6.15	
y39	4374.873	2597.0	24.87	
	4559.922	613.0	5.87	
y41	4560.939	4557.0	43.64	
y43	4806.000	1116.0	10.69	
y44	4934.047	1577.0	15.10	
y48	5365.190	2188.0	20.95	
b37	5383.727	772.0	7.39	
	5475.194	690.0	6.61	
y49	5493.219	5848.0	56.00	
	5514.774	450.0	4.31	
y50	5564.286	2949.0	28.24	
b49	5641.803	1965.0	18.82	
b50	5770.847	2715.0	26.00	
	5854.891	1526.0	14.61	
b51	5871.905	1584.0	15.17	
b52	5942.924	10442.0	100.00	
b53	6071.981	4227.0	40.48	
	6374.072	937.0	8.97	
b57	6502.116	1500.0	14.37	
b58	6631.153	1051.0	10.07	

b)

3A_solubile_010714_XT_00001_MHp__1612... 22/12/2016 13:18:53

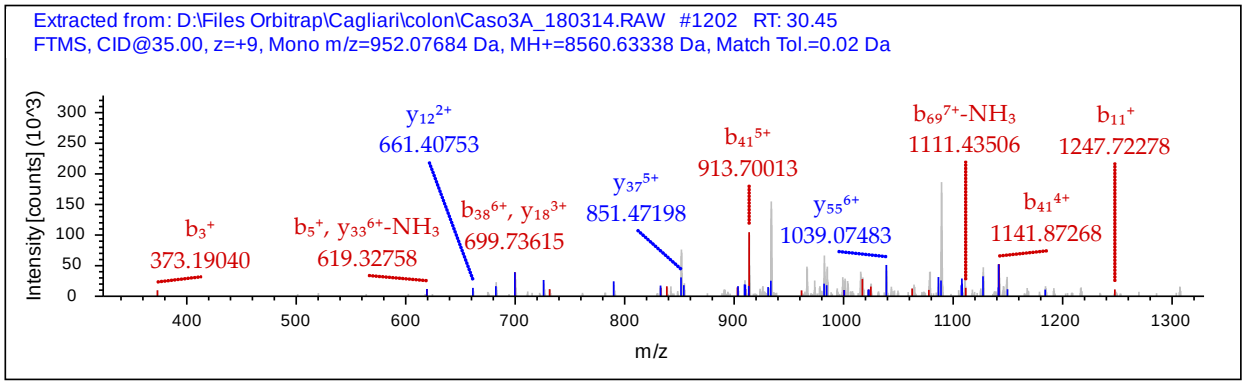
3A_solubile_010714_XT_00001_MHp__161222131853#2 RT: 2.00
F: FTMS + p ESI d Full ms2 1272.14@cid35.00 [340.00-2000.00]
m/z= 1282.687-6192.022

	m/z	Intensity	Relative	Charge
y12	1287.687	1189.0	9.97	
	1327.239	1158.0	9.71	
	1388.069	1549.0	12.99	
y43	4805.992	902.0	7.57	
y44	4934.045	1106.0	9.28	
y47	5249.146	5330.0	44.71	
y48	5364.162	5549.0	46.55	
b37	5383.722	1308.0	10.97	
	5475.204	3068.0	25.74	
y49	5493.229	6022.0	50.52	
	5512.772	2298.0	19.28	
y50	5564.243	4359.0	36.57	
b49	5641.810	2695.0	22.61	
y51	5665.303	2954.0	24.78	
b50	5770.862	4386.0	36.79	
	5853.891	3473.0	29.13	
b51	5871.882	2272.0	19.06	
	5925.914	909.0	7.63	
b52	5942.937	11921.0	100.00	
b53	6071.983	7301.0	61.24	
b54	6187.022	5330.0	44.71	

c)

b		Acetyl		y	b			y	b			y
130.0499	1	**S**	101	11393.1546	5383.7351	47	**E**	55	6181.4799	10658.7987	93	**Q**	9	905.4323
259.0925	2	**E**	100	11306.1226	5512.7777	48	**E**	54	6052.4373	10786.8937	94	**K**	8	777.3737
387.1874	3	**K**	99	11177.0800	5641.8203	49	**E**	53	5923.3947	10887.9413	95	**T**	7	649.2788
474.2195	4	**S**	98	11048.9850	5770.8629	50	**E**	52	5794.3521	11016.9839	96	**E**	6	548.2311
573.2879	5	**V**	97	10961.9530	5871.9106	51	**T**	51	5665.3095	11131.0269	97	**N**	5	419.1885
702.3305	6	**E**	96	10862.8846	5942.9477	52	**A**	50	5564.2619	11188.0483	98	**G**	4	305.1456
773.3676	7	**A**	95	10733.8420	6071.9903	53	**E**	49	5493.2248	11259.0854	99	**A**	3	248.1241
844.4047	8	**A**	94	10662.8049	6187.0172	54	**D**	48	5364.1822	11346.1175	100	**S**	2	177.0870
915.4418	9	**A**	93	10591.7677	6244.0387	55	**G**	47	5249.1552	---	101	**A**	1	90.0550
1044.4844	10	**E**	92	10520.7306	6373.0813	56	**E**	46	5192.1338					
1157.5685	11	**L**	91	10391.6880	6502.1239	57	**E**	45	5063.0912					
1244.6005	12	**S**	90	10278.6040	6631.1664	58	**E**	44	4934.0486					
1315.6376	13	**A**	89	10191.5719	6746.1934	59	**D**	43	4805.0060					
1443.7326	14	**K**	88	10120.5348	6875.2360	60	**E**	42	4689.9790					
1558.7595	15	**D**	87	9992.4399	6932.2574	61	**G**	41	4560.9364					
1671.8436	16	**L**	86	9877.4129	7061.3000	62	**E**	40	4503.9150					
1799.9385	17	**K**	85	9764.3289	7190.3426	63	**E**	39	4374.8724					
1928.9811	18	**E**	84	9636.2339	7319.3852	64	**E**	38	4245.8298					
2057.0761	19	**K**	83	9507.1913	7434.4122	65	**D**	37	4116.7872					
2185.1711	20	**K**	82	9379.0963	7563.4548	66	**E**	36	4001.7603					
2314.2136	21	**E**	81	9251.0014	7692.4974	67	**E**	35	3872.7177					
2442.3086	22	**K**	80	9121.9588	7821.5399	68	**E**	34	3743.6751					
2541.3770	23	**V**	79	8993.8638	7950.5825	69	**E**	33	3614.6325					
2670.4196	24	**E**	78	8894.7954	8079.6251	70	**E**	32	3485.5899					
2799.4622	25	**E**	77	8765.7528	8208.6677	71	**E**	31	3356.5473					
2927.5572	26	**K**	76	8636.7102	8323.6947	72	**D**	30	3227.5047					
2998.5943	27	**A**	75	8508.6153	8438.7216	73	**D**	29	3112.4778					
3085.6263	28	**S**	74	8437.5781	8567.7642	74	**E**	28	2997.4508					
3241.7274	29	**R**	73	8350.5461	8624.7857	75	**G**	27	2868.4082					
3369.8224	30	**K**	72	8194.4450	8721.8384	76	**P**	26	2811.3868					
3498.8650	31	**E**	71	8066.3500	8792.8755	77	**A**	25	2714.3340					
3654.9661	32	**R**	70	7937.3074	8905.9596	78	**L**	24	2643.2969					
3783.0611	33	**K**	69	7781.2063	9034.0546	79	**K**	23	2530.2128					
3911.1560	34	**K**	68	7653.1114	9190.1557	80	**R**	22	2402.1178					
4040.1986	35	**E**	67	7525.0164	9261.1928	81	**A**	21	2246.0167					
4139.2670	36	**V**	66	7395.9738	9332.2299	82	**A**	20	2174.9796					
4238.3354	37	**V**	65	7296.9054	9461.2725	83	**E**	19	2103.9425					
4367.3780	38	**E**	64	7197.8370	9590.3151	84	**E**	18	1974.8999					
4496.4206	39	**E**	63	7068.7944	9719.3577	85	**E**	17	1845.8573					
4625.4632	40	**E**	62	6939.7518	9834.3846	86	**D**	16	1716.8147					
4754.5058	41	**E**	61	6810.7092	9963.4272	87	**E**	15	1601.7878					
4868.5487	42	**N**	60	6681.6666	10034.4643	88	**A**	14	1472.7452					
4925.5702	43	**G**	59	6567.6237	10149.4913	89	**D**	13	1401.7081					
4996.6073	44	**A**	58	6510.6022	10246.5441	90	**P**	12	1286.6811					
5125.6499	45	**E**	57	6439.5651	10374.6390	91	**K**	11	1189.6284					
5254.6925	46	**E**	56	6310.5225	10530.7401	92	**R**	10	1061.5334					

Fig. S9. Ubiquitin.
$[M+H]^+$ 8560.633 Da (Theor. 3788.831 Da). MS/MS analysis performed on different m/z ions in different samples.
- Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+6H]^{6+}$ monoisotopic m/z: 952.08 Da (+1 mmu/+1.05 ppm). RT: 30.45 min. Analysis performed by Proteome discoverer software. XCorr:3.66.
Fragment match tolerance used for search: 0.6 Da; Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (4):
- Polyubiquitin-C OS=Homo sapiens GN=UBC PE=1 SV=3 - [UBC_HUMAN]
- Polyubiquitin-B OS=Homo sapiens GN=UBB PE=1 SV=1 - [UBB_HUMAN]
- Ubiquitin-60S ribosomal protein L40 OS=Homo sapiens GN=UBA52 PE=1 SV=2 - [RL40_HUMAN]
- Ubiquitin-40S ribosomal protein S27a OS=Homo sapiens GN=RPS27A PE=1 SV=2 - [RS27A_HUMAN]



#1	b	b²	b³	b	b	b	Seq.	y	y²	y³	y	y	y	y	#2
1	132.048	66.528	44.687	33.767	27.215	22.847	M								76
2	260.106	130.557	87.374	65.782	52.827	44.190	Q	8429.584	4215.296	2810.533	2108.151	1686.723	1405.770	1205.090	75
3	373.190	187.099	125.068	94.053	75.444	63.038	I	8301.525	4151.266	2767.847	2076.137	1661.111	1384.427	1186.796	74
4	520.259	260.633	174.091	130.820	104.858	87.549	F	8188.441	4094.724	2730.152	2047.866	1638.494	1365.580	1170.641	73
5	619.327	310.167	207.114	155.587	124.671	104.061	V	8041.373	4021.190	2681.129	2011.099	1609.080	1341.068	1149.631	72
6	747.422	374.215	249.812	187.611	150.290	125.410	K	7942.304	3971.656	2648.106	1986.332	1589.267	1324.557	1135.478	71
7	848.470	424.739	283.495	212.873	170.500	142.251	T	7814.209	3907.608	2605.408	1954.308	1563.648	1303.208	1117.179	70
8	961.554	481.281	321.190	241.144	193.117	161.098	L	7713.162	3857.085	2571.725	1929.046	1543.438	1286.366	1102.744	69
9	1062.602	531.804	354.872	266.406	213.326	177.940	T	7600.078	3800.542	2534.031	1900.775	1520.821	1267.519	1086.589	68
10	1119.623	560.315	373.879	280.661	224.730	187.443	G	7499.030	3750.019	2500.348	1875.513	1500.612	1250.678	1072.153	67
11	1247.718	624.363	416.578	312.685	250.349	208.792	K	7442.009	3721.508	2481.341	1861.258	1489.208	1241.174	1064.007	66
12	1348.766	674.887	450.260	337.947	270.559	225.634	T	7313.914	3657.460	2438.643	1829.234	1463.589	1219.825	1045.708	65
13	1461.850	731.429	487.955	366.218	293.176	244.481	I	7212.866	3606.937	2404.960	1803.972	1443.379	1202.984	1031.273	64
14	1562.898	781.952	521.637	391.480	313.385	261.322	T	7099.782	3550.395	2367.265	1775.701	1420.762	1184.136	1015.118	63
15	1675.982	838.494	559.332	419.751	336.002	280.170	L	6998.734	3499.871	2333.583	1750.439	1400.553	1167.295	1000.683	62
16	1805.024	903.016	602.346	452.012	361.811	301.677	E	6885.650	3443.329	2295.888	1722.168	1377.936	1148.448	984.528	61
17	1904.093	952.550	635.369	476.779	381.624	318.188	V	6756.607	3378.807	2252.874	1689.907	1352.127	1126.941	966.093	60
18	2033.135	1017.071	678.383	509.039	407.433	339.695	E	6657.539	3329.273	2219.851	1665.140	1332.314	1110.429	951.940	59
19	2130.188	1065.598	710.734	533.302	426.843	355.871	P	6528.496	3264.752	2176.837	1632.880	1306.505	1088.922	933.506	58
20	2217.220	1109.114	739.745	555.060	444.250	370.376	S	6431.444	3216.225	2144.486	1608.616	1287.095	1072.747	919.641	57
21	2332.247	1166.627	778.087	583.817	467.255	389.547	D	6344.412	3172.709	2115.475	1586.858	1269.688	1058.241	907.208	56
22	2433.295	1217.151	811.770	609.079	487.465	406.389	T	6229.385	3115.196	2077.133	1558.102	1246.683	1039.070	890.775	55
23	2546.379	1273.693	849.464	637.350	510.082	425.236	I	6128.337	3064.672	2043.451	1532.840	1226.473	1022.229	876.340	54
24	2675.421	1338.214	892.479	669.611	535.890	446.743	E	6015.253	3008.130	2005.756	1504.569	1203.856	1003.382	860.185	53
25	2789.464	1395.236	930.493	698.122	558.699	465.750	N	5886.210	2943.609	1962.742	1472.308	1178.048	981.874	841.751	52
26	2888.533	1444.770	963.516	722.889	578.512	482.262	V	5772.167	2886.587	1924.727	1443.797	1155.239	962.867	825.459	51
27	3016.628	1508.817	1006.214	754.912	604.131	503.611	K	5673.099	2837.053	1891.705	1419.030	1135.426	946.356	811.306	50
28	3087.665	1544.336	1029.893	772.672	618.339	515.450	A	5545.004	2773.006	1849.006	1387.006	1109.807	925.007	793.007	49
29	3215.760	1608.384	1072.591	804.695	643.958	536.799	K	5473.967	2737.487	1825.327	1369.247	1095.599	913.167	782.859	48
30	3328.844	1664.926	1110.286	832.966	666.575	555.647	I	5345.872	2673.440	1782.629	1337.223	1069.980	891.818	764.559	47
31	3456.902	1728.955	1152.972	864.981	692.186	576.990	Q	5232.788	2616.898	1744.934	1308.952	1047.363	872.971	748.405	46
32	3571.929	1786.468	1191.315	893.738	715.192	596.161	D	5104.729	2552.868	1702.248	1276.938	1021.752	851.628	730.110	45
33	3700.024	1850.516	1234.013	925.762	740.811	617.510	K	4989.702	2495.355	1663.906	1248.181	998.746	832.456	713.678	44
34	3829.067	1915.037	1277.027	958.022	766.619	639.017	E	4861.607	2431.307	1621.207	1216.157	973.127	811.107	695.379	43
35	3886.088	1943.548	1296.034	972.278	778.024	648.521	G	4732.565	2366.786	1578.193	1183.897	947.319	789.600	676.944	42
36	3999.172	2000.090	1333.729	1000.549	800.640	667.368	I	4675.543	2338.275	1559.186	1169.641	935.914	780.097	668.798	41
37	4096.225	2048.616	1366.080	1024.812	820.051	683.544	P	4562.459	2281.733	1521.491	1141.370	913.298	761.249	652.643	40
38	4193.278	2097.143	1398.431	1049.075	839.461	699.719	P	4465.406	2233.207	1489.140	1117.107	893.887	745.074	638.779	39
39	4308.305	2154.656	1436.773	1077.832	862.467	718.890	D	4368.354	2184.680	1456.789	1092.844	874.477	728.898	624.914	38
40	4436.364	2218.685	1479.459	1109.846	888.079	740.233	Q	4253.327	2127.167	1418.447	1064.087	851.471	709.727	608.481	37
41	4564.422	2282.715	1522.146	1141.861	913.690	761.576	Q	4125.268	2063.138	1375.761	1032.073	825.859	688.384	590.187	36
42	4720.523	2360.765	1574.179	1180.886	944.910	787.593	R	3997.210	1999.108	1333.075	1000.058	800.248	667.041	571.893	35
43	4833.607	2417.307	1611.874	1209.157	967.527	806.441	L	3841.108	1921.058	1281.041	961.033	769.028	641.024	549.593	34
44	4946.691	2473.849	1649.569	1237.428	990.144	825.288	I	3728.024	1864.516	1243.346	932.762	746.411	622.177	533.438	33
45	5093.760	2547.384	1698.591	1274.195	1019.558	849.799	F	3614.940	1807.974	1205.652	904.491	723.794	603.329	517.283	32
46	5164.797	2582.902	1722.270	1291.955	1033.765	861.639	A	3467.872	1734.440	1156.629	867.723	694.380	578.818	496.274	31
47	5221.818	2611.413	1741.278	1306.210	1045.169	871.142	G	3396.835	1698.921	1132.950	849.964	680.173	566.979	486.125	30
48	5349.913	2675.460	1783.976	1338.234	1070.788	892.492	K	3339.813	1670.410	1113.943	835.709	668.768	557.475	477.980	29
49	5477.972	2739.490	1826.662	1370.248	1096.400	913.835	Q	3211.718	1606.363	1071.244	803.685	643.149	536.126	459.680	28
50	5591.056	2796.032	1864.357	1398.519	1119.017	932.682	L	3083.660	1542.334	1028.558	771.670	617.538	514.783	441.386	27
51	5720.099	2860.553	1907.371	1430.780	1144.826	954.189	E	2970.576	1485.791	990.863	743.399	594.921	495.935	425.231	26
52	5835.126	2918.066	1945.713	1459.537	1167.831	973.360	D	2841.533	1421.270	947.849	711.139	569.112	474.428	406.797	25
53	5892.147	2946.577	1964.721	1473.792	1179.235	982.864	G	2726.506	1363.757	909.507	682.382	546.107	455.257	390.364	24
54	6048.248	3024.628	2016.754	1512.817	1210.455	1008.881	R	2669.485	1335.246	890.500	668.127	534.703	445.754	382.218	23
55	6149.296	3075.152	2050.437	1538.079	1230.665	1025.722	T	2513.384	1257.195	838.466	629.101	503.483	419.737	359.918	22
56	6262.380	3131.694	2088.131	1566.350	1253.282	1044.569	L	2412.336	1206.672	804.783	603.839	483.273	402.895	345.483	21
57	6349.412	3175.210	2117.142	1588.108	1270.688	1059.075	S	2299.252	1150.130	767.089	575.568	460.656	384.048	329.328	20
58	6464.439	3232.723	2155.484	1616.865	1293.694	1078.246	D	2212.220	1106.614	738.078	553.810	443.250	369.543	316.895	19
59	6627.502	3314.255	2209.839	1657.631	1326.306	1105.423	Y	2097.193	1049.100	699.736	525.054	420.244	350.372	300.462	18
60	6741.545	3371.276	2247.853	1686.142	1349.115	1124.430	N	1934.129	967.568	645.381	484.288	387.632	323.194	277.168	17
61	6854.629	3427.818	2285.548	1714.413	1371.732	1143.278	I	1820.087	910.547	607.367	455.777	364.823	304.187	260.876	16
62	6982.688	3491.848	2328.234	1746.427	1397.343	1164.621	Q	1707.002	854.005	569.672	427.506	342.206	285.340	244.721	15
63	7110.783	3555.895	2370.932	1778.451	1422.962	1185.970	K	1578.944	789.976	526.986	395.491	316.595	263.997	226.427	14
64	7239.825	3620.416	2413.947	1810.712	1448.771	1207.477	E	1450.849	725.928	484.288	363.468	290.976	242.648	208.128	13
65	7326.857	3663.932	2442.957	1832.470	1466.177	1221.982	S	1321.806	661.407	441.274	331.207	265.167	221.140	189.693	12
66	7427.905	3714.456	2476.640	1857.732	1486.387	1238.824	T	1234.774	617.891	412.263	309.449	247.761	206.635	177.260	11
67	7540.989	3770.998	2514.335	1886.003	1509.004	1257.671	L	1133.727	567.367	378.580	284.187	227.551	189.794	162.824	10
68	7678.048	3839.528	2560.021	1920.267	1536.415	1280.514	H	1020.643	510.825	340.886	255.916	204.934	170.946	146.669	9
69	7791.132	3896.070	2597.716	1948.538	1559.032	1299.361	L	883.584	442.295	295.199	221.651	177.523	148.103	127.090	8
70	7890.201	3945.604	2630.738	1973.306	1578.846	1315.873	V	770.500	385.753	257.505	193.380	154.906	129.256	110.935	7
71	8003.285	4002.146	2668.433	2001.577	1601.463	1334.720	L	671.431	336.219	224.482	168.613	135.092	112.745	96.782	6
72	8159.386	4080.196	2720.467	2040.602	1632.683	1360.737	R	558.347	279.677	186.787	140.342	112.475	93.897	80.627	5
73	8272.470	4136.739	2758.161	2068.873	1655.300	1379.584	L	402.246	201.627	134.754	101.317	81.255	67.880	58.327	4
74	8428.571	4214.789	2810.195	2107.898	1686.520	1405.601	R	289.162	145.085	97.059	73.046	58.638	49.033	42.172	3
75	8485.592	4243.300	2829.202	2122.154	1697.924	1415.105	G	133.061	67.034	45.025	34.021	27.418	23.016	19.872	2
76							G	76.039	38.523	26.018	19.765	16.014	13.513	11.726	1

