Università degli Studi di Cagliari

PhD DEGREE

In

Life, Environmental and Drug Sciences

Biomedical Curriculum

XXXII Cycle

Potential Salivary Biomarkers In Mastocytosis: A Proteomics Approach

BIO/10

Presented by: Doctor Simone Serrao

PhD coordinator: Professor Simona Distinto

Tutor: Professor Tiziana Cabras

Final exam academic year 2018-2019

Thesis defense: January – February 2020

Contents

Abstract

Mastocytosis is a myeloproliferative neoplasm characterized by infiltration of clonally derived mast cells in different tissues. According to mast cells localization, it is possible to discriminate cutaneous mastocytosis (CM) from systemic mastocytosis (SM), the latter involving at least an extracutaneous organ, like bone marrow, liver, spleen and gastrointestinal tract. Some of the SM patient can develop also cutaneous lesions (SM+C). Oral cavity is commonly involved in the symptomatology. Disease classification is often tricky. In the first part of this thesis, in order to highlight possible qualitative/quantitative modifications of the salivary proteome associated to the different forms of the disease, we investigated salivary samples collected from 6 CM, 35 SM patients, among which 8 with only systemic symptoms (SM-C) and 27 with both systemic and cutaneous symptoms (SM+C), and 48 healthy controls by a top-down proteomic approach. Low-resolution HPLC-ESI-MS analysis of the acid soluble fractions of saliva highlighted different proteomic profiles in the three patients' groups, showing that the salivary samples of the patients were characterized by a down-regulation of peptides and proteins involved in the homeostasis and defense of the oral cavity, and in the innate immunity and in inflammation not only in the oral cavity but at systemic level, such as aPRPs, statherin, histatins and cystatins. Only two proteins with regulatory roles in the innate immunity and inflammation, S100A8 and antileukoproteinase, resulted up-regulated in patients differently to all the other salivary proteins analyzed, suggesting the establishment of a response by the organism to the injuries caused by the disease. Interestingly, some differences have been found among the patients in the concentration of α-defensins 1, thymosin β-4, and the truncated forms of cystatin D-R26 variant, and some truncated forms of P-B and statherin. Correlation between the protein/peptide levels and tryptase concentration evidenced that acidic PRPs, statherin and P-B fragments, and cystatin D-R26 des1-5 correlated positively just in SM-C group, while thymosin β-4 correlated negatively. Since the interesting data on cystatin D, in the second part of the thesis I focused on the characterization of the salivary protein complex aggregating to the cystatin D-C26 variant (named by us SIC-D). Indeed, the C-26 variant is usually undetectable in acid soluble fraction of saliva but measurable in whole saliva. Pools of whole saliva from 4 CM, 3 SM-C, 14 SM+C, and 20 sex/age matched healthy controls, were submitted to immunoprecipitation with cystatin D-C26 antibody followed by SDS-PAGE/western-blot under reducing and non-

reducing conditions. Since the low volume of CM samples, the tryptic digestion, and the nano-HPLC-high-resolution-MS/MS analysis were performed only in SM-C, SM+C and control samples. The quantitative comparison was performed with Proteome Discoverer 2.2 software. SIC-D included 44 proteins, among which IgA, IgG, PIgR, annexins, α-defensin 1/2, S100A8, carbonic anhydrase 6, prolactin-inducible protein, lysozyme C and dermicidin. Several qualitative/quantitative differences were highlighted with respect to controls and between the two patient groups. The most relevant were: all the patients exhibited lower levels of IgA, PIgR, DMBT-1 and S100A8 than controls, but higher levels of IgG, α-defensins 1/2 and carbonic anhydrase 6. The highest level of cystatin D-C26 was found in SM+C patients, which were different from SM-C for annexin A2. Both SM-C and SM+C showed the presence of antileukoproteinase and S100A14.

The results on the acid-soluble fraction of saliva and the preliminary results on the SIC-D complex are promising in order to find candidate markers able to discriminate the different forms of mastocytosis.

Introduction

Mast cells (MCs) origin from are hematopoietic stem cells that mature in vascularized tissues and show high estimated life. Morphologically, they are characterized by numerous electron dense cytoplasmatic granules containing enzymes and cytokines; on the surface they can express a wide range of receptors such as the high-affinity IgE receptor (FCεRI), the gamma Fc receptor (Fcγ), complement, cytokine, hormone receptors and Toll-like receptors. This characteristic permits MCs to secrete a large range of biologically active products that promote or inhibit the immune responses[1]. They are specialized secretory cells of the innate immune system and play an important role in host defense producing and releasing proinflammatory mediators, chemotactic factors and immunoregulatory cytokines[2]. After the bound of IgE on Fc receptors, they immediately release a large number of secretory granules that contain histamine, serotonin and other inflammatory mediators, which are important in both immediate and late-phase of inflammatory reaction[3]. MCs can produce three types of effector mediators: pre-formed mediators stored in granules like serotonin, histamine, tryptase and chymase; de-novo synthetized mediators after stimulation, such as the lipid mediators PAF and PDG2; cytokines including IL-1, IL-3, IL-5, IL-8, IL-18, TNF-α, TGF-β, VEGF[4]. MCs develop from $CD34^+$ pluripotent progenitor cells[5,6]; bone marrow-derived MC progenitors enter in to the bloodstream and then migrate to the peripheral tissue where they acquire their mature phenotype and become terminally differentiated. Mature MCs tend to stay in tissue near blood vessels, nerves and glandular structure, therefore MCs are particularly abundant in skin, gastrointestinal tract, respiratory tract and lymphoid tissue[7]. MCs growth and survival depend on the presence of stem cell factor (SCF), which bind to c-KIT, a specific transmembrane tyrosine kinase receptor on the surfaces of MCs. The interaction between ligand and receptor promotes as well as survival of MCs, proliferation, adhesion and chemotaxis[8].

SCF and c-KIT receptor

SCF is a grow factor expressed by fibroblast and endothelial cells, promoting proliferation, migration, survival and differentiation of hematopoietic progenitors, melanocytes and germ cells. It exists as two forms, bounded to the membrane or in a soluble form. The first includes an extracellular domain, a transmembrane domain and an intracellular region. The soluble

form is obtained as a result of a proteolytic cleavage in the extracellular domain of 165 amino acids[9]. Several proteases have been hypnotized as responsible of this proteolytic cleavage, such as matrix metalloprotease-9[10], chymase-1[11] and several members of ADAMs (a disintegrin and metalloproteinases) family like ADAM17 and ADAM33[12]. Soluble SCF exists in homo-dimeric conformation, with two dimers that interact head-head to form an elongated dimer. This dimerization is a dynamic process, which could have a regulatory role in dimerization and activation of c-KIT receptor[13].

The c-KIT receptor belongs to the type III receptor tyrosine kinase family characterized by the presence of five immunoglobulin-like domains and the presence of a kinase insert of 70-100 amino acids. The c-KIT receptor is encoded by *c-KIT* gene and it is a transmembrane protein with an extracellular domain constituted by five immunoglobulin-like domains followed by a single transmembrane region. The first part of intracellular domain, the juxta-membrane region, has great importance for the regulation of c-KIT kinase activity because it forms a hairpin loop that inserts into the active site and blocks itself the kinase activity. The kinase domain consists of two subdomains, tyrosine kinase 1 and 2, separated by a kinase insert sequence. The C-terminal region ends the proteins. Most of the phosphorylation sites are located in the juxta-membrane region, in the kinase insert and in the C-terminal region[8]. The binding of SCF homodimer to the first three immunoglobulin-like domains leads to a conformational change that allows the receptor dimerization and activation thanks to immunoglobulin-like domain 4 and 5 interaction. These protein-protein interactions between transmembrane regions and kinase domains simplify their activation and phosphorylation. The phosphorylation cascade starts in the juxta-membrane region, removing the auto-inhibitory loop[9]. Then, the phosphorylation continues in an orderly manner. The phospho-tyrosine residues act as docking sites for signaling molecules SRC and SHC kinase, phosphoinositide 3-kinase (PIP3K) and phospholipase Cγ, which activate MAP kinase pathway that lead to intracellular calcium concentration and activation of transcriptional factors for MCs proliferation[14].

Various diseases are associated with an increased number of MCs, their activation, or both. Basing on the clinical findings and on the histological features and clonality of MCs it is possible to recognized disorders caused by expansion of clonal MCS and disorders caused by MCs activation[15].

Mastocytosis

Mastocytosis is defined as an abnormal clonal MCs expansion and accumulation in various tissues such as the bone marrow and the skin. The prevalence of the disease in unknown, but recent population-based studies estimate 1 case per 10000 people[16,17]. It comprises a heterogeneous group of disorders that the World Health Organization (WHO) classifies into subvariants, summarized in Table 1: cutaneous mastocytosis (CM), where there is not systemic involvement, systemic variants (SM), and localized MC tumors[18–20]. CM is usually diagnosed in childhood and has a good prognosis, in many case the skin lesion disappear spontaneously during puberty[21]. WHO divides CM into macupapular CM, also termed urticaria pigmentosa (UP), diffuse cutaneous mastocytosis (DCM), and localized mastocytoma of skin. Differently, SM usually develops in adults, it is characterized by the infiltration of MCs in various internal organs, such as bone marrow (BM) and gastrointestinal tract (GI)[18] and usually occurs as a chronic and indolent disease, also called "indolent systemic mastocytosis" (ISM). In some patients is possible to diagnose more advanced type of SM, such as aggressive SM (ASM), SM with associated hematologic neoplasm (SM-AHN), and MC leukemia (MCL) [22,23].

Table 1: World Health Organization Classification of Mastocytosis (2016)[20].

Categories	Subtypes	Diagnostic Criteria	Prognosis
CM	• Macupapular CM = UP • DCM • Mastocytoma of skin	No systemic involvement (most patients are children)	+/- Good
SM	ISM	• No high MC burden and end-organ damages • Most patients are adults	Good
	SSM	• High MC burden, no end-organ damages	Good
	SM-AHN	• SM criteria and WHO diagnostic criteria for AHN are fulfilled • Frequently associated to myeloid AHNs (MPN, MDS, MPN/MDS), rarely to lymphoid AHNs	Depends on the type of SM and of AHN
	ASM	• At least 1 end organ damage	Poor
	MCL	• BM smear > 20% of MCS • PB smear >10% MCs	Very poor
MCS		• Rare form of high-grade solid MC tumor • Very atypical MCs	Very poor

Epidemiology

Mastocytosis in considered has an orphan disease, affecting less than 200000 people in USA. Because of its rareness and because of the criteria that has been established in 2001, there are not many epidemiologic studies to define the precise incidence, point prevalence or cumulative prevalence of mastocytosis in general population[17]. Mastocytosis can occur in children and adult and even if it can occur at any age, in the 50% of the cases it occurs in the first 2 years of life[24,25]. Mastocytosis at birth is uncommon[26]. In adults, the diagnosis is usually between 20 and 50 years of age. Since the chronic nature of the disease and the quite low regression rate in adults, the presence of the disease in adult patients rises with age. There are no gender differences in both incidence and age of onset[25].

Causes and pathogenesis

The interaction between SCF and c-KIT receptor seems to play and essential role in the development and onset of mastocytosis[27]. In fact, patients affected by mastocytosis show activating mutation in the gene *c-KIT*, which encodes for c-KIT receptor, in bone marrow, skin and peripheral blood. The most common oncogenic mutations in c-KIT are in the juxta-membrane region, Val560Gly (V560G) or in the kinase domain, Asp816Val (D816V) (Figure 1). The mechanism of the constitutive activation of c-KIT is not fully understood. Recent *in silico* studies on the D816V mutated protein show that the mutation in the kinase domain results in a structural change of the auto-inhibitory loop and in a weaker bond between juxta-membrane and kinase domain[28]. Consequently, the juxta-membrane region, that has an inhibitory activity on the kinase domain, cannot suppress the enzymatic activity of c-KIT. Another study, performed on the V560G mutated receptor, highlights that the receptor can dimerize without the ligand bond[29]. In both hypotheses, the signals of survival and proliferation on MCs are still on and it is possible to link it to mastocytosis onset.