Fig. S10. SH3BP-1 protein

$[M+H]^+$ 10343.20 Da (Theor. 10343.24). MS/MS analysis performed on different m/z ions in different samples.

a) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+8H]^{8+}$ monoisotopic m/z: 1294.41 Da. RT: 40.10 min.

b) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+10H]^{10+}$ monoisotopic m/z: 1035.64 Da. RT: 40.12 min.

c) Analysis manually performed and validate by comparison with the high-resolution MS/MS simulation of the software MS-Product available on the ProteinProspector website (http://prospector.ucsf.edu/prospector/mshome.htm). In green are highlighted the b and y Ions matching with the experimental ones.

a)

FrescoA_caso1_a(040413)_XT_00001_MHp_ 22/12/2016 14:20:13

FrescoA_caso1_a(040413)_XT_00001_MHp_#2 RT: 2.00
F: FTMS + p ESI d Full ms2 1294.41@cid35.00 [345.00-2000.00]
m/z= 715.428-10146.092

	m/z	Intensity	Relative	Charge
y6	720.428	3888.0	5.35	
y10	1176.659	2222.0	3.05	
	1217.475	3507.0	4.82	
	1231.335	2242.0	3.08	
	1233.474	1618.0	2.22	
	1234.366	2051.0	2.82	
	1234.629	2310.0	3.18	
	1249.918	4056.0	5.58	
	1250.348	2360.0	3.24	
	1251.135	3672.0	5.05	
	1354.672	5508.0	7.57	
	1373.977	4180.0	5.75	
	1391.241	3465.0	4.76	
	1393.132	1871.0	2.57	
	1429.291	5201.0	7.15	
	1458.244	1746.0	2.40	
	3121.628	2012.0	2.77	
	4276.602	957.0	1.32	
b55	6177.189	958.0	1.32	
	8742.382	3392.0	4.66	
b79	8811.414	3927.0	5.40	
b80	8910.480	7017.0	9.65	
b82	9168.561	8421.0	11.58	
b83	9281.616	4481.0	6.16	
b84	9383.674	2185.0	3.00	
	9495.706	10823.0	14.88	
b85	9623.818	14378.0	19.77	
b86				
b87	9752.801	23518.0	32.33	
	9882.942	1442.0	1.98	
b88	9899.924	18669.0	25.67	
	9900.842	3508.0	4.82	
	9995.972	8672.0	11.92	
b89	10013.001	72740.0	100.00	
b90	10141.092	8171.0	11.23	

b)

FrescoA_caso1_a(040413)_XT_00001_MHp_... 22/12/2016 14:23:53

FrescoA_caso1_a(040413)_XT_00001_MHp__161222142353#2 RT: 2.00
F: FTMS + p ESI d Full ms2 1035.64@cid35.00 [275.00-2000.00]
m/z= 935.211-10145.111

	m/z	Intensity	Relative	Charge
	940.211	3454.0	18.40	
	940.353	3375.0	17.98	
	990.199	1442.0	7.68	
	1016.541	4021.0	21.42	
	1041.217	1924.0	10.25	
	1053.441	2453.0	13.07	
	1056.872	2100.0	11.19	
	1076.564	1713.0	9.12	
	1096.442	1325.0	7.06	
	1097.768	1406.0	7.49	
	1098.333	2418.0	12.88	
	1099.777	3092.0	16.47	
	1112.008	3745.0	19.95	
	1112.340	3607.0	19.21	
	1114.343	5883.0	31.34	
	1187.636	1883.0	10.03	
	1204.483	1819.0	9.69	
	1217.196	1494.0	7.96	
	1217.326	2464.0	13.13	
	1217.761	2185.0	11.64	
	1220.120	1371.0	7.30	
	1278.643	1465.0	7.80	
	1321.140	1164.0	6.20	
	1618.753	1087.0	5.79	
	2468.251	715.0	3.81	
	2571.252	1927.0	10.26	
y23	2659.294	4388.0	23.37	
y24	2822.351	1892.0	10.08	
	3161.600	1510.0	8.04	
	6573.419	3027.0	16.12	
	6895.577	2416.0	12.87	
	7107.761	9576.0	51.01	
b65	7221.772	852.0	4.54	
b66	7278.795	1421.0	7.57	
b67	7392.815	873.0	4.65	
	7503.868	935.0	4.98	
b68	7521.882	2987.0	15.91	
b69	7684.943	9574.0	51.00	
b70	7787.949	2776.0	14.79	
B79	8811.399	467.0	2.49	
b84	9383.676	936.0	4.99	
	9736.849	1772.0	9.44	
b87	9752.854	4115.0	21.92	
	9881.956	434.0	2.31	
	9882.905	1275.0	6.79	
b88	9899.929	5483.0	29.21	
	9995.984	1186.0	6.32	
b89	10013.374	18773.0	100.00	
	10123.131	2777.0	14.79	
b90	10140.111	1642.0	8.75	

c)

b		Acetyl	y		b			y	
130.0499	1	**S**	92	10301.2310	5333.7837	47	**D**	46	5125.4922
187.0713	2	**G**	91	10214.1990	5462.8263	48	**E**	45	5010.4652
300.1554	3	**L**	90	10157.1775	5593.8667	49	**M**	44	4881.4226
456.2565	4	**R**	89	10044.0935	5749.9678	50	**R**	43	4750.3821
555.3249	5	**V**	88	9887.9924	5821.0050	51	**A**	42	4594.2810
718.3883	6	**Y**	87	9788.9239	5934.0890	52	**L**	41	4523.2439
805.4203	7	**S**	86	9625.8606	6005.1261	53	**A**	40	4410.1598
906.4680	8	**T**	85	9538.8286	6062.1476	54	**G**	39	4339.1227
993.5000	9	**S**	84	9437.7809	6176.1905	55	**N**	38	4282.1013
1092.5684	10	**V**	83	9350.7489	6273.2433	56	**P**	37	4168.0583
1193.6161	11	**T**	82	9251.6805	6401.3383	57	**K**	36	4071.0056
1250.6375	12	**G**	81	9150.6328	6472.3754	58	**A**	35	3942.9106
1337.6696	13	**S**	80	9093.6113	6573.4230	59	**T**	34	3871.8735
1493.7707	14	**R**	79	9006.5793	6670.4758	60	**P**	33	3770.8258
1622.8133	15	**E**	78	8850.4782	6767.5286	61	**P**	32	3673.7731
1735.8973	16	**I**	77	8721.4356	6895.5872	62	**Q**	31	3576.7203
1863.9923	17	**K**	76	8608.3515	7008.6712	63	**I**	30	3448.6617
1951.0243	18	**S**	75	8480.2566	7107.7396	64	**V**	29	3335.5777
2079.0829	19	**Q**	74	8393.2245	7221.7826	65	**N**	28	3236.5092
2207.1415	20	**Q**	73	8265.1660	7278.8040	66	**G**	27	3122.4663
2294.1735	21	**S**	72	8137.1074	7393.8310	67	**D**	26	3065.4448
2423.2161	22	**E**	71	8050.0754	7521.8895	68	**Q**	25	2950.4179
2522.2845	23	**V**	70	7921.0328	7684.9529	69	**Y**	24	2822.3593
2623.3322	24	**T**	69	7821.9643	7787.9621	70	**C**	23	2659.2960
2779.4333	25	**R**	68	7720.9167	7844.9835	71	**G**	22	2556.2868
2892.5174	26	**I**	67	7564.8156	7960.0105	72	**D**	21	2499.2653
3005.6014	27	**L**	66	7451.7315	8123.0738	73	**Y**	20	2384.2384
3120.6284	28	**D**	65	7338.6474	8252.1164	74	**E**	19	2221.1751
3177.6498	29	**G**	64	7223.6205	8365.2005	75	**L**	18	2092.1325
3305.7448	30	**K**	63	7166.5990	8512.2689	76	**F**	17	1979.0484
3461.8459	31	**R**	62	7038.5041	8611.3373	77	**V**	16	1831.9800
3574.9300	32	**I**	61	6882.4030	8740.3799	78	**E**	15	1732.9116
3702.9886	33	**Q**	60	6769.3189	8811.4170	79	**A**	14	1603.8690
3866.0519	34	**Y**	59	6641.2603	8910.4854	80	**V**	13	1532.8319
3994.1105	35	**Q**	58	6478.1970	9039.5280	81	**E**	12	1433.7635
4107.1945	36	**L**	57	6350.1384	9167.5866	82	**Q**	11	1304.7209
4206.2629	37	**V**	56	6237.0543	9281.6295	83	**N**	10	1176.6623
4321.2899	38	**D**	55	6137.9859	9382.6772	84	**T**	9	1062.6194
4434.3740	39	**I**	54	6022.9590	9495.7612	85	**L**	8	961.5717
4521.4060	40	**S**	53	5909.8749	9623.8198	86	**Q**	7	848.4876
4649.4646	41	**Q**	52	5822.8429	9752.8624	87	**E**	6	720.4291
4764.4915	42	**D**	51	5694.7843	9899.9308	88	**F**	5	591.3865
4878.5344	43	**N**	50	5579.7574	10013.0149	89	**L**	4	444.3180
4949.5715	44	**A**	49	5465.7144	10141.1099	90	**K**	3	331.2340
5062.6556	45	**L**	48	5394.6773	10254.1939	91	**L**	2	203.1390
5218.7567	46	**R**	47	5281.5933	---	92	**A**	1	90.0550

Fig. S11. FABP1

$[M+H]^+$ 14111.48 Da (Theor. 14111.39). MS/MS analysis performed on different m/z ions in different samples.

a) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+10H]^{10+}$ monoisotopic m/z: 1412.95 Da. RT: 39.92 min.

b) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+13H]^{13+}$ monoisotopic m/z: 1086.97 Da. RT: 40.03 min.

c) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+15H]^{15+}$ monoisotopic m/z: 942.10 Da. RT: 39.90 min.

d) Deconvoluted annotated MH^+ spectrum list of high-resolution MS/MS of the ion $[M+16H]^{16+}$ monoisotopic m/z: 883.47 Da. RT: 40.08 min.

e) Analysis manually performed and validate by comparison with the high-resolution MS/MS simulation of the software MS-Product available on the ProteinProspector website (http://prospector.ucsf.edu/prospector/mshome.htm). In green are highlighted the b and y Ions matching with the experimental ones.

a)

3C_superTrip_010716_XT_00001_MHp__161... 22/12/2016 14:42:44

3C_superTrip_010716_XT_00001_MHp__161222144244#2 RT: 2.00
F: FTMS + p ESI d Full ms2 1412.95@cid35.00 [375.00-2000.00]
m/z= 707.330-13568.235

	m/z	Intensity	Relative	Charge		m/z	Intensity	Relative	Charge
b6	712.330	10337.0	16.63			1603.509	12579.0	20.24	
b9	1081.529	10265.0	16.52			1639.450	6718.0	10.81	
	1124.651	8022.0	12.91			1732.930	6789.0	10.92	
	1259.832	26397.0	42.48			1743.382	6808.0	10.96	
	1262.631	8640.0	13.90			1746.014	8165.0	13.14	
	1306.042	7489.0	12.05			1992.146	5832.0	9.38	
	1308.716	10361.0	16.67			2094.195	15536.0	25.00	
	1342.694	10835.0	17.44		y19	2206.276	48503.0	78.05	
	1343.031	10640.0	17.12		y20	2319.363	29013.0	46.69	
	1356.438	8561.0	13.78			2321.364	6087.0	9.80	
	1369.694	7696.0	12.38		y21	2434.398	11766.0	18.93	
	1384.838	7542.0	12.14		y22	2491.410	32818.0	52.81	
	1385.341	11249.0	18.10			2493.414	11351.0	18.27	
	1385.745	8556.0	13.77		y23	2606.455	49697.0	79.97	
	1430.153	6691.0	10.77		y24	2718.544	48812.0	78.55	
	1438.767	15716.0	25.29		y26	2948.630	55915.0	89.98	
	1439.199	9404.0	15.13		y28	3134.731	62143.0	100.00	
	1448.885	9840.0	15.83		y29	3263.823	9379.0	15.09	
	1449.435	13377.0	21.53			3670.863	4120.0	6.63	
	1455.136	24462.0	39.36		y37	4179.573	8930.0	14.37	
	1456.895	7504.0	12.08		y40	4536.952	5365.0	8.63	
	1460.775	8567.0	13.79			5037.616	6234.0	10.03	
	1460.882	9269.0	14.92		y45	5062.834	3679.0	5.92	
	1462.695	9354.0	15.05			6072.479	6264.0	10.08	
	1467.169	10607.0	17.07		b56	6255.261	6483.0	10.43	
	1468.675	7245.0	11.66		b58	6468.401	10753.0	17.30	
	1469.322	6547.0	10.54			6927.795	15062.0	24.24	
	1475.493	10657.0	17.15		b98	10978.673	15497.0	24.94	
	1475.629	8997.0	14.48		b107	11907.175	6116.0	9.84	
	1476.605	11504.0	18.51		y121	13563.235	4470.0	7.19	
	1477.826	9965.0	16.04						
	1478.328	11857.0	19.08						
	1479.057	15571.0	25.06						
	1487.651	12232.0	19.68						
	1487.778	6423.0	10.34						
	1489.399	8709.0	14.01						
	1489.778	7180.0	11.55						
	1490.294	7797.0	12.55						
	1494.939	9069.0	14.59						
	1501.207	13074.0	21.04						
	1503.805	9189.0	14.79						
	1504.021	17255.0	27.77						
	1506.187	6382.0	10.27						
	1510.870	17517.0	28.19						
	1513.344	8527.0	13.72						
	1520.099	7827.0	12.60						
	1523.505	8841.0	14.23						
	1525.916	9548.0	15.36						
	1534.620	12503.0	20.12						
	1536.357	8340.0	13.42						
b13	1539.821	13222.0	21.28						
	1541.427	7290.0	11.73						
	1545.155	7694.0	12.38						
	1552.130	8481.0	13.65						
	1560.343	7489.0	12.05						
	1567.068	6213.0	10.00						
	1596.357	11273.0	18.14						

b) 3C_superTrip_010716_XT_00001_MHp__161... 22/12/2016 14:47:36

3C_superTrip_010716_XT_00001_MHp__161222144736#2 RT: 2.00
F: FTMS + p ESI d Full ms2 1086.97@cid35.00 [285.00-2000.00]
m/z= 708.335-13843.380

	m/z	Intensity	Relative	Charge		m/z	Intensity	Relative	Charge
	713.335	5132.0	5.14			1203.437	13082.0	13.09	
	824.431	10504.0	10.51			1203.732	9447.0	9.45	
b7	840.390	24394.0	24.41			1204.632	11620.0	11.63	
	935.969	6271.0	6.28			1206.945	9087.0	9.09	
b8	953.470	6007.0	6.01			1207.915	8047.0	8.05	
	1032.766	5635.0	5.64			1210.252	6527.0	6.53	
	1063.457	9587.0	9.59			1212.622	7698.0	7.70	
	1063.620	4980.0	4.98			1214.156	9486.0	9.49	
	1118.766	7981.0	7.99			1216.973	8061.0	8.07	
	1120.490	6892.0	6.90			1218.189	8812.0	8.82	
	1121.932	5506.0	5.51			1225.422	6056.0	6.06	
	1122.598	6908.0	6.91			1227.310	6651.0	6.66	
	1125.251	7011.0	7.02			1232.505	5960.0	5.96	
	1133.266	4966.0	4.97			1238.211	6218.0	6.22	
	1137.265	5232.0	5.24			1238.971	7931.0	7.94	
	1137.394	4911.0	4.91			1240.641	6179.0	6.18	
	1139.194	3062.0	3.06			1246.211	8891.0	8.90	
	1141.866	2699.0	2.70			1246.713	9672.0	9.68	
	1143.532	8977.0	8.98			1250.953	4909.0	4.91	
	1143.608	5951.0	5.95			1252.655	8382.0	8.39	
	1145.123	11394.0	11.40		y10	1259.831	6676.0	6.68	
	1145.855	10525.0	10.53			1259.906	4928.0	4.93	
	1146.027	14574.0	14.58			1261.663	5924.0	5.93	
	1148.866	6479.0	6.48			1292.570	5553.0	5.56	
	1149.742	14248.0	14.26			1298.093	5793.0	5.80	
	1151.116	8663.0	8.67			1299.570	5528.0	5.53	
	1153.120	13012.0	13.02		y12	1431.881	24354.0	24.37	
	1155.449	7519.0	7.52		y13	1544.964	10593.0	10.60	
	1158.604	9409.0	9.42		y14	1646.012	40873.0	40.90	
	1160.177	11270.0	11.28		y15	1777.054	42964.0	42.99	
	1163.611	14263.0	14.27			1879.104	12704.0	12.71	
	1167.169	13012.0	13.02			1975.125	7166.0	7.17	
b10	1168.562	2612.0	2.61			1992.145	54062.0	54.10	
	1170.618	7890.0	7.90			2075.184	23605.0	23.62	
	1172.128	4633.0	4.64			2077.179	15500.0	15.51	
	1176.625	15975.0	15.99		y18	2093.191	99935.0	100.00	
	1177.414	12929.0	12.94		y19	2206.276	42758.0	42.79	
	1180.127	10472.0	10.48			2208.286	9306.0	9.31	
	1180.403	10510.0	10.52			2275.187	4506.0	4.51	
	1183.536	15423.0	15.43			2343.249	4085.0	4.09	
	1183.638	20766.0	20.78			2442.737	5883.0	5.89	
	1183.724	13336.0	13.34			3523.229	5521.0	5.52	
	1184.259	10993.0	11.00		b33	3785.873	7346.0	7.35	
	1188.615	8654.0	8.66			4651.813	4713.0	4.72	
	1191.410	11643.0	11.65			4659.790	6893.0	6.90	
	1191.624	21788.0	21.80			4748.154	3075.0	3.08	
	1192.024	17555.0	17.57			5988.556	2684.0	2.69	
	1192.146	10533.0	10.54		b61	6839.589	6103.0	6.11	
	1194.628	9410.0	9.42			7069.591	10193.0	10.20	
	1194.821	13300.0	13.31		b63	7086.970	6876.0	6.88	
	1195.390	9366.0	9.37			8456.714	3303.0	3.31	
	1197.132	13088.0	13.10			11612.968	2697.0	2.70	
	1197.820	21441.0	21.45			11638.005	3793.0	3.80	
	1197.999	10939.0	10.95		b106	11793.057	8533.0	8.54	
	1199.505	6504.0	6.51			11878.135	3284.0	3.29	
	1202.518	19356.0	19.37		y116	12943.728	8089.0	8.09	
	1202.624	9323.0	9.33		y117	13031.845	33096.0	33.12	
						13141.867	8408.0	8.41	
						13749.291	10055.0	10.06	
						13817.269	4701.0	4.70	
						13838.380	10211.0	10.22	

c) 3C_superTrip_010716_XT_00001_MHp__161... 22/12/2016 14:57:16

3C_superTrip_010716_XT_00001_MHp__161222145716#2 RT: 2.00
F: FTMS + p ESI d Full ms2 942.10@cid35.00 [245.00-2000.00]
m/z= 707.341-13988.640

	m/z	Intensity	Relative	Charge
b6	712.341	10550.0	10.90	
	805.422	14861.0	15.35	
	824.430	14164.0	14.63	
b7	840.389	62578.0	64.63	
	956.115	9544.0	9.86	
	971.604	8869.0	9.16	
	978.312	10780.0	11.13	
	979.023	12322.0	12.73	
	981.023	19733.0	20.38	
	981.095	28624.0	29.56	
	981.843	17861.0	18.45	
	983.384	26937.0	27.82	
	985.384	18429.0	19.03	
	992.281	9862.0	10.19	
	993.146	9350.0	9.66	
	993.437	11710.0	12.09	
	993.772	12182.0	12.58	
	994.273	10500.0	10.85	
	996.225	44432.0	45.89	
	996.568	34022.0	35.14	
	997.075	45137.0	46.62	
	998.962	24358.0	25.16	
	1000.037	9101.0	9.40	
	1000.250	15080.0	15.58	
	1003.535	16583.0	17.13	
	1005.112	19494.0	20.13	
	1007.775	14685.0	15.17	
	1010.234	20492.0	21.17	
	1010.690	18495.0	19.10	
	1021.241	12229.0	12.63	
	1021.626	12997.0	13.42	
	1022.315	17442.0	18.02	
	1022.548	27631.0	28.54	
	1023.233	9547.0	9.86	
	1038.592	17115.0	17.68	
	1039.568	13358.0	13.80	
	1043.643	13348.0	13.79	
	1045.371	7268.0	7.51	
	1045.912	16573.0	17.12	
	1050.031	9651.0	9.97	
	1055.652	8734.0	9.02	
	1056.460	11020.0	11.38	
	1061.659	8325.0	8.60	
	1061.733	8183.0	8.45	
	1066.076	8587.0	8.87	
	1068.074	11336.0	11.71	
	1069.572	30795.0	31.81	
	1069.652	26453.0	27.32	
	1070.823	7102.0	7.34	
	1072.553	8867.0	9.16	
	1077.745	10713.0	11.07	
	1086.247	11891.0	12.28	
	1092.867	10967.0	11.33	
	1151.598	9563.0	9.88	
	1153.120	10511.0	10.86	
	1158.611	7785.0	8.04	
b10	1168.563	7609.0	7.86	
	1176.607	10793.0	11.15	
	1262.010	8372.0	8.65	
y12	1431.880	22832.0	23.58	
y13	1544.964	8707.0	8.99	
y14	1646.012	54634.0	56.43	
	1748.007	6311.0	6.52	
	1766.018	14939.0	15.43	
	1778.054	21489.0	22.20	
	2075.175	5660.0	5.85	
	2076.177	11914.0	12.31	
y18	2093.193	68540.0	70.79	
y19	2206.277	33290.0	34.38	
	2355.200	6546.0	6.76	
	2491.397	7456.0	7.70	
y26	2928.655	8803.0	9.09	
	4051.410	7016.0	7.25	
	4837.093	5019.0	5.18	
	7025.735	7845.0	8.10	
	7795.433	4059.0	4.19	
	11791.980	3650.0	3.77	
	12687.479	15117.0	15.61	
	12798.789	15317.0	15.82	
	12816.782	31618.0	32.66	
	12926.512	12312.0	12.72	
	12943.923	16869.0	17.42	
	12975.340	4128.0	4.26	
	13013.597	25845.0	26.69	
	13032.002	96818.0	100.00	
	13142.900	23896.0	24.68	
y118	13159.913	25113.0	25.94	
	13732.714	20183.0	20.85	
	13750.265	37474.0	38.71	
	13816.287	14552.0	15.03	
y124	13824.299	5399.0	5.58	
	13836.257	39806.0	41.11	
	13983.640	4461.0	4.61	

d) 3C_superTrip_010716_XT_00001_MHp__161... 22/12/2016 15:00:03

3C_superTrip_010716_XT_00001_MHp__161222150003#2 RT: 2.00
F: FTMS + p ESI d Full ms2 883.47@cid35.00 [230.00-2000.00]
m/z= 706.347-15566.996

	m/z	Intensity	Relative	Charge
b6	711.347	4628.0	10.63	
	718.390	7649.0	17.56	
b7	840.389	34289.0	78.73	
	904.698	5103.0	11.72	
	909.220	3753.0	8.62	
	910.356	7260.0	16.67	
	911.923	4225.0	9.70	
	917.026	20377.0	46.79	
	918.292	9034.0	20.74	
	918.564	5146.0	11.82	
	920.356	7159.0	16.44	
	920.492	15432.0	35.43	
	922.292	9672.0	22.21	
	924.995	9994.0	22.95	
	928.844	6986.0	16.04	
	929.496	7005.0	16.08	
	929.571	9866.0	22.65	
	932.716	13619.0	31.27	
	933.799	10112.0	23.22	
	935.275	4646.0	10.67	
	935.639	4383.0	10.06	
	937.717	7145.0	16.41	
	946.228	5256.0	12.07	
	947.732	8845.0	20.31	
	948.220	13039.0	29.94	
	949.139	7084.0	16.27	
	949.440	8581.0	19.70	
	949.510	8575.0	19.69	
	949.653	7213.0	16.56	
	950.347	7430.0	17.06	
	952.486	5487.0	12.60	
b8	953.470	3411.0	7.83	
	956.437	7052.0	16.19	
	957.601	8037.0	18.45	
	959.011	9270.0	21.28	
	969.507	6027.0	13.84	
	972.013	5715.0	13.12	
	972.089	7671.0	17.61	
	972.176	11695.0	26.85	
	972.262	4626.0	10.62	
	973.428	8605.0	19.76	
	973.509	10723.0	24.62	
	974.326	4570.0	10.49	
	975.517	5797.0	13.31	
	983.433	7710.0	17.70	
	987.150	10927.0	25.09	
	987.300	3881.0	8.91	
	987.385	7459.0	17.13	
	987.537	8550.0	19.63	
	987.685	12603.0	28.94	
	991.103	15479.0	35.54	
	991.954	4940.0	11.34	
	1002.527	7171.0	16.46	
	1002.934	4998.0	11.48	
	1003.187	7277.0	16.71	
	1005.170	4278.0	9.82	
	1007.806	5200.0	11.94	
	1008.530	8558.0	19.65	
	1008.600	4944.0	11.35	
	1011.251	7129.0	16.37	
	1036.671	5242.0	12.04	
	1047.276	4882.0	11.21	
	1047.457	4019.0	9.23	
	1048.556	7517.0	17.26	
	1048.643	4906.0	11.26	
	1049.759	4220.0	9.69	
	1051.100	5176.0	11.88	
	1059.566	4672.0	10.73	
	1080.565	4559.0	10.47	
	1081.531	2341.0	5.38	
	1094.286	3907.0	8.97	
	1103.638	4241.0	9.74	
	1150.552	3464.0	7.95	
	1168.561	5054.0	11.60	
	1425.662	5030.0	11.55	
	1431.881	14359.0	32.97	
y13	1545.968	2692.0	6.18	
y14	1646.012	21144.0	48.55	
	1648.020	1420.0	3.26	
y15	1777.054	7764.0	17.83	
	1850.984	7548.0	17.33	
	1974.124	3684.0	8.46	
y17	1992.144	15164.0	34.82	
	2075.182	4414.0	10.13	
	2076.182	8099.0	18.60	
y19	2093.179	43553.0	100.00	
	2188.265	8979.0	20.62	
y20	2206.276	43140.0	99.05	
y21	2319.359	8374.0	19.23	
y22	2435.389	2465.0	5.66	
y23	2492.410	3848.0	8.84	
	2715.457	2760.0	6.34	
	2894.996	3132.0	7.19	
	2922.366	4886.0	11.22	
	3908.447	2084.0	4.78	
	11758.341	6272.0	14.40	
	11776.048	10721.0	24.62	
	11777.128	4050.0	9.30	
	11795.116	2665.0	6.12	
	11904.851	2513.0	5.77	
	12325.432	1908.0	4.38	
	12926.818	4974.0	11.42	
y116	12944.853	9660.0	22.18	
	13000.896	2390.0	5.49	
	13014.914	12292.0	28.22	
y117	13031.878	30366.0	69.72	
	13125.898	3149.0	7.23	
	13141.906	8877.0	20.38	
	13159.932	11260.0	25.85	
	13818.280	5429.0	12.47	
y118	13835.000	12131.0	27.85	