After the activation and degranulation, MCs generate and secrete a wide range of mediators that are responsible of the typical symptoms of the disease. Histamine acts through 4 different receptors, H1 to H4, and leads to vasopermeability, vasodilatation, GI and bronchial muscle contraction, gastric acid production and pruritus. MCs also secrete proteases, among which the most important is the tryptase, which is constitutively secreted by them. Mastocytosis'

patients usually have augmented levels of serum tryptase and histamine[30,31], as well interleukin 3 and 16, and tumor necrosis factor-α[32].

Clinical features

Cutaneous hallmarks

Skin is usually the most common organ involved in mastocytosis, it is often the first sign of the disease and in children it could be the only manifestation of the disease. The most frequent skin manifestation both in children and adults is urticaria pigmentosa (UP), characterized by yellow/brown macules mainly on the trunk and legs[33]. Symptoms include pruritus, flushing and blistering. UP lesions are detectable in more than 90% of ISM patients and less than 50% of SM-AHNMD or ASM. Regression of the lesion might not indicate disease recovery[19,33,34].

GI Symptoms

GI symptoms are quite common in patients affected by mastocytosis. The typical symptoms include diarrhea and bloating, followed by nausea and abdominal pain[35–37]. Some patients are also are also usual to develop duodenal ulcers because of the increase of gastric acid secretion caused by higher histamine production and secretion from MCs[35,36]. The frequency of GI symptoms do not seem correlated to the presence of D816V mutation nor to the levels of tryptase nor to the age of onset of mastocytosis[36,38].

Bone marrow pathology

SM is usually diagnosed following a bone marrow histopathologic analysis, as proposed by WHO (Table 1). ISM, the most common variant of SM is diagnosed when criteria for mastocytosis are underlined and there are no evidences of an associated clonal hematologic disorder nor sever liver disease and lymphadenopathy[20].

Hepatic and splenic findings

Liver, spleen and lymph nodes are damaged in patients affected by SM, especially in the most aggressive forms[39]. Usually, MCs infiltrates could be found in the paratrabecular compartment

of the spleen and in the paracortex of lymph nodes. Liver involvement is as well common, but severe liver disease is present in patients with aggressive forms[40].

Diagnosis

Diagnosis of mastocytosis is based on clinical manifestation, histopathology analysis and laboratory evaluation, and then is classified as a variant according to WHO criteria (Tables 1 and 2)[41]. The diagnosis of CM is based on a general macroscopic visual analysis followed by a lesioned skin biopsy to confirm the characteristic histopathology pattern. CM must be distinguished from diseases which share similar characteristics; UP lesions have increased MCs number in the dermal papillae below macules and papules, especially near blood vessels in the upper dermis[42]. MCs may also appear as nodular infiltrates within the papillary dermis and subcutaneous tissues and may also be found in increased numbers in the normal-appearing skin between lesions of UP[42]. The differences in the histologic pattern in CM are related to the density of MCs infiltrates.

The actual procedure to diagnose SM consists of a medical examination, which includes a bone marrow biopsy, a serum tryptase level and an analysis to look for an activating mutation in *c-KIT*, preferably on bone marrow mononuclear cells. The diagnosis of SM requires fulfilling of the major criterion, consisting of multifocal dense MCs aggregates, and 1 minor criterion, or 3 minor criteria[20] (see Table 2). In pediatric cases, bone marrow biopsy is not supported warmly unless there is an evidence of systemic disease, confirmed by hepatosplenomegaly, lymphadenopathy or unexplained peripheral blood abnormalities. In addition, it is recommended an immunohistochemistry analysis to identify $CD25^+$ MCs, because it has been demonstrated that CD25 is expressed on most MCs in mastocytosis[43]. The analysis of other tissues, like spleen, liver or GI mucosa, could help to determine the expansion of MCs, but are not typically necessary. In case of involvement, the histopathologic profile of biopsy from these tissues is similar to that in the bone marrow and the MCs infiltrates are often $CD25^{+}$[35,36,44].

Despite the efforts from the WHO in providing more specific and accurate criteria, in clinical practical distinguish the different form of mastocytosis disease is not so easy. Sometimes, it is possible to find patients that show symptoms of both CM and SM forms (SM+C), with MCs accumulation in the skin as well as other extracutaneous organs. The absent of a strong method of diagnosis and classification and the overlap of some clinical symptoms in the

different forms of mastocytosis make urgent the discovery and characterization of potential disease-specific biomarkers. A valid and efficient tool for this challenge is constituted by proteomic-focused approach, even though a few proteomic studies have been performed on mastocytosis disease[45].

Therapy

Symptomatic therapy

In most patients affected by CM or SM, symptomatic therapy is required independently from the subvariant to control mediator secretion or mediator effects. The therapy that results to be efficient in most of the symptomatic patients is a combination of histamine receptor (HR1 and HR2) antagonists[46]. However, in a few patients, symptoms remain even at maximum doses of HR blockers. In fact, some of these patients suffer from coexisting allergic disorders and, for them, it is important to know the spectrum of triggers/allergens to avoiding the allergen contact. Sometimes, additional glucocorticosteroids[47] could help suppressing the adverse reactions[32,48]. In those patients with detectable IgE, immunotherapy is usually recommended[49,50]. It is important to know of the full spectrum of mediator-related symptoms in SM, the organ systems that can be involved, potential drug side effects and the exact diagnosis and subvariant because in many patients SM-related symptoms are confused with unrelated features and symptoms, highlighting coexisting disorders[51].

Targeted therapies

Several targeted drugs have been developed to treat patients affected by advanced MC disorders. Many of these are oriented against the mutated form of KIT D816V[52–54]. Drugs like midostaurin, desatinib[55,56] and nilotinib[57,58] are reported to inhibit the growth of MC lines and primary neoplastic MCs with c-KIT D816V mutation *in vitro*[56]. However, these effects are usually moderate and ephemeral and are followed by a fallback[59,60]. Presently, clinical trials are focused on the scan of the real antineoplastic potential of these drugs in ASM and MCL. An issue of that, in ASM and MCL, c-KIT-independent pathways could play an important role in disease progression[61], so it was not surprising that some treated patients who responded at first, relapsed early[60]. Therefore, recent studies are focusing on drug combination directed against KITD816V and other relevant KIT-independent kinases, like Lyn and Btk.

Prognosis

Generally, children have more favorable prognosis[19,62–64]. Children with CM usually recover without any treatment in adolescence[19,51,65]. In adults, the healing of the cutaneous wounds could be correlated with the reduction of clinical manifestation but may not indicate the disease regression. In fact, in many case even if the skin lesions decrease, the bone marrow continues to show MCs involvement and the diagnosis of SM is maintained due to WHO criteria[34,51].

Figure 1: Description of c-KIT receptor[66].

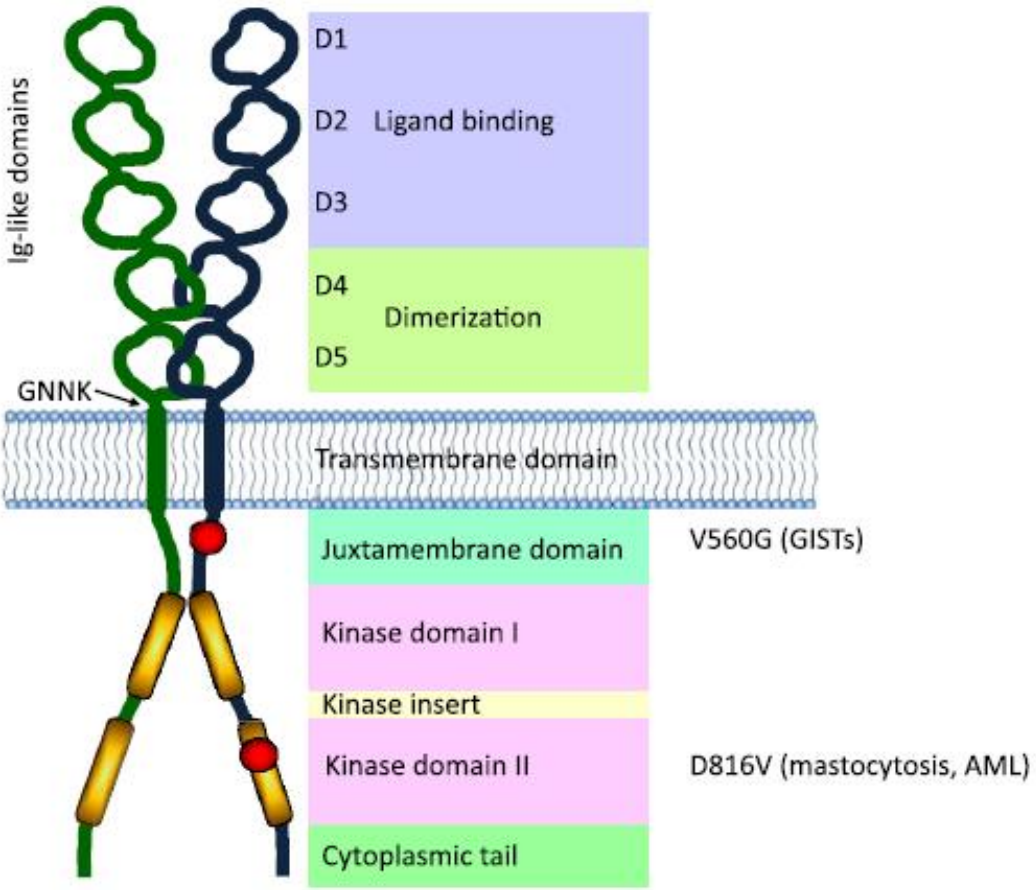


Table 2: Diagnostic criteria for cutaneous and systemic mastocytosis[20].

CM	
Urticaria pigmentosa, maculopapular CM, DCM, or solitary mastocytoma and typical MCs infiltrates in a multifocal or diffuse pattern on skin biopsy	
SM	
Major and 1 minor, or 3 minors, criteria are required for the diagnosis of SM	
Major criterion	Multifocal infiltrates of MCs detected in bone marrow or other extracutaneous organs, and confirmed by tryptase immunohistochemistry
Minor criteria	More than 25% of MCs in the infiltrate have atypical morphology or more than 25% of MCs in bone marrow are immature or atypical
	Detection of *c-KIT* point mutation at codon 816 in bone marrow, blood or other extracutaneous organs
	MCs in bone marrow, blood or other extracutaneous organs co-express CD117 with CD2 or CD25
	Serum tryptase concentration higher than 20ng/ml

SM criteria were defined by the WHO in 2001 and have been confirmed in the WHO updates of 2008 and 2016

Proteomics in detection and characterization of potential biomarkers

The term proteomics indicates the qualitative and quantitative study of the total functionally protein pattern expressed by a genome in a certain period of a cell, a tissue or an organism. A proteomic study includes the characterization of all the isoforms and post-translational modifications (PTMs), such as phosphorylation, glycosylation, proteolytic cleavages, and others, of a certain protein. Since the expression of a protein and/or PTMs can change depending on the cell type and answering to different endogenous and exogenous signals, to perform a proteomic study could be helpful to understand how and when these changes occurs, and how the proteins play different roles in different pathways during the life cycle of a cell, a tissue or an organism. Thus, a proteomic study may be considered as a valid tool to investigate different pathological and physiological states and highlight potential biomarkers. In fact, a proteomic study can analyze at the same time hundreds or thousands of proteins and peptides that are part of a biological system (tissue, cell culture, serum, plasma, urine, saliva, cerebrospinal fluid, synovial liquid). Moreover, it allows the structural characterization of proteins, their quantitation and the qualitative and quantitative comparison of the same protein pattern in different conditions (physiological, environmental, pharmacological and pathological). One of the most important application of proteomics is the identification of biomarkers useful for prognostic and diagnostic aim, to monitor the response of a patient to therapeutic treatments, and as indicators of disease state. Proteomics-based approaches could be helpful in different clinical aspects, such as discovery of modified pathways in a specific disease, identification of individuals inclined to manifest a disease, to respond to specific therapeutic program or develop side effects[67–69](Table 3).

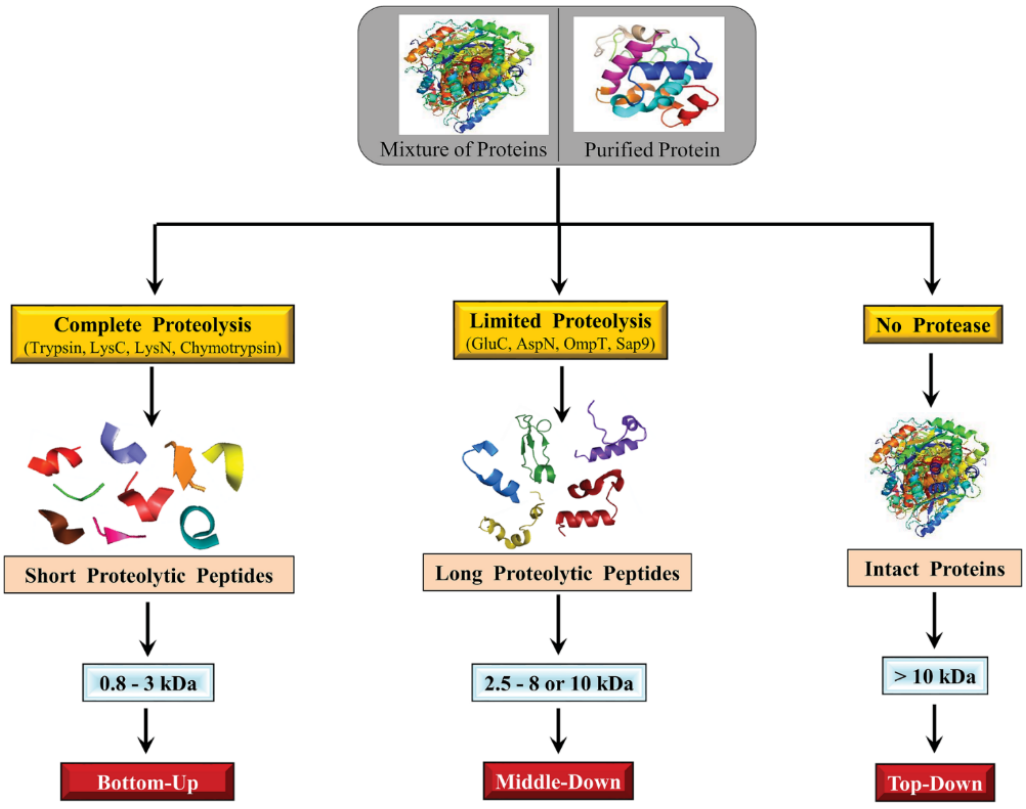
Proteomics studies take advantage of high yield, sensitivity, accuracy and selectivity of technologically advanced methodologies, such as mass spectrometry, bioinformatic tools and proteomics and genetics databases. Nowadays, are present different proteomics platforms, each of which present advantages and disadvantages, distinct in top-down, bottom-up and middle-down[70,71](Figure 2). Top-down approaches are based on the analysis of the intact sequence of the proteins, avoiding sample degradation. These methods are useful because they will give more information about PTMs, which are important in the comprehension of protein biological functions. On the other hand, top-down analysis does not allow the characterization of high molecular weight or glycosylated proteins. In fact, high molecular

weight proteins do not provide a complete MS/MS fragmentation, and the obtained MS/MS spectra are too complex to be read by bioinformatic software[72]. In bottom-up platforms, the samples are subjected to a complete proteolytic digestion by using specific enzymes such us trypsin before the analysis of the peptide fragments by high-throughput analytical methods. Then, the presence of a protein in the sample is assumed when one or more of its specific fragments are detected. This procedure gives rise to numerous peptides, depending on the complexity of the sample under study. Thus, many tryptic peptides are quite often not detected, leading to an inadequate sequence coverage and to an important loss of information, such as PTMs and proteoforms presence[71]. Nerveless, this approach is useful to sequence or characterize purified proteins, since not many peptides would be generated upon complete trypsin digestion of a protein compare to the complete digestion of a mixture of several proteins. Middle-down approach combines the feature of bottom-up and top-down approaches with the aim to arrive at optimum conditions. Middle-down techniques study truncated peptides obtained by limited proteolysis or restricted digestion, thanks to the optimization of the incubation time of proteolysis or manipulating the relative concentration of enzyme and substrate[71]. Thus, the size of resulting peptides would be greater than the ones that are usually originated by bottom-up approaches. Moreover, the number of proteolytic peptides in samples would be relatively lesser than in bottom-up, resulting in a less samples complexity and therefore in an enhanced probability of detecting unique peptides[71].

Table 3: Overview of different proteomic studies.

Proteomics	
Interaction proteomics	Identification of protein interactions, protein complexes and interactomes through affinity purification followed by protein mass spectrometry, using tagged protein baits.
	It is useful for biological networks and system biology (cell signaling cascades and gene regulatory networks
Structural proteomics	Comparison of protein structures to identify functions with X-ray crystallography and NMR spectroscopy
	It gives information about protein/protein and protein/drugs interactions
Expression proteomics	Analysis of proteins expression by using 2D-PAGE and mass spectrometry
	It permits the comparison between diseased and healthy tissue to find drug target or diagnostic marker.
Proteogenomics	Improvement of gene annotation through mass spectrometry in order to discover PTMs and proteolytic events.
Biomarkers	Identification of candidate biomarkers, bacterial targeted by the immune response and immunohistochemistry markers of infectious or neoplastic diseases. It uses techniques as western blot, immunohistochemical staining, ELISA and mass spectrometry

Figure 2: Top-down, middle-down and bottom-up approaches[71].



Saliva as a diagnostic fluid functional in proteomics studies

Saliva is a complex body fluid originated mainly from the secretions from the salivary glands, gingival crevicular fluid (GCF) and plasma exudate. Approximatively 90% of saliva is produced by major salivary glands: parotid, submandibular and sublingual glands. Around 10% is produced by minor salivary glands which are distributed in the oral mucosa (lingual, labial, buccal, palatine, gloss palatine)[73]. Salivary flow differs in the stimulated and unstimulated state. In the stimulated state, parotid glands contribute more than 50% of the totally salivary flow; in unstimulated state, submandibular glands contribute for 65% of the flow, parotids gland for 20% and sublingual for 7-8%[74]. The saliva composition is wide, and it includes enzymes, such as amylase, lysozyme, lipase and lactoperoxidase, glycoproteins, lipids (hormones such as testosterone and progesterone), proteases and high concentration of inorganic ions like sodium, chloride, calcium, potassium, bicarbonate, magnesium and phosphate[75]. In saliva, has been identified more than 2000 proteins and peptides[76], 90% of which derives from the secretion of the three major glands and belongs to the class of proline-rich proteins (PRPs), divided in acid, basic and basic glycosylated proteins, α-amylase, mucins, salivary cystatins, cystatin C and D, histatins, statherin and P-B peptide[77] (Figure 3). The remaining 10% of proteins are secreted by minor salivary glands, or derive by exfoliated cells, by leucocytes present in the gingival-crevicular fluid, like S100A proteins, α-defensins and β-thymosins[78], by the plasma exudate, like serum albumin and many others present in very low amount, and by the contribution of the microbial oral flora. During their transit in the secretory pathway, salivary proteins can undergo to some PTMs, including loss of signal peptide and several post translational modification such as phosphorylation, N-terminal acetylation, glycosylation, sulfation and proteolytical cleavages[79]. Further modification of proteins and peptides could occur in the oral cavity by the action of oral enzymes.

Saliva has multiple functions (Figure 3): firstly, it is involved in many physiological processes, like mastication and early food digestion; in addition, it plays important protective functions on oral tissues and teeth, including the control of the demineralization and remineralization of teeth, the formation of the protein pellicles covering the oral surfaces, and the modulation of the adhesion of microorganisms to teeth and mucosa[80]. Indeed, saliva plays an important role in the health maintenance of oral cavity, through its antibacterial and antiviral activity, and in the lubrification and repair of oral mucosa[81].

Considering the great variety functions and its role in oral homeostasis, in the last decades a great number of studies focused on the comprehension of the composition and function in which salivary proteins and peptides are involved in order to understand their mechanism of action, their interaction and relation with other components of the oral cavity[75,79,82]. Most of them are specific of the oral cavity, while others that are not part of the normal salivary constituent are expressed also in other tissues and biological fluids and can reach saliva by several ways, such as passive transfer, diffusion or ultrafiltration. These proteins, such as S100A, α-defensins and β-thymosins, can allow monitoring physiological/pathological alteration not only in oral cavity but also at systemic level. Considering its easy, rapid and non-invasive collection, saliva is considered an ideal candidate for diagnostic and prognostic purposes, both for oral pathologies and systemic diseases[83–85]. Indeed, alteration in salivary profile in different physiological states, like age, gender, diet or circadian variation[75,79,83,86,87] or in pathological states could be virtually helpful for diagnostic aims[82]. Several research groups have been working on this field, showing the possibility to define different protein profiles that characterize the different diseases and underling significant qualitative/quantitative differences in some of the salivary proteins/peptides between patients and controls[88–90].

For example, salivary proteome in subject with oral carcinoma showed increased levels of transferrin[91], many truncated forms of cystatin SN and 46 proteins and peptides with significantly different concentration compared to healthy controls[92]. Moreover, studies performed on saliva from subject affected by head and neck squamous cell carcinoma have identified several potential biomarkers by using SDS-PAGE MALDI TOF/TOF-MS. Other biomarkers involved in inflammation and bone resorption have been characterized coupling 2-DE to MALDI TOF/TOF-MS[93].

Down's syndrome, a frequent genetic disorder in human, have increased risk of health problems associated to this condition. Change in the salivary proteome profile of patients with Down's syndrome, when compared to controls, were highlighted by HPLC-ESI-MS-based top down proteomics approach[94]. In particular, S100A7, A8 and A9, and A12 proteins exhibited significantly increased levels that could be of particular interest as a biomarker of early onset of Alzheimer disease, which is frequently associated to Down syndrome[94].

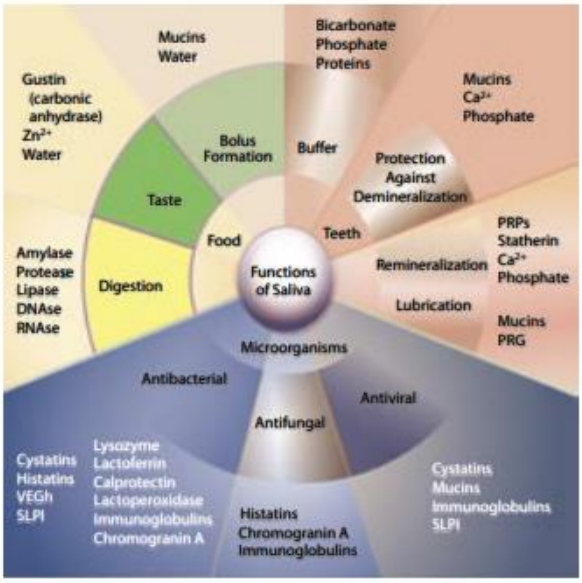
The study of Castagnola et al. on subjects affected by autism spectrum disorders highlighted the hypo-phosphorylation of histatin-1, statherin and different forms of aPRPs in

approximately the 60% of patients[95], indicating a lower activity of the FAM20C kinase responsible for the phosphorylation of the peptides/proteins in the salivary glands. The same kinase was found active in many other tissues, brain included, where could be involved during the proper development in the early life. This study indicates that the analysis of salivary phospho-peptides could help to discriminate a considerable subgroups of autism spectrum disorder patients[95].

SAPHO patients exhibited increased levels and frequency of S100A12 protein, a pro-inflammatory protein whose high expression is probably related to the inflammatory response and to the altered neutrophil responses to functional stimuli that characterize this disease; thus, S100A12 could be used as a salivary biomarker[96].

The salivary proteome of Wilson's disease patients reflexed oxidative stress and inflammatory conditions characteristic of the pathology, underling difference that could be useful clues of disease exacerbation[97].

Figure 3: Composition of human saliva and function of its principal components.