e)

b		Acetyl		y	b			y	b			y
130.0499	1	S	126	14069.3787	5321.7304	47	F	80 8937.7346	10347.2870	93	T	34 3866.1572
277.1183	2	F	125	13982.3467	5449.8253	48	K	79 8790.6662	10494.3554	94	F	33 3765.1095
364.1503	3	S	124	13835.2782	5596.8937	49	F	78 8662.5712	10622.4504	95	K	32 3618.0411
421.1718	4	G	123	13748.2462	5697.9414	50	T	77 8515.5028	10736.4933	96	N	31 3489.9462
549.2667	5	K	122	13691.2247	5811.0255	51	I	76 8414.4551	10849.5774	97	I	30 3375.9032
712.3301	6	Y	121	13563.1298	5912.0731	52	T	75 8301.3711	10977.6723	98	K	29 3262.8192
840.3886	7	Q	120	13400.0665	5983.1103	53	A	74 8200.3234	11064.7044	99	S	28 3134.7242
953.4727	8	L	119	13272.0079	6040.1317	54	G	73 8129.2863	11163.7728	100	V	27 3047.6922
1081.5313	9	Q	118	13158.9238	6127.1638	55	S	72 8072.2648	11264.8204	101	T	26 2948.6238
1168.5633	10	S	117	13030.8652	6255.2587	56	K	71 7985.2328	11393.8630	102	E	25 2847.5761
1296.6219	11	Q	116	12943.8332	6354.3271	57	V	70 7857.1378	11506.9471	103	L	24 2718.5335
1425.6645	12	E	115	12815.7746	6467.4112	58	I	69 7758.0694	11620.9900	104	N	23 2605.4494
1539.7074	13	N	114	12686.7320	6595.4698	59	Q	68 7644.9853	11678.0115	105	G	22 2491.4065
1686.7758	14	F	113	12572.6891	6709.5127	60	N	67 7516.9267	11793.0384	106	D	21 2434.3850
1815.8184	15	E	112	12425.6207	6838.5553	61	E	66 7402.8838	11906.1225	107	I	20 2319.3581
1886.8555	16	A	111	12296.5781	6985.6237	62	F	65 7273.8412	12019.2066	108	I	19 2206.2740
2033.9239	17	F	110	12225.5410	7086.6714	63	T	64 7126.7728	12120.2542	109	T	18 2093.1900
2164.9644	18	M	109	12078.4726	7185.7398	64	V	63 7025.7251	12234.2972	110	N	17 1992.1423
2293.0594	19	K	108	11947.4321	7242.7613	65	G	62 6926.6567	12335.3449	111	T	16 1878.0993
2364.0965	20	A	107	11819.3371	7371.8039	66	E	61 6869.6353	12466.3853	112	M	15 1777.0517
2477.1806	21	I	106	11748.3000	7500.8464	67	E	60 6740.5927	12567.4330	113	T	14 1646.0112
2534.2020	22	G	105	11635.2159	7603.8556	68	C	59 6611.5501	12680.5171	114	L	13 1544.9635
2647.2861	23	L	104	11578.1945	7732.8982	69	E	58 6508.5409	12737.5385	115	G	12 1431.8794
2744.3389	24	P	103	11465.1104	7845.9823	70	L	57 6379.4983	12852.5655	116	D	11 1374.8580
2873.3815	25	E	102	11368.0577	7975.0249	71	E	56 6266.4142	12965.6495	117	I	10 1259.8310
3002.4241	26	E	101	11239.0151	8076.0726	72	T	55 6137.3716	13064.7180	118	V	9 1146.7470
3115.5081	27	L	100	11109.9725	8207.1130	73	M	54 6036.3240	13211.7864	119	F	8 1047.6786
3228.5922	28	I	99	10996.8884	8308.1607	74	T	53 5905.2835	13339.8813	120	K	7 900.6101
3356.6508	29	Q	98	10883.8043	8365.1822	75	G	52 5804.2358	13495.9825	121	R	6 772.5152
3484.7457	30	K	97	10755.7458	8494.2248	76	E	51 5747.2143	13609.0665	122	I	5 616.4141
3541.7672	31	G	96	10627.6508	8622.3197	77	K	50 5618.1717	13696.0985	123	S	4 503.3300
3669.8621	32	K	95	10570.6293	8721.3882	78	V	49 5490.0768	13824.1935	124	K	3 416.2980
3784.8891	33	D	94	10442.5344	8849.4831	79	K	48 5391.0084	13980.2946	125	R	2 288.2030
3897.9732	34	I	93	10327.5074	8950.5308	80	T	47 5262.9134	---	126	I	1 132.1019
4026.0681	35	K	92	10214.4234	9049.5992	81	V	46 5161.8657				
4083.0896	36	G	91	10086.3284	9148.6676	82	V	45 5062.7973				
4182.1580	37	V	90	10029.3069	9276.7262	83	Q	44 4963.7289				
4269.1900	38	S	89	9930.2385	9389.8103	84	L	43 4835.6703				
4398.2326	39	E	88	9843.2065	9518.8529	85	E	42 4722.5863				
4511.3167	40	I	87	9714.1639	9575.8743	86	G	41 4593.5437				
4610.3851	41	V	86	9601.0798	9690.9013	87	D	40 4536.5222				
4738.4437	42	Q	85	9502.0114	9804.9442	88	N	39 4421.4953				
4852.4866	43	N	84	9373.9529	9933.0392	89	K	38 4307.4523				
4909.5081	44	G	83	9259.9099	10046.1232	90	L	37 4179.3574				
5037.6030	45	K	82	9202.8885	10145.1916	91	V	36 4066.2733				
5174.6619	46	H	81	9074.7935	10246.2393	92	T	35 3967.2049				

Supporting Information Section II

Results of the bottom-up proteomics analysis performed on proteins of interest present in the >30 KDa and in the <30 KDa fractions.

Fig. S1. Thymosin β4, fragment 21-39.
$[M+H]^+$ 2229.12430 Da (Theor. 2229.1245) RT: 15.93 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+3H]^{3+}$ monoisotopic m/z: 743.71295 Da (-0.1 mmu/-0.13 ppm). Analysis performed by Proteome discoverer software
Identified with: Sequest HT (v1.3); XCorr:3.71, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.8 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Thymosin beta-4 OS=Homo sapiens GN=TMSB4X PE=1 SV=2 - [TYB4_HUMAN]

#1	b^+	b^{2+}	b^{3+}	Seq.	y^+	y^{2+}	y^{3+}	#2
1	102.05496	51.53112	34.68984	T				19
2	231.09756	116.05242	77.70404	E	2128.07691	1064.54209	710.03049	18
3	332.14524	166.57626	111.38660	T	1999.03431	1000.02079	667.01629	17
4	460.20382	230.60555	154.07279	Q	1897.98663	949.49695	633.33373	16
5	589.24642	295.12685	197.08699	E	1769.92805	885.46766	590.64753	15
6	717.34139	359.17433	239.78531	K	1640.88545	820.94636	547.63333	14
7	831.38432	416.19580	277.79962	N	1512.79048	756.89888	504.93501	13
8	928.43709	464.72218	310.15055	P	1398.74755	699.87741	466.92070	12
9	1041.52116	521.26422	347.84524	L	1301.69478	651.35103	434.56978	11
10	1138.57393	569.79060	380.19616	P	1188.61071	594.80899	396.87509	10
11	1225.60596	613.30662	409.20684	S	1091.55794	546.28261	364.52416	9
12	1353.70093	677.35410	451.90516	K	1004.52591	502.76659	335.51349	8
13	1482.74353	741.87540	494.91936	E	876.43094	438.71911	292.81516	7
14	1583.79121	792.39924	528.60192	T	747.38834	374.19781	249.80096	6
15	1696.87528	848.94128	566.29661	I	646.34066	323.67397	216.11840	5
16	1825.91788	913.46258	609.31081	E	533.25659	267.13193	178.42371	4
17	1953.97646	977.49187	651.99700	Q	404.21399	202.61063	135.40951	3
18	2083.01906	1042.01317	695.01120	E	276.15541	138.58134	92.72332	2
19				K	147.11281	74.06004	49.70912	1

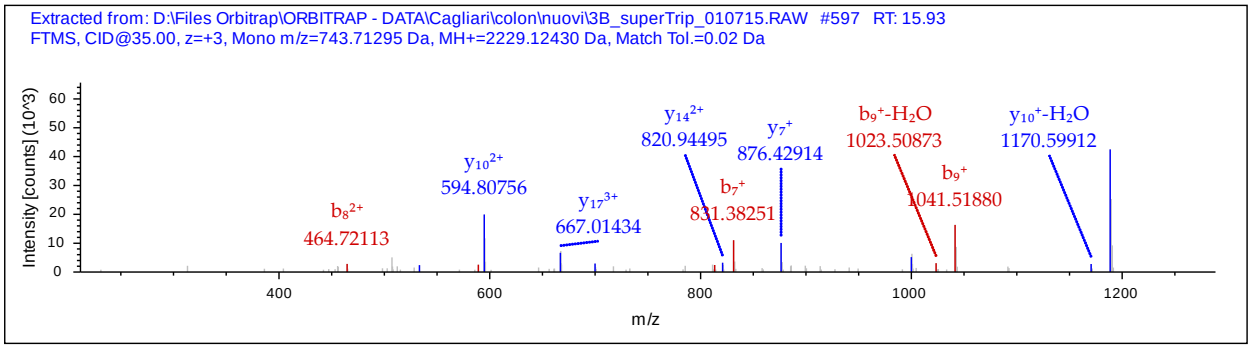


Fig. S2a. Prothymosin α, Fragment 91-101.
$[M+H]^+$ 1437.5496 Da (Theor. 1437.5499) RT: 9.32 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 719.27844 Da (-0.22 mmu/-0.31 ppm). Analysis performed by Proteome discoverer software
Identified with: Sequest HT (v1.3); XCorr:4.14, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.8 Da
Fragments used for search: b; b-H_2O; y; y-H_2O; y-NH_3
Protein references (1):
- Prothymosin alpha OS=Homo sapiens GN=PTMA PE=1 SV=2 - [PTMA_HUMAN]

	b^+	b^{2+}	Seq	y^+	y2+	
1	72.04440	36.52584	A			13
2	143.08152	72.04440	A	1366.51293	683.76010	12
3	272.12412	136.56570	E	1295.47581	648.24154	11
4	387.15107	194.07917	D	1166.43321	583.72024	10
5	502.17802	251.59265	D	1051.40626	526.20677	9
6	631.22062	316.11395	E	936.37931	468.69329	8
7	746.24757	373.62742	D	807.33671	404.17199	7
8	861.27452	431.14090	D	692.30976	346.65852	6
9	976.30147	488.65437	D	577.28281	289.14504	5
10	1075.36989	538.18858	V	462.25586	231.63157	4
11	1190.39684	595.70206	D	363.18744	182.09736	3
12	1291.44452	646.22590	T	248.16049	124.58388	2
13			K	147.11281	74.06004	1

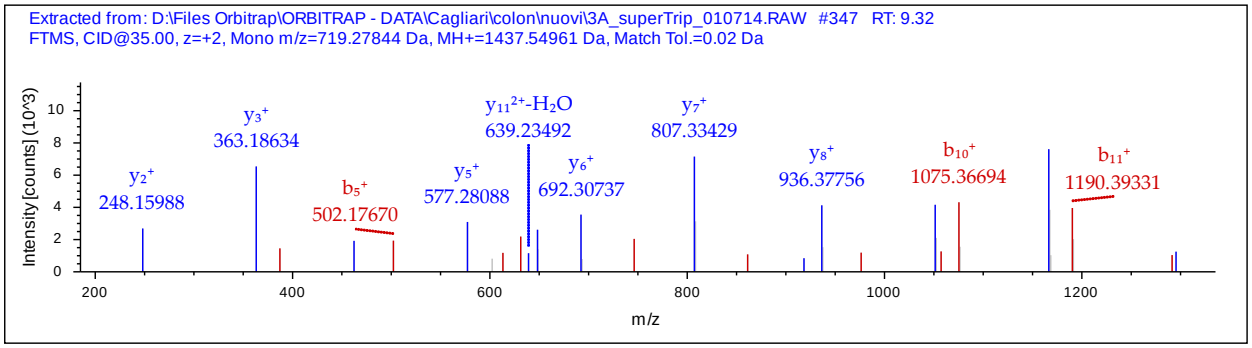


Fig. S2b. Prothymosin α, Fragment 91-102.
$[M+H]^+$ 1565.6434 Da (Theor. 1565.6449) RT: 7.05 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 783.32532 Da (-0.83 mmu/-1.06 ppm). Analysis performed by Proteome discoverer software
Identified with: Sequest HT (v1.3); XCorr:4.11, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.8 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Prothymosin alpha OS=Homo sapiens GN=PTMA PE=1 SV=2 - [PTMA_HUMAN]

	b^+	b^{2+}	Seq	y^+	y2+	
1	72.04440	36.52584	A			14
2	143.08152	72.04440	A	1494.60790	747.80759	13
3	272.12412	136.56570	E	1423.57078	712.28903	12
4	387.15107	194.07917	D	1294.52818	647.76773	11
5	502.17802	251.59265	D	1179.50123	590.25425	10
6	631.22062	316.11395	E	1064.47428	532.74078	9
7	746.24757	373.62742	D	935.43168	468.21948	8
8	861.27452	431.14090	D	820.40473	410.70600	7
9	976.30147	488.65437	D	705.37778	353.19253	6
10	1075.36989	538.18858	V	590.35083	295.67905	5
11	1190.39684	595.70206	D	491.28241	246.14484	4
12	1291.44452	646.22590	T	376.25546	188.63137	3
13	1419.53949	710.27338	K	275.20778	138.10753	2
14			K	147.11281	74.06004	1