Proline-rich proteins

PRPs represent the major fraction of human saliva, more than 60% in weight of the salivary proteome; it is possible to distinguish acid PRPs (aPRPs), basic PRPs (bPRPs) and basic glycosylated PRPs (gPRPs). This protein family shows high abundance of proline residues, which represent with glycine and glutamine 70-80% of the sequence[98].

aPRPs, secreted from parotid (70%) and submandibular and sublingual glands (30%), are expressed by two loci, *PRH1* and *PRH2*, located in 12p13 chromosome; *PRH1* encodes for PIF-s, Db-s and Pa isoforms, *PHR2* encodes for PRP1 and PRP2 isoforms. PRP-1, PRP-2, PIF-s (usually indicated as PRP-1 type) and Db-s can undergo to proteolytic cleavage in Arg-106 by the action of a pro-protein convertase, producing a 44 amino acids peptide, called P-C peptide and corresponding to the common C-terminus, and four truncated isoforms called PRP-3, PRP-4, PIF-f (PRP-3 type) and Db-f. Pa isoform, which contains a cysteine residues at position 103 instead of arginine due to the loss of the consensus di-basic sequence of the convertase, cannot undergo to proteolysis but is found in dimeric form through a disulfide bond. The main isoforms have a pyro glutamination at N-terminus and are phosphorylated in Ser-8 and Ser-22 by FAM20C kinase[99]; non-phosphorylated, mono-phosphorylated of both PRP-1 and PRP-3, and tri-phosphorylated on Ser-17 of PRP-1 proteoforms are also detectable[100]. Dimerization of aPRPs can occur after secretion due to the action of transglutaminase[101]. These proteins regulate the calcium homeostasis of the oral cavity and participate on the formation of protein pellicle that covers and protects oral mucosa and tooth enamel. aPRPs are able to interact with some bacterial species and participate to selective colonization of oral microflora[99].

bPRPs are the more complex family of salivary protein in terms of genetic variability and post-translational modifications[99]. These proteins are secreted by parotid glands and are the expression product of four loci, *PRB1*, *PRB2*, *PRB3* and *PRB4*, located in 12p13.2 chromosome near aPRPs genes. In all of the loci are present at least three alleles called *S* (small), *M* (medium) and *L* (large)[102,103]; the fourth allele called *VL* (very large) is described only for *PRB1* and *PRB3* loci[104]. Each allele encodes for a big pre-pro-protein that undergoes to proteolytic cleavage from a convertase and a C-carboxypeptidase that generates mature bPRPs. Some of them can be phosphorylated by Fam20C kinase or N- or O-glicosilated[99]. From *PBR1* pro-proteins derive II-2 peptide (from *S*, *M*, *L* alleles), P-E peptides (*S* allele), IB-6 protein (*S* allele), Ps-1 (*M* allele) and Ps-2 proteins (*L* allele). From *PBR2* pro-protein are obtained IB-1, P-J, P-H,

P-F peptides and IB8a protein (*L* allele). *PRB3* and *PRB4* pro-proteins produce glycosylated proteins and P-D peptide (*S*, *M*, *L* alleles). bPRPs functions are not completely known, but some of them bind to tannins preventing their absorption and toxic effect in the gastrointestinal tract[105]. Moreover, these proteins are also involved in bitter taste sensitivity[106,107].

Statherin e P-B peptide

Statherin is a 43 amino acids peptide secreted by both parotid and submandibular glands[108], encoded by *STATH* gene located in 4q13-19 chromosome[109]. The peptide is detectable in saliva as a di-phosphorylated at Ser-2 and Ser-3, mono-phosphorylated on Ser-2, and non-phosphorylated proteoforms. In adult human saliva are also detectable both C- and N-terminal truncated proteoforms[110]. SV-1 (lacking the C-terminal Phe43) and statherin des-Thr42-Phe43 generate during maturation in secretory granules[108], while the N-terminal removal generates statherin des-Asp1, statherin des 1-9, statherin des 1-10 and statherin des 1-13. Statherin has high affinity for calcium ions and, as aPRPs, plays an important role in calcium homeostasis in the oral cavity and in calcium and phosphate transport during salivary gland secretion[109,111,112]. In saliva it has also been observed a cyclic form called cyclic-statherin, that it is formed for the action of type-2 transglutaminase naturally present in oral cavity, which is involved in the formation of the protein pellicle covering the enamel and of oral mucosa[101].

P-B is a small peptide rich in proline encoded by *PROL3* gene, located on chromosome 4q13.3[108], very close to the statherin gene, suggesting a functional relationship with this proteins. Its amino acid composition is similar to that one of salivary PRPs from which differs for the presence of three tyrosine residues in its sequence[110]. P-B peptide is principally secreted by submandibular and sublingual glands and in a smaller amount by parotid[108]. There are different truncated forms of this peptide in saliva, whose cleavage probably occurs after secretion[108]. Differently from statherin, none specific function for P-B peptide and its fragment has been proposed[108,110].

Histatins

Histatins are small cationic peptides secreted by major and minor salivary gland, whose name derives from the high number of histidine residues that are present in their sequences. These peptides derive from the expression of two genes, *HIS1* and *HIS2*, located in 4q13 chromosome. The products of these two genes are respectively histain-1 (His-1) and histatin-

3 (His-3) that, despite the high sequence homology, follow different maturation pathways and show different biological activity. His-1 is a peptide of 28 amino acids phosphorylated in Ser-2 probably by FAM20C kinase that recognized he consensus sequence Ser2-Asp3-Glu4[100,108]. His-1 is detectable in saliva both phosphorylated and non-phosphorylated form. Moreover, sulfated proteoforms on Tyr residues of His-1 were characterized in saliva[113]. Nowadays it is difficult to understand the biological significance of these modifications, in any case, the role of His-1 to participate to wound healing and cell migration has been demonstrated[114,115]. His-3 is non-phosphorylated, indeed it does not has the consequence recognizable by FAM20C; smaller histatins are generated by multiple cleavages of His-3[116]; among them, histatin-5 (His-3 1-24) is the most abundant[108,116]. Zhou and colleagues propose that the proteolytic cleavage happens to combine action of convertase and peptidase located in trans cisternae of Golgi and in secreted vesicles[117]. Histatins play an important role in homeostasis and in innate defense of horal cavity, acting as a strong antibacterial and antifungal agent[116], especially against Candida Albicans[117].

Cystatins

Cystatins are a superfamily of inhibitory cysteine proteases[118]. The family includes type 1 cystatin (cystatin A and B), type 2 cystatins (C, D, E, F, S, SN, SA), type 3 cystatin or kininogens. Cystatin A and B are inhibitors of papain and cathepsin L, S and H.

Cystatin A is expressed in epidermises[119], oral mucosa[120], lymphocytes, neutrophils[121], lymphoid tissue[122], while cystatin B is widely expressed in different cells and tissues[123,124]. It is possible to detect two isoform of cystatin A, acetylated and non-acetylated on its N-terminal[83,125]; moreover, Manconi et al. highlighted the presence of another variant, cystatin A T96→M, that could be both acetylated or not acetylated in N-terminus[125]. Cystatin A is a proteases inhibitor involved in cellular proliferation and augmented levels have been correlated to atopic dermatitis[126] and have been found in noncancerous tissues surrounding hepatocellular carcinoma cells[127]. Moreover, polymorphism in the gene for cystatin A have been associated to different inflammatory skin disease such as congenital exfoliative ichthyosis[128] and acral peeling skin syndrome[129].

Cystatin B is N-terminally acetylated and usually it is detectable in adult saliva only as oxidized proteoforms on its Cys-3 residue: glutathionylated (SSG), cysteinylated (SSC) and disulfide dimer (S-S)[130]. In addition to his cathepsin inhibitory power, cystatin B shows other important

functions, as prevention of apoptosis[131], reduction of oxidative stress[132] and neuroprotective role acting like a chaperon on Aβ peptide[133].

Type 2 cystatins comprise five salivary cystatins (S, S1, S2, SA, SN) and the cystatins C and D. Cystatin S, SN and SA are produced mainly by submandibular glands, present 90% of sequence homology and are the consequence of genic expression of a genic cluster in 20p11.21 chromosome[108]. Cystatin S may be mono-phosphorylated on Ser-3 (cystatin S1, represent 60-70% of the total amount) or di-phosphorylated on Ser-1 and Ser-3 (cystatin S2, 20-30%)[86,108]. Recently, other PTMS of salivary cystatins were characterized in saliva[125]. Cystatins are inhibitors of cysteine protease, suggesting a role on oral cavity of protection against pathogens and on the control of lysosomal cathepsins[125]. This activity is marked for cystatin SN and less evident for SA and S, which have more tendency on calcium binding; for this reason, these proteins seems have, with statherin, an important role on mineral homeostasis of teeth[134].

Cystatin C is a powerful inhibitor of cysteine-protease such as cathepsin B, H, L and S; it is made of 120 amino acids and it is widely expressed in all cells and tissues, salivary glands included. It seems to have a protective function on connective tissue, preventing its degradation from enzymes that escape from lysing cells or whose secretion has been altered by malignant cells[135]. Recent study has demonstrated that oligomerization of this protein leads to the formation of amyloids plaques in older people. This pathological pathway is quick in case of a mutant protein[136].

Cystatin D is a peptide made of 122 amino acids and acts as a strong inhibitor against cathepsin S, H and L[137,138]. Furthermore, cystatin D is involved in the inhibition of proliferation, migration and invasion of colon carcinoma cells[139], regulation of antigen presenting cells activity[140] and modulation of gene expression related to its nuclear activity localization[141]. This peptide exists in two natural variants with cysteine or arginine in position 26 because of gene polymorphism[137], indicated in this manuscript as cystatin D-R26 and cystatin D-C26. Manconi et al. identified in acid-soluble fraction of saliva different proteoforms of cystatin D: cystatin D-R26 lacking the first four amino acids from the N-terminus (des1-4) and carrying two disulfide bonds; cystatin D-R26 des 1-8 with two disulfide bonds; cystatin D-R26 des1-5 with the N-terminal glutamine residue converted to pyro-glutamic acid and with two disulfide bonds[125]. Cystatin D-C26 has been never observed in salivary samples by HPLC-ESI-MS-based top-down proteomic analysis of acid-soluble fraction of salivary proteins, probably because

the presence of the cysteine lead protein precipitation during acid-treatment. The presence of the cysteine residue could lead to the formation of insoluble adducts with many interactors. A preliminary study, performed with the aim to characterize both cystatin D variants in adult saliva, highlighted the participation of cystatin D-C26 in the formation of a protein complex of ca 250kDa by interaction with protein partners. The identified protein partners are involved both in innate/acquired immune system and inflammatory responses, such as immunoglobulins, polymeric immunoglobulin receptor (PIgR), lysozyme C, S100A8 and S100A9[142].

Antileukoproteinase

Antileukoproteinase or human secretory leukocyte protease inhibitor (SLPI) is a basic non-glycosylated protein made of 107 amino acids residue rich in cysteine that participates to innate immune system. SPLI is the genic product of *SLPI* gene, located in 20q12-13.2 chromosome[143]. This protein is produced by neutrophils, macrophages, pancreas β-cells, epithelial cells associated to respiratory tract mucosa, parotid and submandibular gland[143]. It was identified in many secretions like whole saliva, seminal fluid, breast milk, synovial fluid, tears and cerebrospinal fluid[83,143]. The mainly function of SLPI is protect tissues against the consequences of inflammatory conditions: it protects tissue inhibiting elastase and trypsin secreted from neutrophils, which can lead to wide damage and tissue degradation as well as in disease like cystic fibrosis, pulmonary emphysema, chronic bronchitis and bacterial pneumonias[143]. In fact, thanks to its cationic character, it is able to aggregate onto microbial membranes and lead to destruction of the membrane via formation of ion channel, transmembrane pores or membrane rupture[144].

α-defensins

Defensins are cysteine-rich, cationic small peptides that contain 18-45 amino acid residues. Mammalian defensins are classified in alpha (α), beta (β) e theta (θ) depending on the pattern of disulfide bridges formation between six conserved cysteine residues[145]. α-defensins (HNP – Human Neutrophil Peptides) are salivary peptides rich in cysteine and tyrosine with 3 disulfide bounds[145]. Recent studies have found that in saliva are present α-defensin 1, 2, 3 and 4 and β-defensin 1 and 2[146]. Salivary α-defensins comes from GCF, in which defensins are the principal secretory component of the neutrophils[147] and represent 40% of total protein amount[148].

GCF origins from gingival plexus and carries also serum proteins, inflammation mediators, metabolites and degradation products of host cells[147]. α-defensins are expressed in neutrophils in which they play a role in the oxygen-independent killing of phagocytosed microorganism linked to their antimicrobial and antiviral activity, therefore they play an important role in innate immunity responses[149]. Their first interaction with bacteria is mediated by their positive net charge that match with the negative net charge of the bacterial cell membrane. The electrostatic forces lead to bound, aggregation and integration of these peptides into the lipid bilayer. The integration of defensins into the bacterial membranes results in the formation of ion channels, transmembrane pores and membrane rupture, resulting in the destruction of the bacteria[150,151]. Defensins can also act as regulators of cellular volume and chemotaxis and inhibitors of natural-killer cells[145].