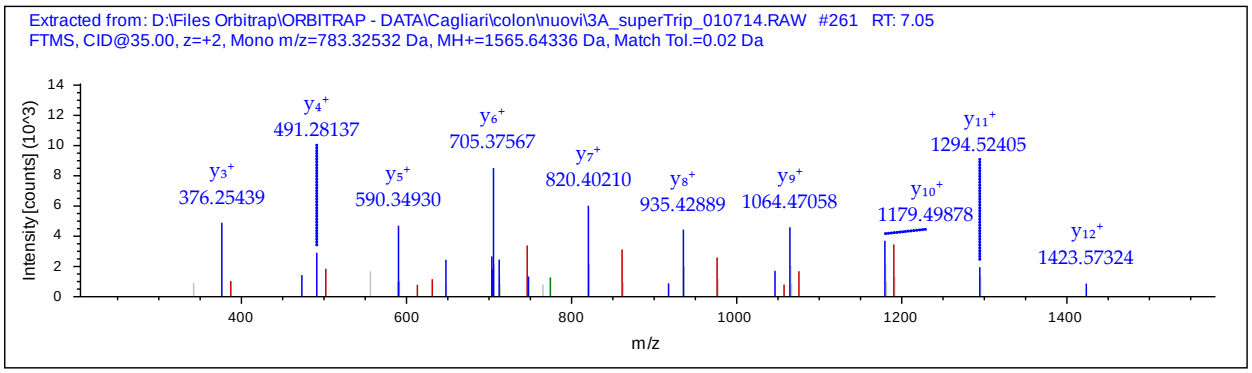


Fig. S3a. SH3BGRL3 protein, Fragment 6-15.
$[M+H]^+$ 1056.53093 Da (Theor. 1056.5320) RT: 14.84 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 528.76910 Da (-0.56 mmu/-1.05 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:3.29, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; y; y-H_2O; y-NH_3
Protein references (1):
- SH3 domain-binding glutamic acid-rich-like protein 3 OS=Homo sapiens GN=SH3BGRL3 PE=1 SV=1 - [SH3L3_HUMAN]

#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	100.07570	50.54149	V			10
2	263.13902	132.07315	Y	957.46362	479.23545	9
3	350.17105	175.58916	S	794.40030	397.70379	8
4	451.21873	226.11300	T	707.36827	354.18777	7
5	538.25076	269.62902	S	606.32059	303.66393	6
6	637.31918	319.16323	V	519.28856	260.14792	5
7	738.36686	369.68707	T	420.22014	210.61371	4
8	795.38833	398.19780	G	319.17246	160.08987	3
9	882.42036	441.71382	S	262.15099	131.57913	2
10			R	175.11896	88.06312	1

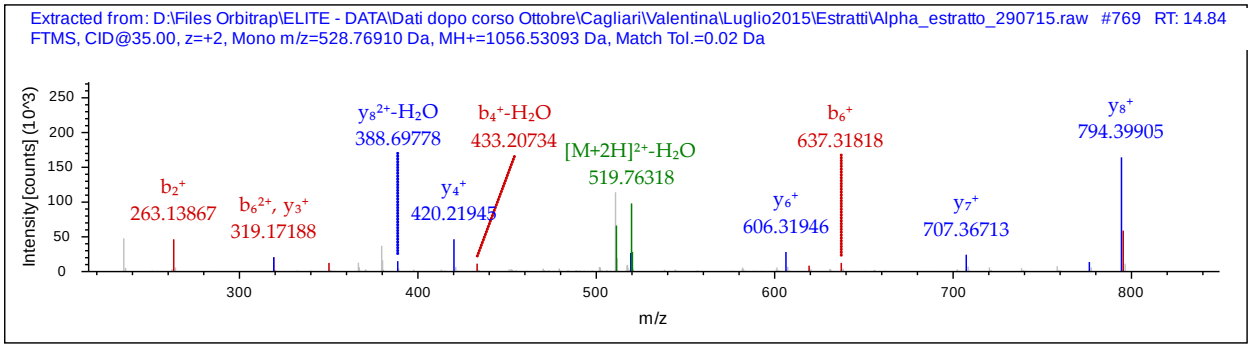


Fig. S3b. SH3BGRL3 protein, Fragment 33-51.
$[M+H]^+$ 2307.14078 Da (Theor. 2307.1397) RT: 25.66 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+3H]^{3+}$ monoisotopic m/z: 769.71844 Da (+0.3 mmu/+0.39 ppm). Analysis performed by Proteome discoverer software
Identified with: Sequest HT (v1.3); XCorr:3.52, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.8 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- SH3 domain-binding glutamic acid-rich-like protein 3 OS=Homo sapiens GN=SH3BGRL3 PE=1 SV=1 - [SH3L3_HUMAN]

b^+	b^{2+}	b^{3+}	Seq.	y^+	y^{2+}	y^{3+}	#2
114.09135	57.54931	38.70197	I				19
242.14993	121.57860	81.38816	Q	2194.05580	1097.53154	732.02345	18
405.21325	203.11026	135.74260	Y	2065.99722	1033.50225	689.33726	17
533.27183	267.13955	178.42879	Q	1902.93390	951.97059	634.98282	16
646.35590	323.68159	216.12348	L	1774.87532	887.94130	592.29662	15
745.42432	373.21580	249.14629	V	1661.79125	831.39926	554.60193	14
860.45127	430.72927	287.48861	D	1562.72283	781.86505	521.57913	13
973.53534	487.27131	325.18330	I	1447.69588	724.35158	483.23681	12
1060.56737	530.78732	354.19397	S	1334.61181	667.80954	445.54212	11
1188.62595	594.81661	396.88017	Q	1247.57978	624.29353	416.53144	10
1303.65290	652.33009	435.22248	D	1119.52120	560.26424	373.84525	9
1417.69583	709.35155	473.23679	N	1004.49425	502.75076	335.50293	8
1488.73295	744.87011	496.91583	A	890.45132	445.72930	297.48862	7
1601.81702	801.41215	534.61052	L	819.41420	410.21074	273.80958	6
1757.91814	879.46271	586.64423	R	706.33013	353.66870	236.11489	5
1872.94509	936.97618	624.98655	D	550.22901	275.61814	184.08119	4
2001.98769	1001.49748	668.00075	E	435.20206	218.10467	145.73887	3
2133.02819	1067.01773	711.68091	M	306.15946	153.58337	102.72467	2
			R	175.11896	88.06312	59.04450	1

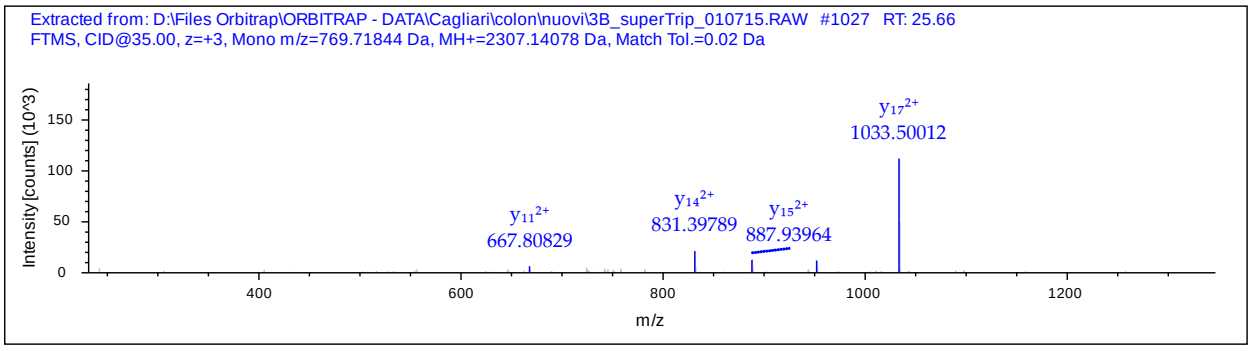


Fig. S3c. SH3BGRL3 protein, Fragment 59-91.
$[M+H]^+$ 3815.79929 Da (Theor. 3815.8108) RT: 36.10 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+3H]^{3+}$ monoisotopic m/z: 1272.60461 Da (-3.92 mmu/-3.08 ppm). Analysis performed by Proteome discoverer software
C13-Carbamidomethyl (57.02146 Da)
Identified with: Sequest HT (v1.3); XCorr:6.57, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- SH3 domain-binding glutamic acid-rich-like protein 3 OS=Homo sapiens GN=SH3BGRL3 PE=1 SV=1 - [SH3L3_HUMAN]

b^+	b^{2+}	b^{3+}	Seq.	y^+	y^{2+}	y^{3+}	#2
72.04440	36.52584	24.68632	A				33
173.09208	87.04968	58.36888	T	3744.77393	1872.89060	1248.92949	32
270.14485	135.57606	90.71980	P	3643.72625	1822.36676	1215.24693	31
367.19762	184.10245	123.07072	P	3546.67348	1773.84038	1182.89601	30
495.25620	248.13174	165.75692	Q	3449.62071	1725.31399	1150.54509	29
608.34027	304.67377	203.45161	I	3321.56213	1661.28470	1107.85889	28
707.40869	354.20798	236.47441	V	3208.47806	1604.74267	1070.16420	27
821.45162	411.22945	274.48872	N	3109.40964	1555.20846	1037.14140	26
878.47309	439.74018	293.49588	G	2995.36671	1498.18699	999.12709	25
993.50004	497.25366	331.83820	D	2938.34524	1469.67626	980.11993	24
1121.55862	561.28295	374.52439	Q	2823.31829	1412.16278	941.77761	23
1284.62194	642.81461	428.87883	Y	2695.25971	1348.13349	899.09142	22
1444.65259	722.82993	482.22238	C-Carbamidomethyl	2532.19639	1266.60183	844.73698	21
1501.67406	751.34067	501.22954	G	2372.16573	1186.58650	791.39343	20
1616.70101	808.85414	539.57185	D	2315.14426	1158.07577	772.38627	19
1779.76433	890.38580	593.92629	Y	2200.11731	1100.56229	734.04395	18
1908.80693	954.90710	636.94049	E	2037.05399	1019.03063	679.68951	17
2021.89100	1011.44914	674.63518	L	1908.01139	954.50933	636.67531	16
2168.95942	1084.98335	723.65799	F	1794.92732	897.96730	598.98062	15
2268.02784	1134.51756	756.68080	V	1647.85890	824.43309	549.95782	14
2397.07044	1199.03886	799.69500	E	1548.79048	774.89888	516.93501	13
2468.10756	1234.55742	823.37404	A	1419.74788	710.37758	473.92081	12
2567.17598	1284.09163	856.39684	V	1348.71076	674.85902	450.24177	11
2696.21858	1348.61293	899.41104	E	1249.64234	625.32481	417.21896	10
2824.27716	1412.64222	942.09724	Q	1120.59974	560.80351	374.20476	9
2938.32009	1469.66368	980.11155	N	992.54116	496.77422	331.51857	8
3039.36777	1520.18752	1013.79411	T	878.49823	439.75275	293.50426	7
3152.45184	1576.72956	1051.48880	L	777.45055	389.22891	259.82170	6
3280.51042	1640.75885	1094.17499	Q	664.36648	332.68688	222.12701	5
3409.55302	1705.28015	1137.18919	E	536.30790	268.65759	179.44082	4
3556.62144	1778.81436	1186.21200	F	407.26530	204.13629	136.42662	3
3669.70551	1835.35639	1223.90669	L	260.19688	130.60208	87.40381	2
			K	147.11281	74.06004	49.70912	1

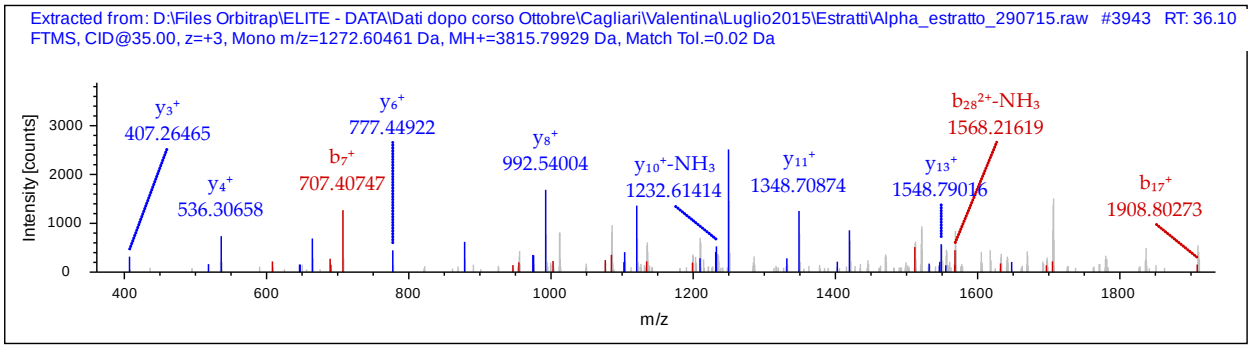


Fig. S4a. FABP1, Fragment 21-31.
$[M+H]^+$ 1210.70391 Da (Theor. 1210.7041) RT: 25.12 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 605.85559 Da (-0.16 mmu/-0.27 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:3.49, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Fatty acid-binding protein, liver OS=Homo sapiens GN=FABP1 PE=1 SV=1 - [FABPL_HUMAN]

b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
72.04440	36.52584	A			11
185.12847	93.06787	I	1139.66711	570.33719	10
242.14994	121.57861	G	1026.58304	513.79516	9
355.23401	178.12064	L	969.56157	485.28442	8
452.28678	226.64703	P	856.47750	428.74239	7
581.32938	291.16833	E	759.42473	380.21600	6
710.37198	355.68963	E	630.38213	315.69470	5
823.45605	412.23166	L	501.33953	251.17340	4
936.54012	468.77370	I	388.25546	194.63137	3
1064.59870	532.80299	Q	275.17139	138.08933	2
		K	147.11281	74.06004	1

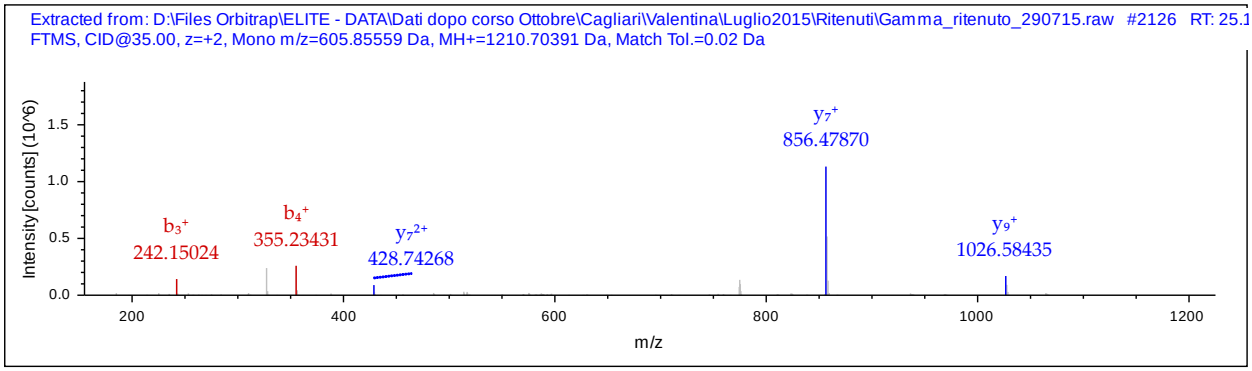


Fig. S4b. FABP1, Fragment 50-57.
$[M+H]^+$ 824.45323 Da (Theor. 824.4512) RT: 17.92 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 412.73026 Da (+0.98 mmu/+2.37 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:1.94, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; y; y-H_2O; y-NH_3
Protein references (1):
- Fatty acid-binding protein, liver OS=Homo sapiens GN=FABP1 PE=1 SV=1 - [FABPL_HUMAN]

b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
148.07570	74.54149	F			8
249.12338	125.06533	T	677.38286	339.19507	7
362.20745	181.60736	I	576.33518	288.67123	6
463.25513	232.13120	T	463.25111	232.12919	5
534.29225	267.64976	A	362.20343	181.60535	4
591.31372	296.16050	G	291.16631	146.08679	3
678.34575	339.67651	S	234.14484	117.57606	2
		K	147.11281	74.06004	1