β-thymosins

β-thymosins are ubiquitous peptides with important intra and extra-cellular functions. Thymosin β-4 (Tβ-4), encoded by *TMSB4X* gene clustered on Xp22.2 chromosome, is the most abundant. It is a moonlighting peptide that plays different biological roles, among them the modulation of actin polymerization and thus the rearrangement of the cytoskeleton, the endothelial cell migration[152], differentiation and maturation of endothelial cells[153], tissue repair and regeneration[154]. Several studies demonstrated its ability to protect heart[155], brain[156] and cornea[157] from tissue damages. Tβ-4 is a strong angiogenesis stimulator, indeed thymosin β-10 (Tβ-10), encoded by *TMSB10* located on 2p11.2 chromosome, inhibits this pathway[158]. Tβ-4 has been identified in whole saliva and tear using immunological techiques[83,159,160]. The study of Castagnola et al. on preterm human newborn saliva evidences a high concentration of Tβ-4 and β-10 in the oral cavity, suggesting a role of these peptides in the development of oral cavity and showing that salivary gland possess specific secretory capability even during prenatal maturation stages. Then, during the middle phases of fetal development the switch of secretion happens, with a progressive decrease of thymosins concentration to adult-like levels [83,161]. In adult, the presence of Tβ-4 and Tβ-10 in salivary fluid is due to the contribution of GCF, where these peptides are released by neutrophils[160]. In saliva, thymosins act as antibacterial, anti-inflammatory and antiapoptotic agent on gingival fibroblasts[162].

S100A proteins

The S100A proteins belong to a low-molecular weight family of protein carrying Ca^{2+}-binding EF-hand motif and are expressed only in the vertebrates[163]. Most of S100-coding genes are clustered on 1q21 chromosome; this organization gave rise to the nomenclature of the S100 proteins: those ones encoded by genes located in the cluster on chromosome 1 are called S100A proteins with number 1-16 depending of the position of the gene in the cluster[164]. The remaining S100 genes are located on chromosome 21q22(S100B), Xp22 (S100G), 4p16 (S100P) and 5q14 (S100Z).

S100A are characterized by two Ca^{2+}-binding EF-hand motifs connected by a linker. The C-terminal motif is typical of Ca^{2+}-binding proteins, while the N-terminal motif owns an extended specific structure for S100 proteins ("pseudo EF hand"), which cause a reduced Ca^{2+} affinity. The bound of calcium provokes structural modifications that allows them to bind and modulate the activity of other proteins[165]. They are constitutively expressed in neutrophils, myeloid cells, platelets, osteoclast and chondrocytes, but can be induce and overexpressed in macrophages, monocytes and fibroblasts in acute and chronic inflammatory and oxidative stress conditions[166,167]. Their involvement has been indicated in a wide range of intracellular and extracellular functions: regulation of calcium homeostasis, cytoskeletal rearrangement, contraction and motility, cell growth and differentiation, membrane organization, promotion of wound repair, protection against microbial proliferation chemotaxis, apoptosis, promotion of wound repair, protection against microbial proliferation, control of ROS formation, inflammation and protein phosphorylation and secretion[166,168]. Their activity could be regulated by formation of homodimers and heterodimers and through PTMs such as phosphorylation, N-terminal acetylation and oxidation of cysteine, tryptophan and methionine residues, which could change their ability to bind to ions or target proteins[166,169,170]. Among the S100Aproteins, S100A7, S100A8, S100A9, S100A11 and S100A12 have been detected in human saliva[79,83,108].

S100A7, also called psoriasin, is present in two isoforms, E-27 and D-27, of which D-27 is present in adult saliva[83]. Both S100A7 isoforms are N-terminal acetylated following the loss of the initial methionine. S100A7 acts like a chemotactic agent and it is able to attract lymphocytes and neutrophils[171]. Its overexpression is related to many epidermal inflammatory diseases, such as atopic dermatitis, mycosis fungoides and inflammatory lichen sclerosis and

atrophicus[172]. Moreover, since its high level of expression in psoriatic lesions, investigators propose that it may have chemotactic role in psoriasis[172]. S100A7 expression is also increased in invasive skin cancers such as squamous cell carcinoma[171].

S100A9 was identified on human granulocytes as four isoforms[173], which were also characterized in saliva from newborns[83]. Two of them, defined long-types, are acetylated following loss of the N-terminal methionine residue; they differ each other for the phosphorylation in threonine 112. The other two, defined as short-types, are acetylated following the loss of the last five amino acid residues and differ from each other in the phosphorylation of threonine 112[83]. Recently, many modified forms of S100A8 and S100A9 have been discovered. In fact, these proteins are readily oxidized by reactive oxygen/nitrogen species, which promote structural changes that modify their functions. For example, oxidation of murine S100A8 to disulfide-linked dimers[174] or intermolecular sulfonamide-linked[175] complexes blocks its chemotactic activity, while the activity of human S100A8 and S100A9 regarding neutrophils migration is regulated by oxidation[176,177]. Lim et al. showed that S100A8 and S100A9 are S-nitrosylated by the physiological NO donor S-nitrosoglutathione (GSNO) *in vitro*[178]. Moreover, my research group characterized derivative of S100A8 and S100A9 with different levels of oxidation via an integrated top-down and bottom-up approach on acid soluble fraction of whole saliva in Wilson's disease patients[97]. S100A8 oxidation includes methionine 1 and 78 (M1, M78), tryptophan 54 (W54) and cysteine 42 (C42). Other proteoforms were discovered, which shows C42 oxidized to sulfonic acid (S100A8-SO_3H). One of them has a further oxidation at W54 (S100A8-SO_3H/W54ox), another has W54 and M78 oxidized (S100A8-SO_3H/W54ox/M78ox) and the last has W54 dioxidized (S100A8-SO_3H/W54diox). All these proteoforms were named hyper-oxidized S100A8[97]. The glutathionylation in Cys-41 of S100A8 was also characterized, as well as the disulfide bound with S100A9 that generates S100A8/A9 dimer[97].

Salivary antibodies and polymeric immunoglobulin receptor (PIgR)

Immunoglobulins (Ig) or antibodies are the antigen recognition molecules of B cells. An Ig is made of 2 identical heavy chains and 2 identical light chains (kappa or lambda), joined by disulfide bridges that bind a heavy chain to a light chain and the two heavy chains together. Each Ig heavy chain has an N-terminal variable region containing the antigen-binding site and a C-terminal constant region that provides effector or signaling functions. Igs are the product

of expression of genes located in three major loci: 14q32.33, 2p11.2 and 22q11.2[179].

Salivary antibodies constitute the first defense against antigens in saliva, in the mucosal layer on the epithelial surfaces and in the acquired enamel on the tooth. In saliva are present two principal classes of immunoglobulins, secretory IgA and IgG[144,180–182]. Secretory IgA, originating from salivary glands and mucosal cells, is primarily in dimeric form even though in some case could be in polymeric form, while salivary IgG is monomeric and it origins from serum and passes in saliva through gingival crevices[182]. Despite IgA and IgG represent the most abundant antibodies, it is possible to find in human saliva very small amount of IgM, IgD and IgE antibodies[180].

Salivary secretory immunoglobulins are produced by specific B cells that stay in the salivary glandular stroma and in the oral mucosa[183,184]; these cells release mainly dimeric and also polymeric IgA and a few pentameric IgM, stabilized by an incorporated 15kDa J chain[182]. Both molecules are then internalized from interstitial space and exported to the saliva by salivary epithelial transport mechanism using the polymeric immunoglobulin receptor (PIgR) and its secretory component (SC), which exhibits high affinity to the J chain of dimeric IgA and pentameric IgM[182]. PIgR is a glycoprotein of 100kDa, which is constitutively expressed in basolateral surface of glandular epithelial cell[182] and mucosal epithelial cells[185]. At the basolateral epithelial surfaces, the J chain binds to SC, leading to molecular transport that starts with the internalization of SC bound to the antibodies and ends with the exocytosis in the apical cell surface into the glandular lumen or into the oral cavity. The process is paired to the cleavage of SC from PIgR[182]. Then, the 80kDa SC could be incorporated to the antibodies, providing resistance against proteolytic degradation[182], or stay unbound exerting several innate defense functions in oral cavity, such as inhibition of epithelial adhesion of bacteria[186] or neutralization of certain bacterial toxins[187].

Aim of the study

This study aroses in order to identify new potentially salivary biomarkers, which could be helpful in both diagnosis and classification of the different forms of mastocytosis. Saliva is an ideal biological fluid since its easy, painless and non-stressed collection. Furthermore, human saliva reflects both oral cavity physio-pathological conditions and systemic conditions, since in saliva are present also peptides and proteins which act in other tissues and organs. Finally, the characterization of salivary proteome has now reach high levels thanks to advanced platforms. Therefore, in this thesis are going to present the results of a study of the quantitative/qualitative variation of the salivary proteome in patients affected by different forms of mastocytosis compared each other and with a controls group.

The aim of the part 1 of the thesis was to use a top-down proteomic approach to highlight qualitative and quantitative differences in the protein profile of acid-soluble fraction of salivary samples analyzed by high-performance liquid chromatography coupled to and electrospray source-ion trap mass spectrometer (HPLC-ESI-IT-MS) between the three groups of patients and the healthy subjects and among the patient groups. This could be helpful to find potentially useful biomarkers to discriminate the different forms of the disease.

The part 2 was focused on the study of protein immune-complexes associated to cystatin D (SIC-D) present in saliva. The study moved through this way since the preliminary data on the characterization of this high molecular weight complex incorporating Igs, PIgR, and several proteins involved in the innate and acquired immunity or in inflammation. Moreover, after the evaluation on the results obtained about cystatin D on the part 1 of this thesis, it was considered interesting to study the formation of SIC-D in the different patient groups under study and underline eventually qualitative and quantitative variations. The SIC-D was detected by western-blotting and purified by immunoprecipitation. The qualitative/quantitative characterization was performed by in-gel trypsin digestion and nano HPLC-ESI-high-resolution-MS/MS analyses.

The experiments and my research activity were performed partly during my abroad period project at the Chemistry Department of the University of Aveiro (Portugal) under the supervision of Professor Francisco Amado, and partly at the University of Cagliari under the supervision of prof. Tiziana Cabras.

PART I: PROTEOMIC INVESTIGATION OF ACID-SOLUBLE FRACTION OF SALIVA FROM PATIENTS AFFECTED BY MASTOCYTOSIS USING A TOP-DOWN APPROACH

Materials and methods

Reagents and instrument

All the chemicals and reagents used for high performance liquid chromatography separation coupled to electrospray-ion trap mass spectrometry (HPLC-ESI-IT-MS) analysis were purchased from Sigma Aldrich (St. Louis, MO, USA). HPCL-ESI-MS analyses were performed with a Surveyor HPLC system connected to an LCQ Advantage mass spectrometer (Thermo Fisher Scientific San Jose, CA, USA). The chromatographic column for the low-resolution HPLC-MS analysis was a Vydac C8 reverse phase (Hesperia, CA, USA) (150 x 2.1mm, particle diameter 5µm).

Study subjects and controls

6 patients with cutaneous mastocytosis (CM) (2 males, 4 females, mean age ± SD: 37.3 ± 15.7) and 35 patients with systemic mastocytosis (SM) (16 males, 19 females, mean age ± SD: 48 ± 16), were enrolled from the Internal Medicine and Immunology outpatient's clinic of University of Cagliari. Among the 35 SM patients, 8 manifested just systemic symptoms (SM-C), whereas the remaining 27 both cutaneous and systemic symptoms (SM+C). 48 sex/age matched healthy controls were enrolled as volunteer among the staff of the Department of Life and Environmental Science, University of Cagliari. The informed consent process agreed with the latest stipulations established by the Declaration of Helsinki. The local review boards approved the study, and in view of its observational nature, a formal ethical committee approval was approved.

All the samples of unstimulated whole saliva were collected during the morning and in fasting condition with a plastic pipette. Salivary samples were immediately treated a solution of 0.2% solution of trifluoroacetic acid (TFA) containing 50µM leuenkephalin as internal standard in 1:1 v/v ratio. Then, each sample was centrifugated 15 minutes at 13400 rpm at 4°C. Finally, the acid supernatant was separated from pellet and stored at -80°C until RP-HPLC-ESI-MS analysis.

Clinical data

Demographic and clinical features of the included patients are reported in Table 4. The patients have been classified in CM (n=6), SM-C (n=8) and SM+C (n=27) basing on the clinical criteria established by WHO[20]. Patients who showed hematologic diseases associated to mastocytosis were excluded from the analysis. CM patients showed low levels of tryptase and no one exhibited mutation in c-KIT receptor. Two of them exhibited also both aphthous ulcers and GI symptoms. On the other hand, all the SM-C and SM+C patients showed high levels of tryptase, that in 16 SM+C and in 2 SM-C patients were coupled by the mutation of c-KIT, mainly D816V. Only the patient #9 exhibited M541L mutation, while the patient #12 displayed both D816V and M541L mutations. Some of the patients included in the group SM+C were primarily classified as cutaneous before the manifestation of systemic symptoms. Then, the genetic analysis and the bone marrow biopsy confirmed the systemic mastocytosis in these patients. Moreover, only a SM-C patient and 9 SM+C patients showed aphthous ulcers, while 50% of SM-C and all the SM+C patients suffered of GI symptoms.

Samples' collection and treatment

All the samples of unstimulated whole saliva were collected according to a standardized protocol. Donors did not eat or drink at least 2h before the collection, which was performed in the morn morning between 10:00 a.m. and 12:00 a.m.. They were invited to sit assuming a relaxed position and to swallow. Whole saliva was collected as it flowed into the anterior floor of the mouth with a soft plastic aspirator for less than 1 min, and transferred to a plastic tube, cooled on ice. The volume of saliva (from 0.5 to 1ml, depending on the donor) was divided to be utilized in the two parts of the study described in this thesis.

a) 100µl of saliva of each sample were immediately diluted in a 1:1 v/v ratio with a 0.2% solution of trifluoroacetic acid (TFA) containing 50µM of leu-enkephalin as internal standard. Then, each sample was centrifuged at 13400 RPM for 15 minutes at 4°C. Finally, the supernatant was separated from the precipitate and stored at -80°C until the RP-HPLC-ESI-MS analysis.

b) The remaining volumes of whole saliva of each sample was treated with PBS buffer (270mM NaCl, 5mM KCl, 20mM Na_2HPO_4, 4mM KH_2PO_4) containing a cocktail of protease inhibitors (mini-cOMPLETE EDTA-free in 1:2.5 v/v ratio with PBS buffer – Sigma-Aldrich/Merck, Darmstadt, Germany), in a 1:1 v/v ratio and stored at -80°C until the immuno-proteomics analysis described in the second part of the thesis.

RP-HPLC-ESI-MS analysis

For the HPLC-ESI-low-resolution MS analysis the chromatographic separation was carried out with the following eluents: A, 0.056% TFA in water; eluent B, acetonitrile/water 80:20 with 0.05% TFA. The gradient applied was linear from 0 to 55% of in 40 minutes, and from 55% to 100% of B in 10 minutes, at a flow rate of 0.1ml/min. During the first 5 minutes of the analysis, the flow was not direct to the mass spectrometer to avoid instrument damages from the high salt concentration. Mass spectra were collected every 3ms in the m/z range 300-2000 in positive ion mode; the MS spray voltage was 5.0 kV, and the capillary temperature was 255 °C.

Low-resolution MS data analysis

HPLC-low-resolution-MS Total Ion Current (TIC) chromatographic profiles were analyzed to selectively search and quantify the peptides/proteins reported in Table 5, which shows UniProt-KB codes, elution times, experimental and theoretical average mass values (Mav) of the proteins/peptides, comprised their PTMs derivatives, included in the study. Table 5 also reports the multiply-charged ions used for eXtracing Ion Current (XIC) search, which were selected excluding values common to other closely eluting proteins. A window of ± 0.5 Da was used to extract XIC peaks. Experimental Mav were obtained by deconvolution of averaged ESI-MS spectra automatically performed by using MagTran 1.0 software[188]. Mav and elution times of proteins/peptides were compared with those determined on salivary samples, under the same experimental conditions, in our previous studies[79]. Experimental Mav were also compared with the theoretical ones available at the UniProt-KB human data-bank (http://us.expasy.org/tools). The proteins and peptides analyzed in this study have been already characterized in other previous studies by my research group. Structural characterization of acidic proline-rich proteins (aPRPs)[79,189], histatins (His-1, His-3, His-3 1/24, His-3 1/25)[108], statherin[108], P-B peptide[108], thymosins β-4 and β-10[160], cystatins A, B, C, D and

S-type[94,125,130], S100A8[97], S100A9[97], S100A7[83], S100A12[83], antileukoproteinase[83], and prolactin-inducible protein (PIP)[190], performed by bottom-up and top-down proteomic platforms based on high-resolution MS/MS analysis, has been already described in our previous proteomic investigation of human saliva.

Quantification

XIC peak areas were integrated by using the following peak parameters: baseline window 15; area noise factor 50; peak noise factor 50; peak height 15% and tailing factor 1.5. Area of the XIC peaks is proportional to the protein concentration, and, under constant analytical conditions, it allows performing relative quantification of the same protein in different samples [100,191]. Estimated percentage error of the XIC procedure was <8%. Eventual dilution errors occurring during sample collection were corrected by normalizing XIC peak areas of peptides/proteins with the XIC peak area of leu-enkephalin used as internal standard.

Statistical analysis

GraphPad Prism software (version 6.0) was used to calculate means and standard deviations of protein XIC peak areas and to perform statistical analysis. Data distributions were tested for normality by D'Agostino-Pearson test. The comparison between the three groups of patients and controls and between the three groups of patients has been carried out with Mann-Whitney and Unpaired t test with or without Welch's correction, depending on the distribution (skewed or normal) and the variance (unequal or homogeneous). Specifically, a non-parametric t-test was chosen if the distribution of XIC peak areas was not Gaussian for at least one of the two groups (Mann-Whitney t-test), while a parametric test was used 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 has been considered significant when the p-value was <0.05. Results from t-tests analysis are also graphically shown by the Volcano plot, which plots $\log_2$ (fold-change) mean value of XIC peak areas from patients/controls ratio or patients/patients ratio on the X-axis against $\log_{10}$ of p-value from the t-test analysis on the Y-axis.

Correlation analyses were also performed with the Spearman or the Pearson tests based on the distribution of the data (skewed or normal). The correlation was performed relating XIC peak area of the proteins/peptides under analysis and the tryptase levels of each patients.

Table 4: Age/gender, diagnosis, tryptase serum level, eventual mutation of c-KIT receptor and presence of oral or gastrointestinal symptoms of each patient at the time of the study.

PATIENT	AGE/ GENDER	DIAGNOSIS	SM WITH CUTANOUS SYMPTOMS	TRYPTASE (µg/L)	MUTATION(S)	ORAL SYMPTOMS	GI SYMPTOMS
#2	59/F	SM		17.5		N	Y
#5	54/F	SM		16.5		N	Y
#6	72/F	SM		41		N	Y
#9	45/M	SM	SM+C	33.6	M541L	N	Y
#10	44/F	SM	SM+C	16.2		Y	Y
#11	58/F	SM	SM+C	20.2	D816V	Y	Y
#12	53/M	SM	SM+C	5.52	D816V/M541L	N	Y
#13	35/F	SM	SM+C	44.7		Y	Y
#14	22/F	CM		5		N	N
#15	38/F	SM	SM+C	7.66	D816V	Y	Y
#16	52/M	SM	SM+C	37.5	D816V	N	Y
#17	46/F	CM		4.45		Y	Y
#18	55/M	SM		27.1	D816V	Y	Y
#19	48/M	SM	SM+C	16.2		N	Y
#20	32/M	SM	SM+C	53.5	D816V	N	Y
#21	44/F	SM		16		N	N
#22	60/F	SM	SM+C	23.6		N	Y
#23	62/F	SM		80.4	D816V	N	N
#24	49/M	SM	SM+C	29.1	D816V	N	Y
#25	49/F	SM	SM+C	37.4		N	Y
#26	35/F	SM	SM+C	16.5		Y	Y
#27	67/F	CM		7.23		N	N
#28	38/F	SM	SM+C	26.2	D816V	Y	Y
#29	57/M	SM	SM+C	77.4	D816V	Y	Y
#30	32/M	SM	SM+C	40.8	D816V	Y	Y
#31	39/M	CM		8.84		Y	N
#35	34/M	SM	SM+C	28.7		Y	Y
#36	52/F	SM	SM+C	16.1	D816V	N	Y
#37	28/F	CM		7.7		N	N
#43	71/F	SM		9.67		N	N
#45	23/M	CM		4.27		N	N
#46	75/F	SM		18.2		N	N
#47	46/F	SM	SM+C	7.26		N	Y
#51	29/M	SM	SM+C	64.5		N	Y
#52	44/M	SM	SM+C	7.41	D816V	N	Y
#58	38/M	SM	SM+C	26.3		N	Y
#59	67/F	SM	SM+C	30.6		N	Y
#60	49/F	SM	SM+C	15.8	D816V	N	Y
#63	56/M	SM	SM+C	60.8	D816V	N	Y
#64	24/M	SM	SM+C	11.4	D816V	N	Y
#66	36/F	SM	SM+C	6.63	D816V	N	Y

Table 5: UniProt-KB code, experimental and theoretical average mass (Mav) values ± standard deviation, and elution times of proteins and peptides analyzed.

PROTEINS/PEPTIDES (SWISS-PROT CODE)	ELUITION TIME (MIN ± 0.5)	EXPER. (THEOR) AV. MASS (Mav)	M/Z (CHARGE)
Acid Proline-Rich Proteins			
PRP-1 di-phos. (P02810)	22.2	15515 ± 2 (15514-15515)	1293.9(+12), 1194.4(+13), 1035.3(+15), 970.7(+16), 913.6(+17)
PRP-1 mono-phos.	22.9	15435 ± 2 (15434-15435)	1287.2(+12), 1188.3(+13), 1030.0(+15), 965.7(+16), 908.9(+17)
PRP-1 non-phos.	23.2	15355 ± 2 (15354-15355)	1280.5(+12), 1182.1(+13), 1024.6(+15), 960.7(+16), 904.2(+17)
PRP-1 tri-phos.	21.6	15595 ± 2 (15594-15595)	1418.7(+11), 1300.5(+12), 1200.6(+13), 1040.6(+15), 975.7(+16)
PRP-3 di-phos. (P02810)	22.8	11161 ± 1 (11161-11162)	1595.5(+7), 1396.2(+8), 1015.7(+11), 931.1(+12), 859.6(+13)
PRP-3 mono-phos.	23.4	11081 ± 1 (11081-11082)	1584.1(+7), 1386.2(+8), 1008.4(+11), 924.5(+12), 853.4(+13)
PRP-3 non-phos.	23.8	11001 ± 1 (11001-11002)	1376.2(+8), 1101.2(+10), 917.8(+12) 786.8(+14)
PRP-3 diphos. desArg106	22.8	11004 ± 1 (11005-11006)	1573.2(+7), 1223.8(+9), 1001.5(+11), 847.6(+13)
P-C peptide (P02810)	15.0	4370.9 ± 0.4 (4370.8)	1457.9(+3), 1093.7(+4)
Statherin			
Statherin di-phos. (P02808)	29.2	5380.0 ± 0.5 (5379.7)	1794.2(+3), 1345.9(+4), 1076.9(+5)
Statherin mono-phos.	28.9	5299.9 ± 0.5 (5299.7)	1767.6(+3), 1325.9(+4), 1060.9(+5)
Statherin non-phos.	28.6	5220.5 ± 0.5 (5219.7)	1741.2(+3), 1306.1(+4), 1045.1(+5)
SV-1 (des-Phe43)	27.8	5232.4 ± 0.5 (5232.5)	1745.1(+3), 1309.1(+4), 1047.5(+5)
Statherin desThr42-Phe43	27.9	5131.2 ± 0.5 (5131.4)	1711.4(+3), 1283.8(+4), 1027.2(+5)
Statherin desAsp1	28.7	5264.7 ± 0.5 (5264.6)	1755.9(+3), 1317.2(+4), 1053.9(+5)
Statherin des1-9	28.5	4127.9 ± 0.4 (4127.6)	1376.9(+3), 1032.9(+4)
Statherin des1-10	28.0	3971.3 ± 0.4 (3971.4)	1986.7(+2), 1324.8(+3)
Statherin des1-13	27.5	3645.2 ± 0.4 (3645.0)	1823.6(+2), 1216.1(+3)
P-B peptide			
P-B peptide (P02814)	30.0	5792.9 ± 0.5 (5792.7)	1932.0(+3), 1449.2(+4), 1159.6(+5)
P-B des1-5	30.3	5215.0 ± 0.5 (5215.1)	1739.4(+3), 1304.8(+4), 1044.0(+5)
P-B des1-7	30.1	5060.1 ± 0.5 (5060.9)	1688.0(+3), 1266.2(+4), 1013.2(+5)
P-B des1-4	30.0	5371.0 ± 0.5 (5371.3)	1791.4(+3), 1343.8(+4), 1075.3(+5)
P-B des1-12	27.5	4549.0 ± 0.5 (4549.3)	1517.5(+3), 1138.3(+4)