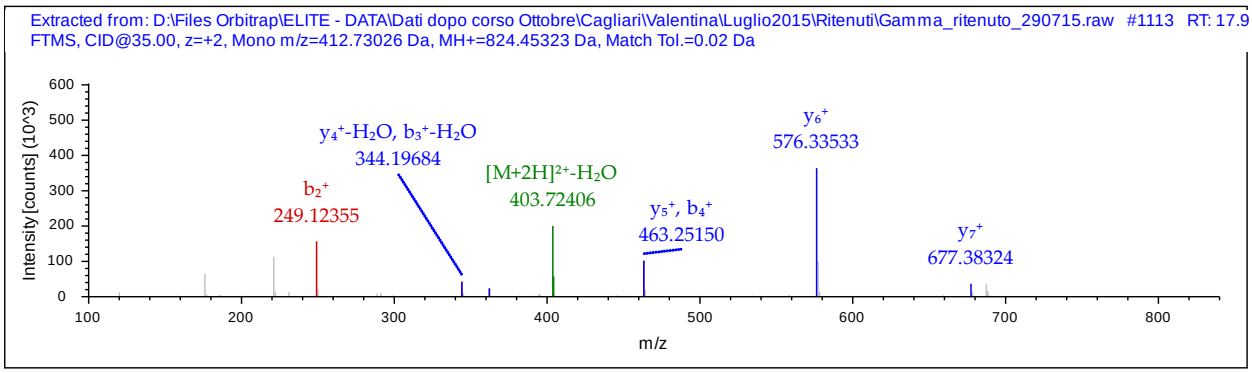


Fig. S4c. FABP1, Fragment 81-90.
$[M+H]^+$ 1102.57573 Da (Theor. 1102.5738) RT: 17.30 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 551.79150 Da (+0.9 mmu/+1.63 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:3.61, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Fatty acid-binding protein, liver OS=Homo sapiens GN=FABP1 PE=1 SV=1 - [FABPL_HUMAN]

1	102.05496	51.53112	T			10
2	201.12338	101.06533	V	1001.52625	501.26676	9
3	300.19180	150.59954	V	902.45783	451.73255	8
4	428.25038	214.62883	Q	803.38941	402.19834	7
5	541.33445	271.17086	L	675.33083	338.16905	6
6	670.37705	335.69216	E	562.24676	281.62702	5
7	727.39852	364.20290	G	433.20416	217.10572	4
8	842.42547	421.71637	D	376.18269	188.59498	3
9	956.46840	478.73784	N	261.15574	131.08151	2
10			K	147.11281	74.06004	1

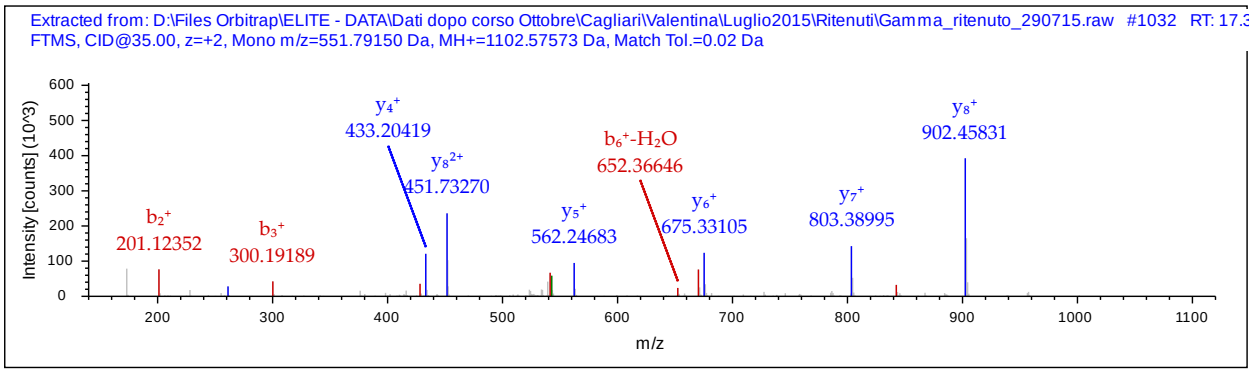


Fig. S4d. FABP1, Fragment 81-97.
$[M+H]^+$ 1791.98709 Da (Theor. 1791.9850) RT: 23.81 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+3H]^{3+}$ monoisotopic m/z: 598.00055 Da (+0.64 mmu/+1.07 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:2.86, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Fatty acid-binding protein, liver OS=Homo sapiens GN=FABP1 PE=1 SV=1 - [FABPL_HUMAN]

b^+	b^{2+}	b^{3+}	Seq.	y^+	y^{2+}	y^{3+}	#2
102.05496	51.53112	34.68984	T				16
201.12338	101.06533	67.71264	V	1690.93749	845.97238	564.31735	15
300.19180	150.59954	100.73545	V	1591.86907	796.43817	531.29454	14
428.25038	214.62883	143.42164	Q	1492.80065	746.90396	498.27173	13
541.33445	271.17086	181.11633	L	1364.74207	682.87467	455.58554	12
670.37705	335.69216	224.13053	E	1251.65800	626.33264	417.89085	11
727.39852	364.20290	243.13769	G	1122.61540	561.81134	374.87665	10
842.42547	421.71637	281.48001	D	1065.59393	533.30060	355.86949	9
956.46840	478.73784	319.49432	N	950.56698	475.78713	317.52718	8
1084.56337	542.78532	362.19264	K	836.52405	418.76566	279.51287	7
1197.64744	599.32736	399.88733	L	708.42908	354.71818	236.81454	6
1296.71586	648.86157	432.91014	V	595.34501	298.17614	199.11985	5
1397.76354	699.38541	466.59270	T	496.27659	248.64193	166.09705	4
1498.81122	749.90925	500.27526	T	395.22891	198.11809	132.41449	3
1645.87964	823.44346	549.29806	F	294.18123	147.59425	98.73193	2
			K	147.11281	74.06004	49.70912	1

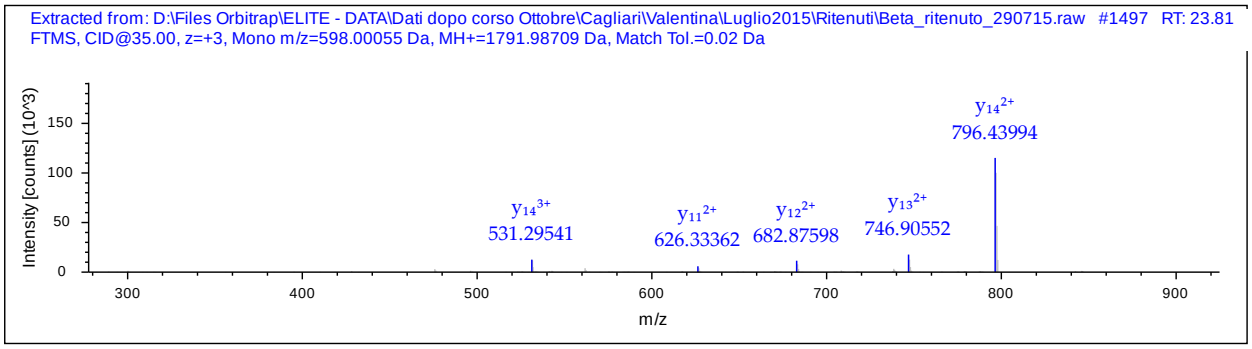


Fig. S5a. Carbonic anhydrase 1, Fragment 20-35.
$[M+H]^+$ 1742.90654 Da (Theor. 1742.9071) RT: 21.45 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 871.95691Da (-0.35 mmu/-0.4 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:4.85, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.8 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Carbonic anhydrase 1 OS=Homo sapiens GN=CA1 PE=1 SV=2 - [CAH1_HUMAN]

#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	114.09135	57.54931	L			16
2	277.15467	139.08097	Y	1629.82317	815.41522	15
3	374.20744	187.60736	P	1466.75985	733.88356	14
4	487.29151	244.14939	I	1369.70708	685.35718	13
5	558.32863	279.66795	A	1256.62301	628.81514	12
6	672.37156	336.68942	N	1185.58589	593.29658	11
7	729.39303	365.20015	G	1071.54296	536.27512	10
8	843.43596	422.22162	N	1014.52149	507.76438	9
9	957.47889	479.24308	N	900.47856	450.74292	8
10	1085.53747	543.27237	Q	786.43563	393.72145	7
11	1172.56950	586.78839	S	658.37705	329.69216	6
12	1269.62227	635.31477	P	571.34502	286.17615	5
13	1368.69069	684.84898	V	474.29225	237.64976	4
14	1483.71764	742.36246	D	375.22383	188.11555	3
15	1596.80171	798.90449	I	260.19688	130.60208	2
16			K	147.11281	74.06004	1

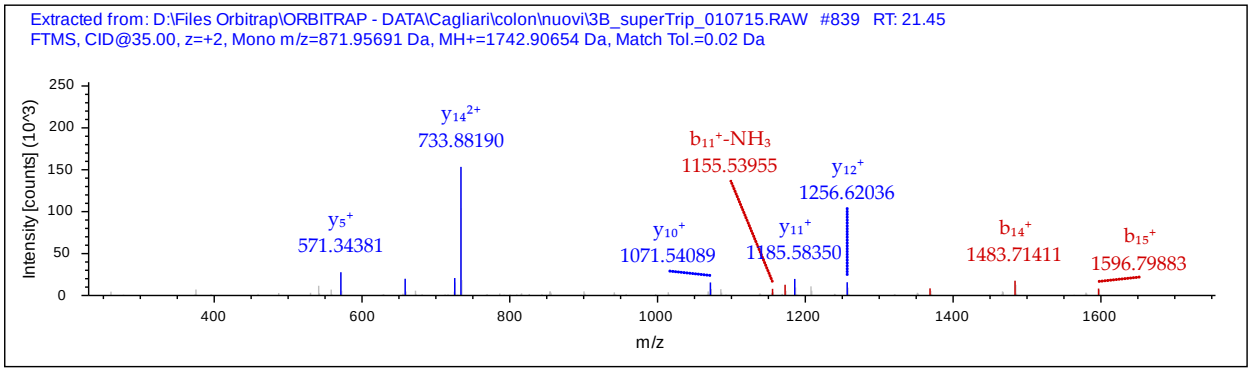
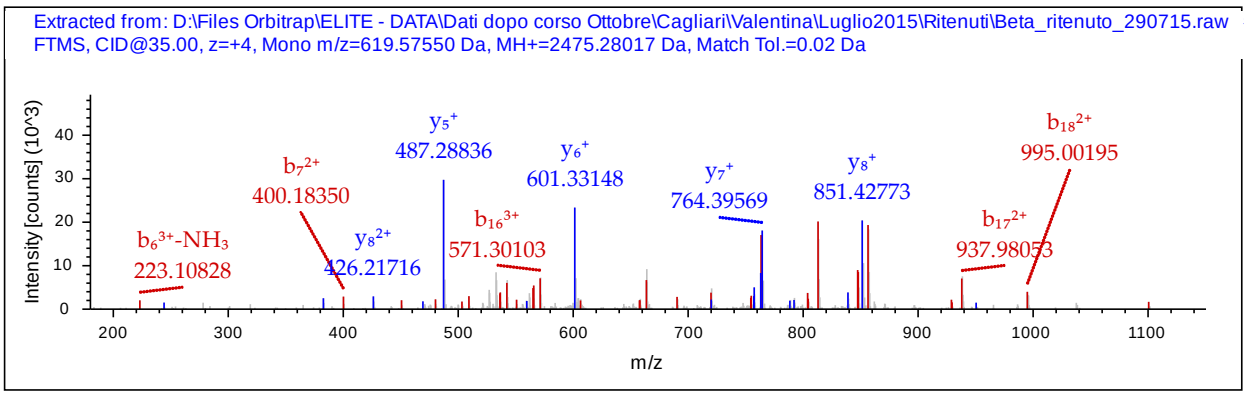


Fig. S5b. Carbonic anhydrase 1, Fragment 36-58.
$[M+H]^+$ 2475.27225 Da (Theor. 2475.2725). MS/MS sequencing results by two different samples:
- Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+4H]^{4+}$ monoisotopic m/z: 619.57550 Da (+1.88 mmu/+3.04 ppm), RT: 19.65 min. Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr: 5.67, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
- Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+3H]^{3+}$ monoisotopic m/z: 825.76227 Da (-0.13 mmu/-0.16 ppm), RT: 18.20 min. Analysis performed by Proteome discoverer software.

Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Carbonic anhydrase 1 OS=Homo sapiens GN=CA1 PE=1 SV=2 - [CAH1_HUMAN]

#1	b^+	b^{2+}	b^{3+}	b^{4+}	Seq.	y^+	y^{2+}	y^{3+}	y^{4+}	#2
1	102.05496	51.53112	34.68984	26.26920	T					23
2	189.08699	95.04713	63.70051	48.02720	S	2374.22496	1187.61612	792.07984	594.31170	22
3	318.12959	159.56843	106.71471	80.28785	E	2287.19293	1144.10010	763.06916	572.55369	21
4	419.17727	210.09227	140.39727	105.54977	T	2158.15033	1079.57880	720.05496	540.29304	20
5	547.27224	274.13976	183.09560	137.57352	K	2057.10265	1029.05496	686.37240	515.03112	19
6	684.33115	342.66921	228.78190	171.83824	H	1929.00768	965.00748	643.67408	483.00738	18
7	799.35810	400.18269	267.12422	200.59498	D	1791.94877	896.47802	597.98777	448.74265	17
8	900.40578	450.70653	300.80678	225.85690	T	1676.92182	838.96455	559.64546	419.98591	16
9	987.43781	494.22254	329.81745	247.61491	S	1575.87414	788.44071	525.96290	394.72399	15
10	1100.52188	550.76458	367.51214	275.88593	L	1488.84211	744.92469	496.95222	372.96599	14
11	1228.61685	614.81206	410.21047	307.90967	K	1375.75804	688.38266	459.25753	344.69497	13
12	1325.66962	663.33845	442.56139	332.17286	P	1247.66307	624.33517	416.55921	312.67123	12
13	1438.75369	719.88048	480.25608	360.44388	I	1150.61030	575.80879	384.20828	288.40803	11
14	1525.78572	763.39650	509.26676	382.20189	S	1037.52623	519.26675	346.51359	260.13702	10
15	1624.85414	812.93071	542.28956	406.96899	V	950.49420	475.75074	317.50292	238.37901	9
16	1711.88617	856.44672	571.30024	428.72700	S	851.42578	426.21653	284.48011	213.61190	8
17	1874.94949	937.97838	625.65468	469.49283	Y	764.39375	382.70051	255.46943	191.85390	7
18	1988.99242	994.99985	663.66899	498.00356	N	601.33043	301.16885	201.11499	151.08807	6
19	2086.04519	1043.52623	696.01991	522.26675	P	487.28750	244.14739	163.10068	122.57733	5
20	2157.08231	1079.04479	719.69895	540.02603	A	390.23473	195.62100	130.74976	98.31414	4
21	2258.12999	1129.56863	753.38151	565.28795	T	319.19761	160.10244	107.07072	80.55486	3
22	2329.16711	1165.08719	777.06055	583.04723	A	218.14993	109.57860	73.38816	55.29294	2
23					K	147.11281	74.06004	49.70912	37.53366	1