PROTEINS/PEPTIDES (SWISS-PROT CODE)	ELUITION TIME (MIN ± 0.5)	EXPER. (THEOR) AV. MASS (Mav)	M/Z (CHARGE)
Histatins			
Histatin-1 (P15515)	21.9	4928.2 ± 0.5 (4928.2)	1644.1(+3), 1233.5(+4)
Histatin-1 non-phos.	22.0	4848.2 ± 0.5 (4848.2)	1617.4(+3), 1213.5(+4)
Histatin-3 (P15516)	17.7	4062.2 ± 0.4 (4062.4)	1355.1(+3), 1016.6(+4)
Histatin-3 1/25	14.3	3192.4 ± 0.3 (3192.5)	1065.1(+3), 799.1(+4)
Histatin-3 1/24	14.6	3036.5 ± 0.3 (3036.3)	1013.2(+3), 760.1(+4)
Cystatins			
Cystatin A (P01040)	31.8	11005.354 ± 2 (11006.5)	1001.59(+11), 1101.59(+10), 1223.94(+9), 1376.81(+8), 1573.36(+7), 1835.42(+6)
Cystatin A N-acetylated	33	11047.43 ± 2 (11048.5)	1005.41(+11), 1105.85(+10), 1228.61(+9), 1382.06(+8), 1579.36(+7), 1842.42(+6)
Cystatin B (P04080)	31.9	11179.3 ± 0.5 (11181.6)	1864.6(+6), 1598.4(+7), 1398.7(+8), 1243.4(+9), 1119.2(+10), 1017.5(+11)
Cystatin B SSG [a]	32.8	11485.8 ± 2 (11486.9)	1915.5(+6), 1642.0(+7), 1436.9(+8), 1277.3(+9), 1149.7(+10), 1045.3(+11)
Cystatin B SSC [b]	32.9	11299.8 ± 2 (11300.7)	1884.5(+6), 1615.4(+7), 1413.6(+8), 1256.7(+9), 1131.1(+10), 1028.6(+11)
Cystatin B S-S dimer [c]	34.3	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)
Cystatin C (P01034)	35.1	13342 ± 2 (13343.1)	1483.57(+9), 1335.32(+10), 1214.02(+11), 1112.93(+12), 1027.40(+13)
Cystatin C Met-ox.	35.3	13358.5 ± 1 (13358.4)	1485.28(+9), 1336.85(+10), 1215.41(+11), 1114.21(+12), 1028.58(+13)
Cystatin D-R26 des1-8 (P28325)	38	13163 ± 2 (13163.0)	1646.40 (+8), 1463.60 (+9), 1317.30 (+10), 1197.60 (+11), 1097.90 (+12), 1013.50 (+13)
Cystatin D-R26 des1-4 (P28325)	37.6	13605± 2 (13605.4)	1701.70 (+8), 1512.70 (+9), 1361.50 (+10), 1232.90 (+11), 1134.80 (+12), 1047.60 (+13)
Cystatin D-R26 des1-5 (P28325)	37.7	13517 ± 2 (13517.3)	1690.70(+8), 1502.90(+9), 1352.70 (+10), 1229.80 (+11), 1127.4 (+12), 1040.40 (+13)
Cystatin S non-phos. (P01036)	35.3	14186 ± 2 (14185)	1774.3(+8), 1577.2(+9), 1419.6(+10), 1290.6(+11), 1183.2(+12), 1092.2(+13), 1014.3(+14)
Cystatin S mono-phos. (P01036)	35.3	14266 ± 2 (14265)	1784.3(+8), 1586.1(+9), 1427.6(+10), 1297.9(+11), 1189.8(+12), 1098.4(+13), 1020.0(+14)
Cystatin S mono-phos. ox.	35.3	14281 ± 2 (14280.7)	1786.40(+8), 1589.70 (+9), 1429.30 (+10), 1299.50 (+11), 1191.30 (+12), 1099.70 (+13)

PROTEINS/PEPTIDES (SWISS-PROT CODE)	ELUITION TIME (MIN ± 0.5)	EXPER. (THEOR) AV. MASS (Mav)	M/Z (CHARGE)
Cystatin S di-phos. (P01036)	35.3	14346 ± 2 (14345)	1794.3(+8), 1595.0(+9), 1435.6(+10), 1305.2(+11), 1196.5(+12), 1104.5(+13), 1025.7(+14)
Cystatin SN (P01037)	34.6	14312 ± 2 (14313)	1790.0(+8), 1591.2(+9), 1432.2(+10), 1302.1(+11), 1193.7(+12), 1101.9(+13), 1023.3(+14)
Cystatin SN ox. (P01037)	34.6	14328 ± 2 (14328)	1792.30(+8), 1593.20 (+9), 1434.00 (+10), 1303.30 (+11), 1195.20 (+12), 1103.30 (+13)
Cystatin SA (P09228)	36.8	14347 ± 2 (14346)	1794.4(+8), 1595.1(+9), 1435.7(+10), 1305.3(+11), 1196.6(+12), 1104.6(+13), 1025.8(+14)
Antileukoproteinase			
Antileukoproteinase (P03973)	26.2	11702.2 ± 1 (11706)	1952.64(+6), 1673.84(+7), 1464.73(+8), 1302.10(+9)
α-Defensins			
α-defensin 1 (P59665)	23.5	3442.5 ± 2 (3442.1)	1772.03(+2), 1148.36(+3), 861.52(+4)
α-defensin 2 (P59665/6)	23.5	3370.4 ± 1 (3370.9)	1686.49(+2), 1124.66(+3), 843.75(+4)
α-defensin 3 (P59666)	23.5	3485 ± 2 (3486.1)	1744.03(+2), 1163.03(+3), 872.52(+4)
α-defensin 4 (P12838)	27.2	33708 ± 1 (3709.4)	1855.71(+2), 1237.48(+3), 928.36(+4)
β-thymosins			
Thymosin β-4 (P62328)	18.5	4944.5 ± 1 (4963)	1655.51(+3), 1241.88(+4), 993.71(+5)
Thymosin β-10 (P62313)	20.8	4934.5 ± 1 (4935)	1646.52(+3), 1235.14(+4), 988.31(+5)
Thymosin β-4 Met-ox.	18.0	4977.5 ± 1 (4979.5)	1660.84(+3), 1245.88(+4), 996.91(+5)
S100A Family			
S100A12 (P80511)	40.0	10444 ± 2 (10443.9)	1306.5(+8), 1161.4(+9), 1045.4(+10), 950.4(+11)
S100A7 (D27) (P31151[g])	37.0	11367 ± 2 (11367.8)	1422.0(+8), 1264.1(+9), 1137.8(+10), 1034.4(+11)
S100A8 (P05109)	40.4	10833 ± 2 (10834.5)	1355.3(+8), 1204.8(+9), 1084.5(+10), 985.9(+11)
S100A8-SO_2H	39.9	10866 ± 2 (10866.5)	1359.3(+8), 1208.4(+9), 1087.7(+10), 988.9(+11)
S100A8-SO_3H/Trp54ox	40.4	10898 ± 2 (10898.6)	1363.3(+8), 1212.0(+9), 1090.9(+10), 991.8(+11)
S100A8 hyperoxidized (SO_3H/ Trp54-diox; SO_3H/ Trp54-ox - Met-ox)	39.3	10915 ± 2 (10914.6)	1365.3(+8), 1213.7(+9), 1092.5(+10), 993.2(+11)
S100A8 SSG [a]	38.3	11140 ± 2 (11139.8)	1393.5(+8), 1238.8(+9), 1115.0(+10), 1013.7(+11)
S100A8 SNO [d]	40.8	10863 ± 2 (10863.5)	1358.9(+8), 1208.1(+9), 1087.3(+10), 988.6(+11)
S100A8/A9 S-S dimer [c]	41.8	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)

PROTEINS/PEPTIDES (SWISS-PROT CODE)	ELUITION TIME (MIN ± 0.5)	EXPER. (THEOR) AV. MASS (Mav)	M/Z (CHARGE)
S100A9(S) (P06702)	42.2	12690 ± 2 (12689.2)	1410.9(+9), 1269.9(+10), 1154.6(+11), 1058.4(+12), 977.1(+13)
S100A9(S) mono-phos.	42.2	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	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	12786 ± 2 (12785.2)	1421.9(+9), 1279.5(+10), 1163.3(+11), 1066.4(+12), 984.5(+13)
S100A9(L) SSG [a]	41.5	13459 ± 2 (13458.1)	1346.8(+10), 1224.5(+11), 1122.5(+12), 1036.3(+13), 962.3(+14)
S100A9(L) SSG [a] mono-phos.	41.5	13539 ± 2 (13538.1)	1354.8(+10), 1231.8(+11), 1129.2(+12), 1042.4(+13), 968.0(+14)
S100A9(L) SSG [a] Met-ox	41.3	13475 ± 2 (13474.1)	1348.4(+10), 1225.9(+11), 1123.8(+12), 1037.5(+13), 963.4(+14)
S100A9(L) SSG [a] Met-ox, mono-phos.	41.3	13555 ± 2 (13554.1)	1356.4(+10), 1233.2(+11), 1130.5(+12), 1043.6(+13), 969.1(+14)
S100A9(L) SSC [b]	41.6	13273 ± 2 (13271.9)	1328.2(+10), 1207.6(+11), 1107.0(+12), 1021.9(+13), 949.0(+14)
S100A9(L) SSC [b] mono-phos.	41.6	13353 ± 2 (13351.9)	1336.2(+10), 1214.8(+11), 1113.7(+12), 1028.1(+13), 954.7(+14)

[a], SSG: glutathionylation on cysteine residue; [b], SSC: cysteinylation on cysteine residue; [c], S-S: formation of a disulfide bond; [d], SNO: nitrosylation on cysteine residue.