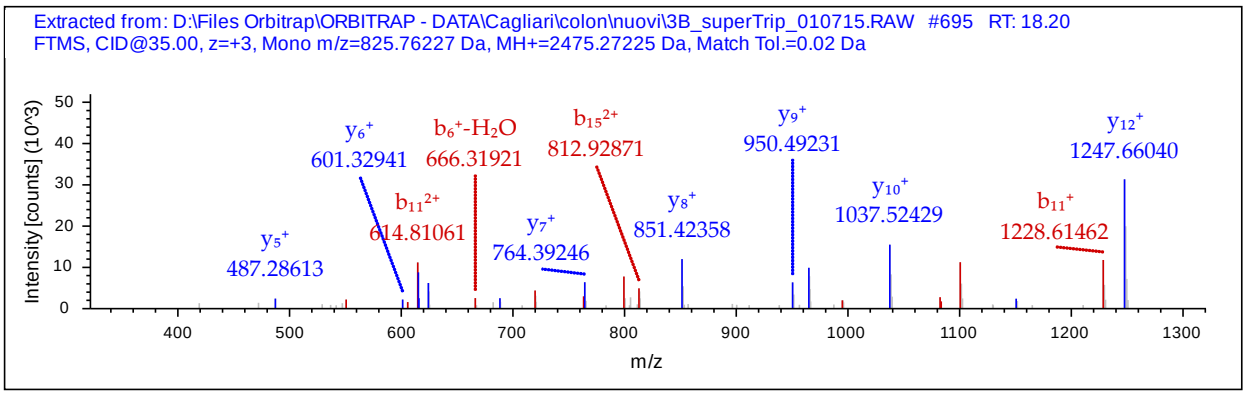


Fig. S5c. Carbonic anhydrase 1, Fragment 41-58.
$[M+H]^+$ 1929.00766 Da (Theor. 1929.0076) RT: 20.30 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+3H]^{3+}$ monoisotopic m/z: 643.67407 Da (+1.09 mmu/+1.7 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr: 6.09, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Carbonic anhydrase 1 OS=Homo sapiens GN=CA1 PE=1 SV=2 - [CAH1_HUMAN]

b^+	b^{2+}	b^{3+}	Seq.	y^+	y^{2+}	y^{3+}	#2
138.06619	69.53673	46.69358	H				18
253.09314	127.05021	85.03590	D	1791.94877	896.47802	597.98777	17
354.14082	177.57405	118.71846	T	1676.92182	838.96455	559.64546	16
441.17285	221.09006	147.72913	S	1575.87414	788.44071	525.96290	15
554.25692	277.63210	185.42382	L	1488.84211	744.92469	496.95222	14
682.35189	341.67958	228.12215	K	1375.75804	688.38266	459.25753	13
779.40466	390.20597	260.47307	P	1247.66307	624.33517	416.55921	12
892.48873	446.74800	298.16776	I	1150.61030	575.80879	384.20828	11
979.52076	490.26402	327.17844	S	1037.52623	519.26675	346.51359	10
1078.58918	539.79823	360.20124	V	950.49420	475.75074	317.50292	9
1165.62121	583.31424	389.21192	S	851.42578	426.21653	284.48011	8
1328.68453	664.84590	443.56636	Y	764.39375	382.70051	255.46943	7
1442.72746	721.86737	481.58067	N	601.33043	301.16885	201.11499	6
1539.78023	770.39375	513.93159	P	487.28750	244.14739	163.10068	5
1610.81735	805.91231	537.61063	A	390.23473	195.62100	130.74976	4
1711.86503	856.43615	571.29319	T	319.19761	160.10244	107.07072	3
1782.90215	891.95471	594.97223	A	218.14993	109.57860	73.38816	2
			K	147.11281	74.06004	49.70912	1

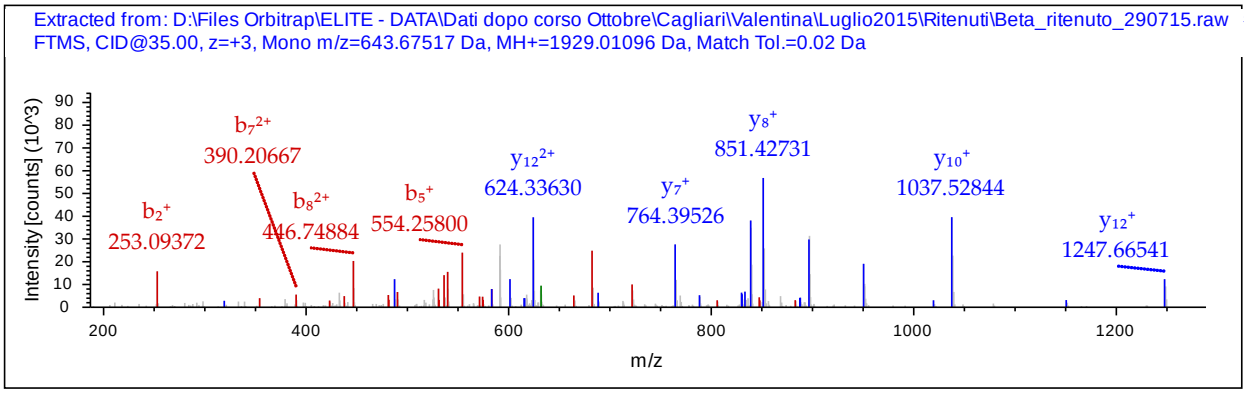


Fig. S5d. Carbonic anhydrase 1, Fragment 59-77.
$[M+H]^+$ 2256.04227Da (Theor. 2256.0428) RT: 23.86 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+3H]^{3+}$ monoisotopic m/z: 752.68561 Da (+1.68 mmu/+2.23 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:6.88, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Carbonic anhydrase 1 OS=Homo sapiens GN=CA1 PE=1 SV=2 - [CAH1_HUMAN]

#1	b^+	b^{2+}	b^{3+}	Seq.	y^+	y^{2+}	y^{3+}	#2
1	130.04988	65.52858	44.02148	E				19
2	243.13395	122.07061	81.71617	I	2127.00032	1064.00380	709.67162	18
3	356.21802	178.61265	119.41086	I	2013.91625	1007.46176	671.97693	17
4	470.26095	235.63411	157.42517	N	1900.83218	950.91973	634.28224	16
5	569.32937	285.16832	190.44797	V	1786.78925	893.89826	596.26793	15
6	626.35084	313.67906	209.45513	G	1687.72083	844.36405	563.24513	14
7	763.40975	382.20851	255.14143	H	1630.69936	815.85332	544.23797	13
8	850.44178	425.72453	284.15211	S	1493.64045	747.32386	498.55167	12
9	997.51020	499.25874	333.17492	F	1406.60842	703.80785	469.54099	11
10	1134.56911	567.78819	378.86122	H	1259.54000	630.27364	420.51818	10
11	1233.63753	617.32240	411.88403	V	1122.48109	561.74418	374.83188	9
12	1347.68046	674.34387	449.89834	N	1023.41267	512.20997	341.80907	8
13	1494.74888	747.87808	498.92114	F	909.36974	455.18851	303.79476	7
14	1623.79148	812.39938	541.93534	E	762.30132	381.65430	254.77196	6
15	1738.81843	869.91285	580.27766	D	633.25872	317.13300	211.75776	5
16	1852.86136	926.93432	618.29197	N	518.23177	259.61952	173.41544	4
17	1967.88831	984.44779	656.63429	D	404.18884	202.59806	135.40113	3
18	2081.93124	1041.46926	694.64860	N	289.16189	145.08458	97.05881	2
19				R	175.11896	88.06312	59.04450	1

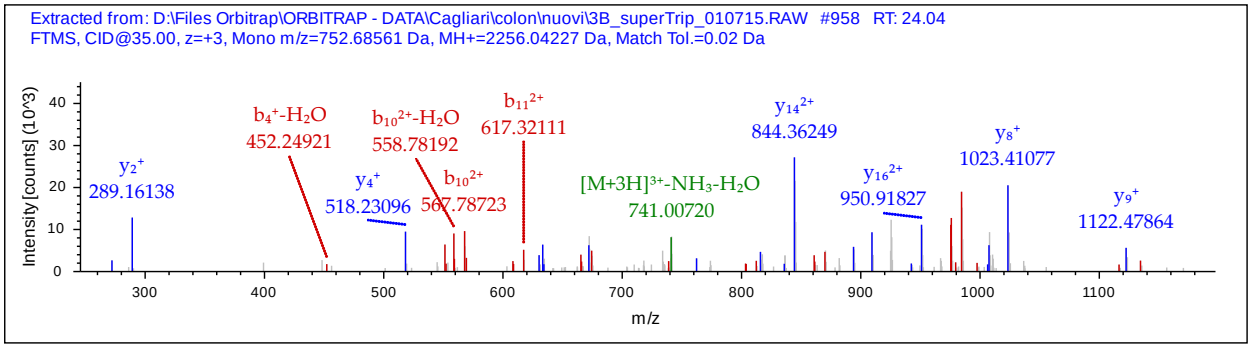


Fig. S5e. Carbonic anhydrase 1, Fragment 82-90.
$[M+H]^+$ 985.43956Da (Theor. 985.4374) RT: 18.65 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 493.22342 Da (+1.07 mmu/+2.17 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:2.06, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; y; y-H_2O; y-NH_3
Protein references (1):
- Carbonic anhydrase 1 OS=Homo sapiens GN=CA1 PE=1 SV=2 - [CAH1_HUMAN]

#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	58.02875	29.51801	G			9
2	115.05022	58.02875	G	928.41595	464.71161	8
3	212.10299	106.55513	P	871.39448	436.20088	7
4	359.17141	180.08934	F	774.34171	387.67449	6
5	446.20344	223.60536	S	627.27329	314.14028	5
6	561.23039	281.11883	D	540.24126	270.62427	4
7	648.26242	324.63485	S	425.21431	213.11079	3
8	811.32574	406.16651	Y	338.18228	169.59478	2
9			R	175.11896	88.06312	1

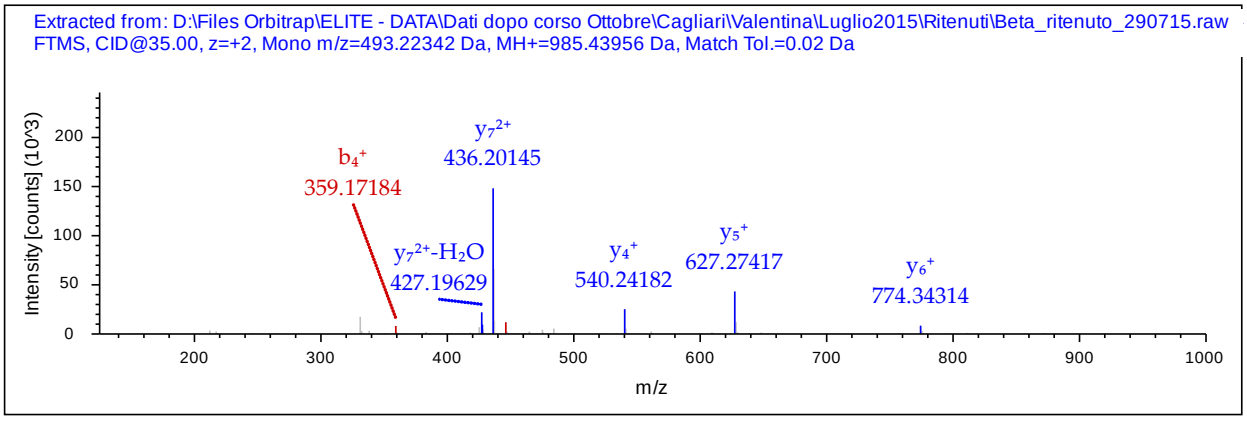


Fig. S5f. Carbonic anhydrase 1, Fragment 161-169.
$[M+H]^+$ 970.59496 Da (Theor. 970.5931) RT: 23.37 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 485.80112 Da (+0.87 mmu/+1.8 ppm). Analysis performed by Proteome discoverer software.
dentified with: Sequest HT (v1.3); XCorr:3.66, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Carbonic anhydrase 1 OS=Homo sapiens GN=CA1 PE=1 SV=2 - [CAH1_HUMAN]

b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
100.07570	50.54149	V			9
213.15977	107.08352	L	871.52479	436.26603	8
328.18672	164.59700	D	758.44072	379.72400	7
399.22384	200.11556	A	643.41377	322.21052	6
512.30791	256.65759	L	572.37665	286.69196	5
640.36649	320.68688	Q	459.29258	230.14993	4
711.40361	356.20544	A	331.23400	166.12064	3
824.48768	412.74748	I	260.19688	130.60208	2
		K	147.11281	74.06004	1

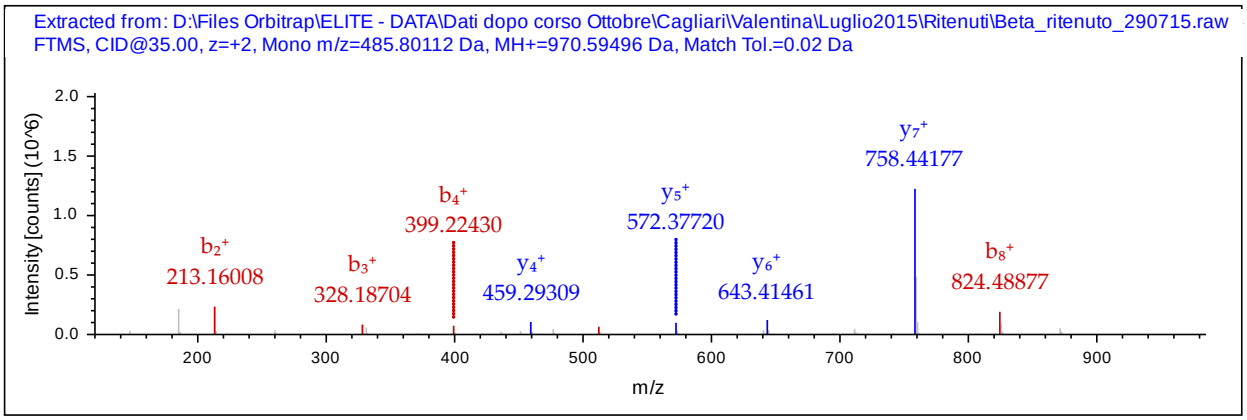


Fig. S5g. Carbonic anhydrase 1, Fragment 174-214.
$[M+H]^+$ 4754.36098 Da (Theor. 4754.3639) RT: 36.65 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+4H]^{4+}$ monoisotopic m/z: 1189.34570 Da (-0.79 mmu/-0.66 ppm). Analysis performed by Proteome discoverer software.
C40-Carbamidomethyl (57.02146 Da)
Identified with: Sequest HT (v1.3); XCorr:3.94, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Carbonic anhydrase 1 OS=Homo sapiens GN=CA1 PE=1 SV=2 - [CAH1_HUMAN]

#1	b^+	b^{2+}	b^{3+}	b^{4+}	Seq.	y^+	y^{2+}	y^{3+}	y^{4+}	#2

1	157.10840	79.05784	53.04098	40.03256	R					41
2	228.14552	114.57640	76.72002	57.79184	A	4598.26302	2299.63515	1533.42586	1150.32121	40
3	325.19829	163.10278	109.07095	82.05503	P	4527.22590	2264.11659	1509.74682	1132.56193	39
4	472.26671	236.63699	158.09375	118.82213	F	4430.17313	2215.59020	1477.39589	1108.29874	38
5	573.31439	287.16083	191.77631	144.08405	T	4283.10471	2142.05599	1428.37309	1071.53163	37
6	687.35732	344.18230	229.79062	172.59479	N	4182.05703	2091.53215	1394.69053	1046.26971	36
7	834.42574	417.71651	278.81343	209.36189	F	4068.01410	2034.51069	1356.67622	1017.75898	35
8	949.45269	475.22998	317.15575	238.11863	D	3920.94568	1960.97648	1307.65341	980.99188	34
9	1046.50546	523.75637	349.50667	262.38182	P	3805.91873	1903.46300	1269.31109	952.23514	33
10	1133.53749	567.27238	378.51735	284.13983	S	3708.86596	1854.93662	1236.96017	927.97195	32
11	1234.58517	617.79622	412.19991	309.40175	T	3621.83393	1811.42060	1207.94949	906.21394	31
12	1347.66924	674.33826	449.89460	337.67277	L	3520.78625	1760.89676	1174.26693	880.95202	30
13	1460.75331	730.88029	487.58929	365.94378	L	3407.70218	1704.35473	1136.57224	852.68100	29
14	1557.80608	779.40668	519.94021	390.20698	P	3294.61811	1647.81269	1098.87755	824.40998	28
15	1644.83811	822.92269	548.95089	411.96498	S	3197.56534	1599.28631	1066.52663	800.14679	27
16	1731.87014	866.43871	577.96156	433.72299	S	3110.53331	1555.77029	1037.51595	778.38878	26
17	1844.95421	922.98074	615.65625	461.99401	L	3023.50128	1512.25428	1008.50528	756.63078	25
18	1959.98116	980.49422	653.99857	490.75075	D	2910.41721	1455.71224	970.81059	728.35976	24
19	2107.04958	1054.02843	703.02138	527.51785	F	2795.39026	1398.19877	932.46827	699.60302	23
20	2293.12890	1147.06809	765.04782	574.03768	W	2648.32184	1324.66456	883.44546	662.83592	22
21	2394.17658	1197.59193	798.73038	599.29960	T	2462.24252	1231.62490	821.41902	616.31609	21
22	2557.23990	1279.12359	853.08482	640.06543	Y	2361.19484	1181.10106	787.73646	591.05417	20
23	2654.29267	1327.64997	885.43574	664.32862	P	2198.13152	1099.56940	733.38202	550.28834	19
24	2711.31414	1356.16071	904.44290	678.58399	G	2101.07875	1051.04301	701.03110	526.02514	18
25	2798.34617	1399.67672	933.45357	700.34200	S	2044.05728	1022.53228	682.02394	511.76978	17
26	2911.43024	1456.21876	971.14826	728.61302	L	1957.02525	979.01626	653.01327	490.01177	16
27	3012.47792	1506.74260	1004.83082	753.87494	T	1843.94118	922.47423	615.31858	461.74075	15
28	3149.53683	1575.27205	1050.51713	788.13966	H	1742.89350	871.95039	581.63602	436.47883	14
29	3246.58960	1623.79844	1082.86805	812.40286	P	1605.83459	803.42093	535.94971	402.21410	13
30	3343.64237	1672.32482	1115.21897	836.66605	P	1508.78182	754.89455	503.59879	377.95091	12
31	3456.72644	1728.86686	1152.91366	864.93707	L	1411.72905	706.36816	471.24787	353.68772	11
32	3619.78976	1810.39852	1207.26810	905.70290	Y	1298.64498	649.82613	433.55318	325.41670	10
33	3748.83236	1874.91982	1250.28230	937.96355	E	1135.58166	568.29447	379.19874	284.65087	9
34	3835.86439	1918.43583	1279.29298	959.72155	S	1006.53906	503.77317	336.18454	252.39022	8
35	3934.93281	1967.97004	1312.31579	984.48866	V	919.50703	460.25715	307.17386	230.63221	7
36	4035.98049	2018.49388	1345.99835	1009.75058	T	820.43861	410.72294	274.15105	205.86511	6
37	4222.05981	2111.53354	1408.02479	1056.27041	W	719.39093	360.19910	240.46849	180.60319	5
38	4335.14388	2168.07558	1445.71948	1084.54143	I	533.31161	267.15944	178.44205	134.08336	4
39	4448.22795	2224.61761	1483.41417	1112.81244	I	420.22754	210.61741	140.74736	105.81234	3
40	4608.25860	2304.63294	1536.75772	1152.82011	C-Carbamidomethyl	307.14347	154.07537	103.05267	77.54132	2
41					K	147.11281	74.06004	49.70912	37.53366	1

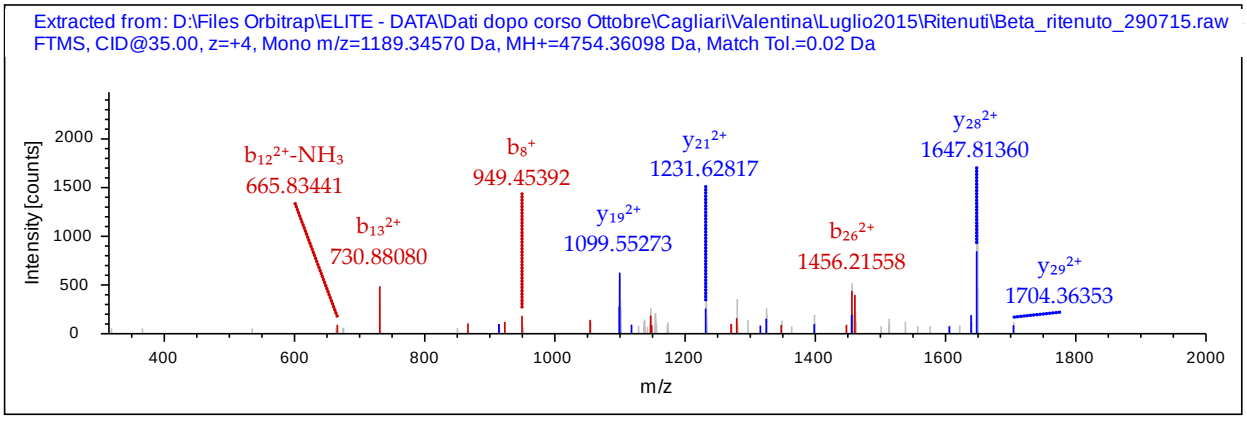


Fig. S5h. Carbonic anhydrase 1, Fragment 215-228.
$[M+H]^+$ 1580.79534 Da (Theor. 1580.7914) RT: 23.41 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 790.90131 Da (+1.9 mmu/+2.4 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:4.68, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Carbonic anhydrase 1 OS=Homo sapiens GN=CA1 PE=1 SV=2 - [CAH1_HUMAN]

b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
130.04988	65.52858	E			14
217.08191	109.04459	S	1451.74894	726.37811	13
330.16598	165.58663	I	1364.71691	682.86209	12
417.19801	209.10264	S	1251.63284	626.32006	11
516.26643	258.63685	V	1164.60081	582.80404	10
603.29846	302.15287	S	1065.53239	533.26983	9
690.33049	345.66888	S	978.50036	489.75382	8
819.37309	410.19018	E	891.46833	446.23780	7
947.43167	474.21947	Q	762.42573	381.71650	6
1060.51574	530.76151	L	634.36715	317.68721	5
1131.55286	566.28007	A	521.28308	261.14518	4
1259.61144	630.30936	Q	450.24596	225.62662	3
1406.67986	703.84357	F	322.18738	161.59733	2
		R	175.11896	88.06312	1

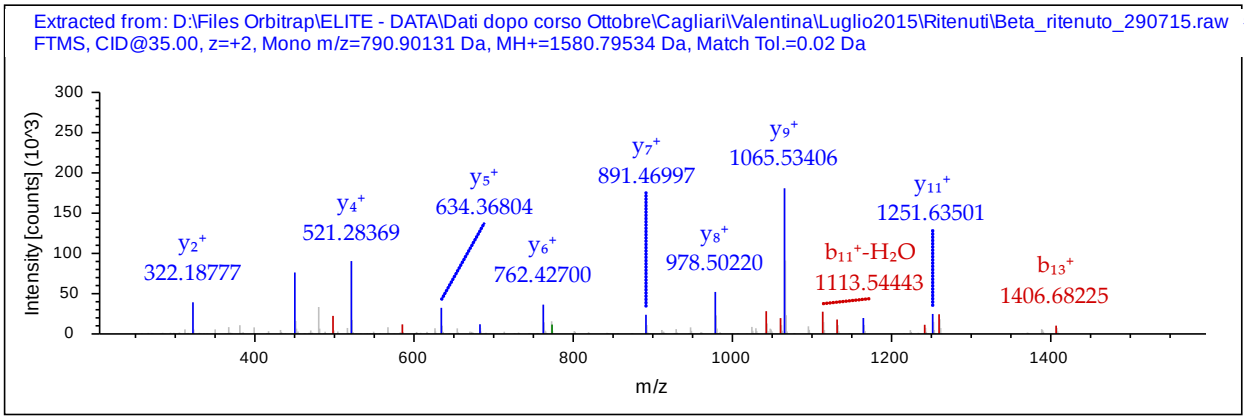


Fig. S5i. Carbonic anhydrase 1, Fragment 229-253.
$[M+H]^+$ 2759.39060 Da (Theor. 2759.3893) RT: 22.31 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+3H]^{3+}$ monoisotopic m/z: 920.46838 Da (+0.38 mmu/+0.41 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:4.86, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.8 Da
Fragments used for search: b; b-H_2O; b-NH_3; y; y-H_2O; y-NH_3
Protein references (1):
- Carbonic anhydrase 1 OS=Homo sapiens GN=CA1 PE=1 SV=2 - [CAH1_HUMAN]

#1	b^+	b^{2+}	b^{3+}	Seq.	y^+	y^{2+}	y^{3+}	#2
1	88.03931	44.52329	30.01795	S				25
2	201.12338	101.06533	67.71264	L	2672.35743	1336.68235	891.45733	24
3	314.20745	157.60736	105.40733	L	2559.27336	1280.14032	853.76264	23
4	401.23948	201.12338	134.41801	S	2446.18929	1223.59828	816.06795	22
5	515.28241	258.14484	172.43232	N	2359.15726	1180.08227	787.05727	21
6	614.35083	307.67905	205.45513	V	2245.11433	1123.06080	749.04296	20
7	743.39343	372.20035	248.46933	E	2146.04591	1073.52659	716.02015	19
8	800.41490	400.71109	267.47648	G	2017.00331	1009.00529	673.00595	18

9	915.44185	458.22456	305.81880	D	1959.98184	980.49456	653.99880	17
10	1029.48478	515.24603	343.83311	N	1844.95489	922.98108	615.65648	16
11	1100.52190	550.76459	367.51215	A	1730.91196	865.95962	577.64217	15
12	1199.59032	600.29880	400.53496	V	1659.87484	830.44106	553.96313	14
13	1296.64309	648.82518	432.88588	P	1560.80642	780.90685	520.94032	13
14	1427.68359	714.34543	476.56605	M	1463.75365	732.38046	488.58940	12
15	1555.74217	778.37472	519.25224	Q	1332.71315	666.86021	444.90923	11
16	1692.80108	846.90418	564.93854	H	1204.65457	602.83092	402.22304	10
17	1806.84401	903.92564	602.95285	N	1067.59566	534.30147	356.53674	9
18	1920.88694	960.94711	640.96716	N	953.55273	477.28000	318.52243	8
19	2076.98806	1038.99767	693.00087	R	839.50980	420.25854	280.50812	7
20	2174.04083	1087.52405	725.35179	P	683.40868	342.20798	228.47441	6
21	2275.08851	1138.04789	759.03435	T	586.35591	293.68159	196.12349	5
22	2403.14709	1202.07718	801.72055	Q	485.30823	243.15775	162.44093	4
23	2500.19986	1250.60357	834.07147	P	357.24965	179.12846	119.75473	3
24	2613.28393	1307.14560	871.76616	L	260.19688	130.60208	87.40381	2
25				K	147.11281	74.06004	49.70912	1

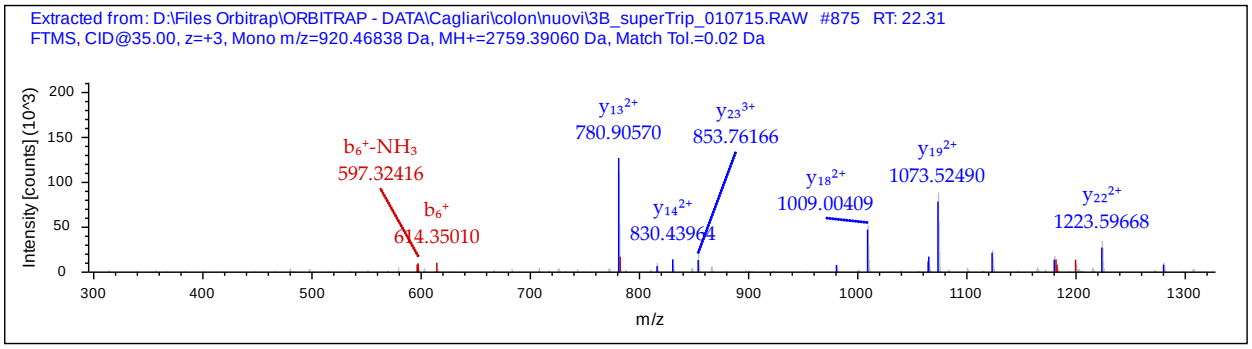
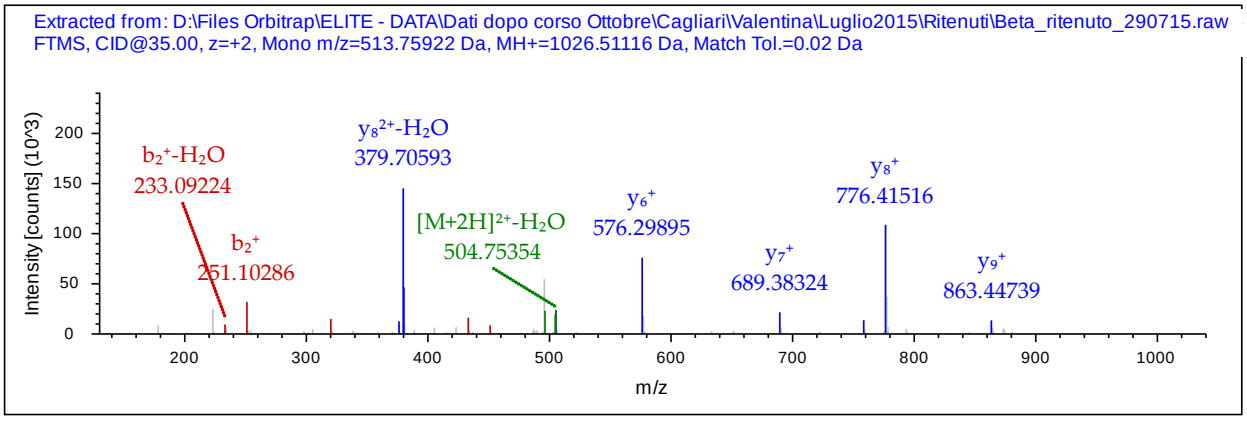


Fig. S5l. Carbonic anhydrase 1, Fragment 229-238.
$[M+H]^+$ 1026.51116 Da (Theor. 1026.5102) RT: 16.49 min. Annotated MH^+ spectrum of high-resolution MS/MS of the ion $[M+2H]^{2+}$ monoisotopic m/z: 513.75922 Da (+0.45 mmu/+0.88 ppm). Analysis performed by Proteome discoverer software.
Identified with: Sequest HT (v1.3); XCorr:3.01, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.5 Da
Fragments used for search: b; b-H_2O; y; y-H_2O; y-NH_3
Protein references (1):
- Carbonic anhydrase 1 OS=Homo sapiens GN=CA1 PE=1 SV=2 - [CAH1_HUMAN]

#1	b^+	b^{2+}	Seq.	y^+	y^{2+}	#2
1	164.07060	82.53894	Y			10
2	251.10263	126.05495	S	863.44693	432.22710	9
3	338.13466	169.57097	S	776.41490	388.71109	8
4	451.21873	226.11300	L	689.38287	345.19507	7
5	522.25585	261.63156	A	576.29880	288.65304	6
6	651.29845	326.15286	E	505.26168	253.13448	5
7	722.33557	361.67142	A	376.21908	188.61318	4
8	793.37269	397.18998	A	305.18196	153.09462	3
9	880.40472	440.70600	S	234.14484	117.57606	2
10			K	147.11281	74.06004	1


Extracted from: D:\Files Orbitrap\ELITE - DATA\Dati dopo corso Ottobre\Cagliari\Valentina\Luglio2015\Ritenuti\Beta_ritenuto_290715.raw
FTMS, CID@35.00, z=+2, Mono m/z=513.75922 Da, MH+=1026.51116 Da, Match Tol.=0.02 Da
Intensity [counts] (10^3)
m/z
b2+-H2O
233.09224
b2+
251.10286
y8 2+-H2O
379.70593
[M+2H]2+-H2O
504.75354
y6+
576.29895
y7+
689.38324
y8+
776.41516
y9+
863.44739