Results

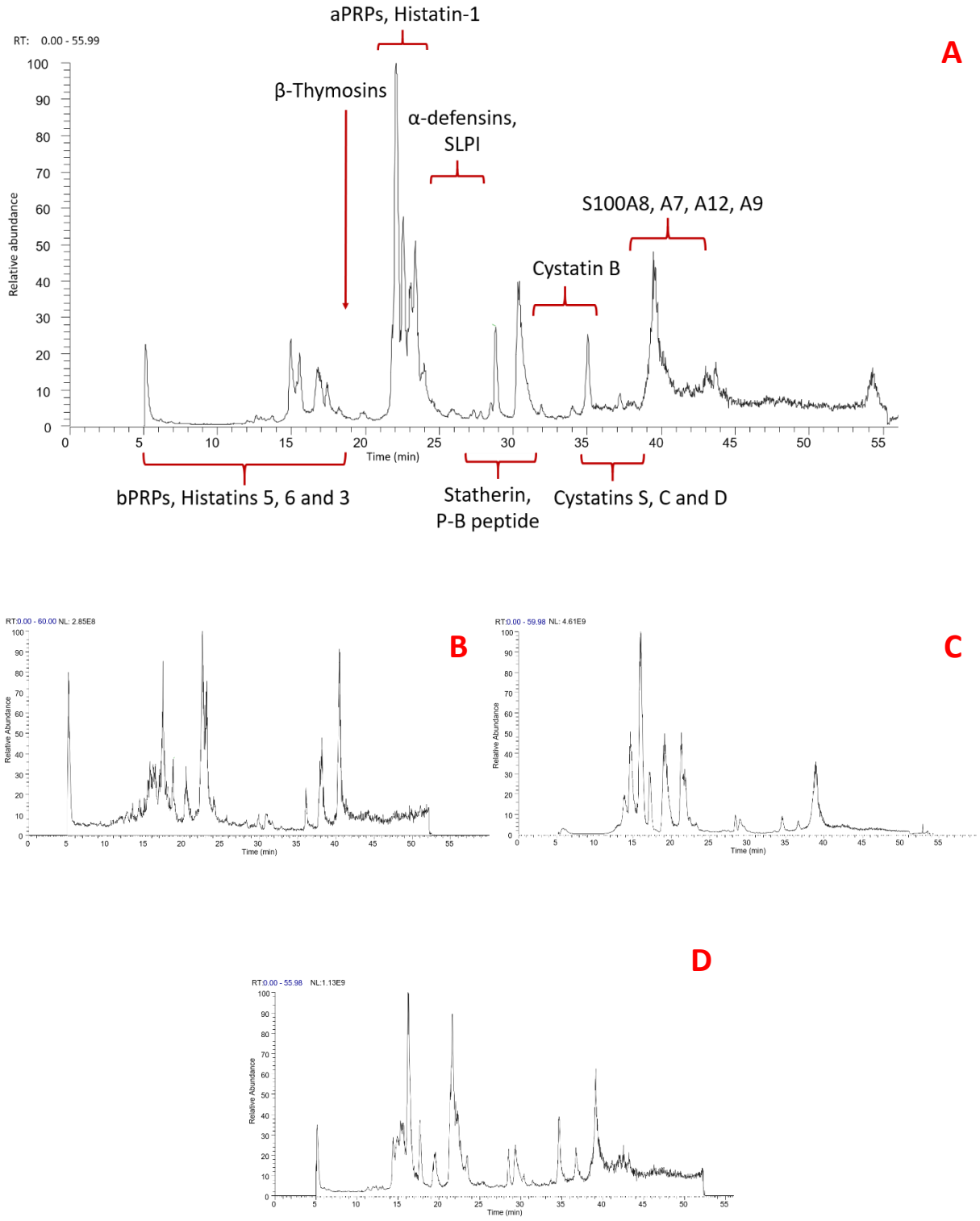
In order to analyze the salivary samples by RP-HPLC-ESI-MS, whole saliva collected from healthy subjects and patients was treated with an acid solution contained 0.2% of TFA. This treatment takes two advantages[100]: 1) the acidic pH of the solution blocks the exogenous and endogenous proteases naturally present in saliva and in this way prevents the degradation of proteins and peptides; 2)TFA is one of the best ion-paring HPLC agents because of its chromatographic performance, UV transparency and ability to dissolve high-molecular weight peptides and proteins. It is relevant to underline that the obliged use of acidic conditions causes the precipitation of high molecular weight salivary proteins, like immunoglobulin and mucins, thus they are excluded from the analyses by these experimental conditions.

However, RP-HPLC-ESI-MS top-down analysis of the acid-soluble fraction of human saliva allowed simultaneously detection of all the soluble proteins/peptides and their proteolytic products, as well as performing a comparative label-free quantification of these components in different sets of samples[75].

The following proteins and peptides were searched and quantified by XIC procedure, as reported in Table 5: histatins (His) 1, 3, His-3 1/24 and His-3 1/25, salivary Cystatins S, S1, S2, SN and SA, cystatin A, B, C and D, statherin, P-B peptide, acid proline-rich proteins (aPRPs), α-defensins 1–4, β-thymosin 4 and 10, antileucoproteinase (SLPI), S100A7 (D27 variant), S100A12, S100A8, S100A9 (short (S) and long (L) proteoforms. Moreover, proteoforms derived from PTMs of these components were searched and quantified: phosphorylated, N-terminal acetylated, oxidized in Met and Trp residues, disulfide dimers (S-S), Cys-glutathionylated (SSG), Cys-cystainylated (SSC), Cys-nitrosylated (SNO), Cys-sulfonic (-SO_3H) and sulfinic acid (SO_2H). bPRPs were not evaluated due to their high variability linked to the physiological status[106].

Figure 4 shows the typical total ion current chromatographic (TIC) profile obtained by HPLC-low-resolution MS analysis of the acid soluble fraction of a salivary samples from a healthy subject (panel A), a patient with cutaneous mastocytosis (CM – panel B), systemic mastocytosis (SM-C – panel C) and systemic mastocytosis with cutaneous lesions (SM+C – panel D). Panel A shows also the elution ranges of the families of proteins and peptides that have been analyzed in salivary samples.

Figure 4 : TIC profile of acidic-soluble fraction of saliva from healty control (panel A), CM patient (panel B), SM-C patient (panel C) and SM+C patient (panel D) obtained by RP-HPLC-ESI-low-resolution MS analysis.



A great advantage of top-down LC-ESI-MS platforms is the possibility to carry out label-free comparative quantitation without any limitation on the number of specimens and subjects under study[192]. Thought the XIC procedure it is possible to extract from the TIC an ion current peak specific of a protein and quantified it by integration of the peak area. This value, which is proportional to the concentration, in constant experimental conditions may be used to quantitatively compare the same protein in different samples[100,191]. Figure 5 shows an example of quantification with XIC procedure. Panel A displays the total ion current of a salivary sample from a healthy person. In panel B is shown the XIC peak of di-phosphorylated statherin generated by the selection of the multiply-charged ions selectively extracted for this peptide. Panel C exhibits mass spectrum (*m/z*) recorded in elution time of XIC peak of di-phosphorylated statherin. In panel D is reported deconvoluted mass spectrum with the Mav of di-phosphorylated statherin.

The patients were divided in two different groups based on the clinical classification, CM and SM. Table 6 summarizes the results of the quantitative analysis carried out comparing XIC peak areas of proteins and peptides measured in the two mastocytosis patients' groups (CM and SM) and healthy controls (Ctrls). The statistical comparison between the XIC peak area values of proteins/peptides measured in CM, SM and controls groups provided the results resumed in Table 6. In the case of cystatins B, C, S1, S2, SN, and S100A8 all the proteoform of a same protein showed the same trend, thus the sum of their XIC peak areas was reported in the Table 3. The patients were also split considering *c-KIT* gene mutation in two groups, patients with and without mutation, and compared each other and with the control group, but the results did not show any difference statistically considerable.

Figure 5: XIC procedure on di-phosphorylated statherin.

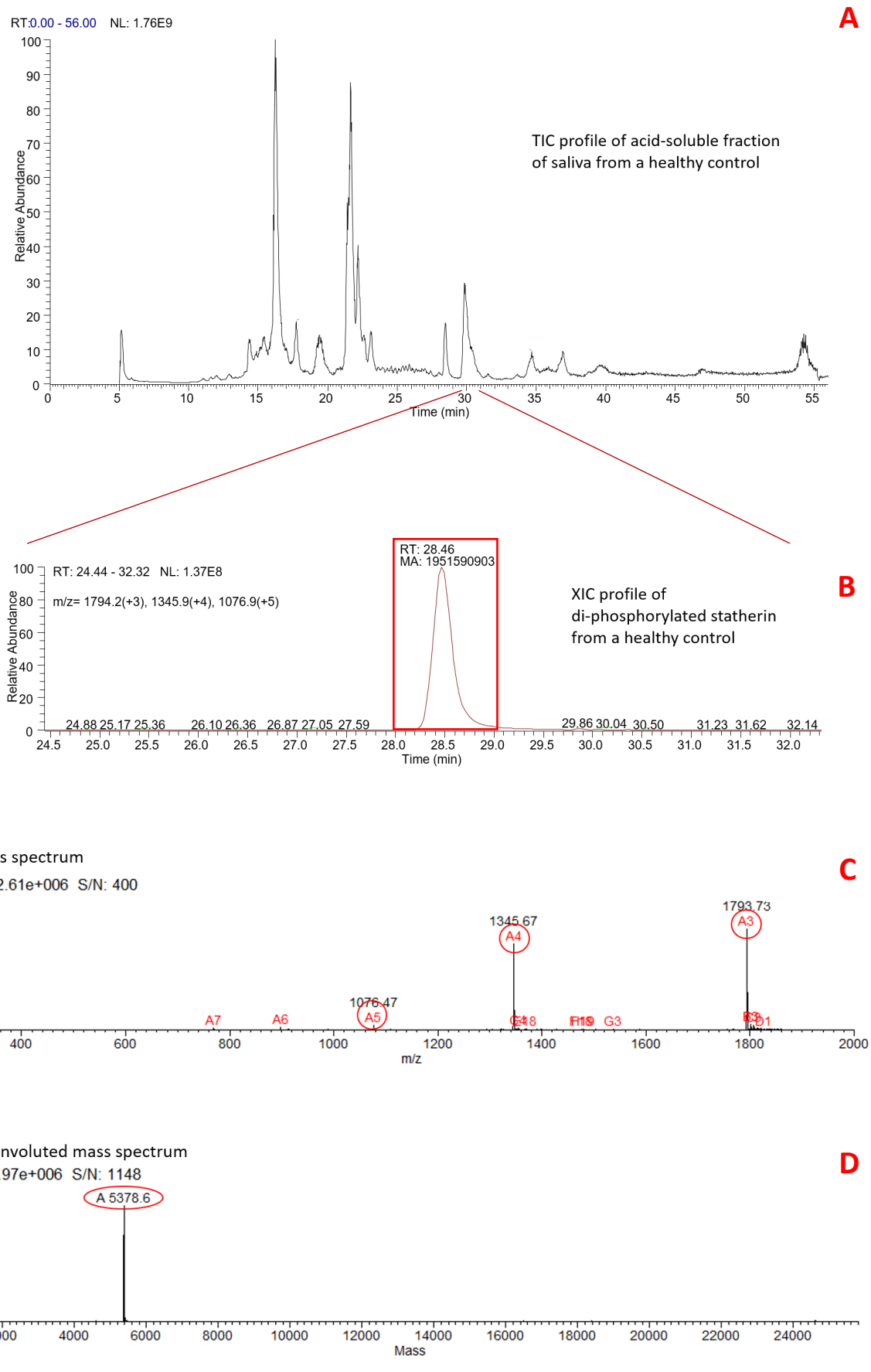


Table 6: XIC peak areas values (mean ± SD), frequencies and p-values obtained by statistical analysis by comparing the two patients' groups with respect to the controls. p-values >0.05 are not statistically significant (•). In some cases, no statistical comparison was possible to carry out because of the absence of protein in one or more groups (NA). Data are reported in bold when frequencies are <2 for group. The numbers in column 1 correspond to those one reported in Figure 8 described the Volcano test results.

	PROTEIN/ PEPTIDE	**XIC PEAK AREA X 10^8 (MEAN ± SD) AND FREQUENCY**			*p-values*	
		CM	SM	Ctrls	CM vs Ctrls	SM vs Ctrls
#1	PRP-1 diphos.	25.6 ± 19.7 (6/6)	28.2 ± 27.8 (35/35)	99.9 ± 70.9 (48/48)	0.001 ↓CM	< 0.0001 ↓SM
#2	PRP-1 monophos.	2.5 ± 3.1 (5/6)	3.7 ± 5.2 (29/35)	12.8 ± 12.9 (47/48)	0.002 ↓CM	< 0.0001 ↓SM
#3	PRP-1 nonphos.	**0 (0/6)**	0.09 ± 0.3 (6/35)	0.5 ± 0.9 (22/48)	**NA**	0.003 ↓SM
#4	PRP-1 triphos.	0.2 ± 0.2 (4/6)	0.4 ± 0.6 (22/35)	2.0 ± 1.9 (44/48)	0.002 ↓CM	< 0.0001 ↓SM
#5	PRP-3 diphos.	5.9 ± 4.4 (6/6)	9.5 ± 11.3 (33/35)	35.1 ± 29.3 (48/48)	0.0002 ↓CM	< 0.0001 ↓SM
#6	PRP-3 monophos.	1.1 ± 1.0 (6/6)	1.5 ± 1.8 (29/35)	4.9 ± 4.1 (47/48)	0.002 ↓CM	< 0.0001 ↓SM
#7	PRP-3 nonphos.	**0 (0/10)**	0.004 ± 0.02 (2/35)	0.4 ± 1.1 (18/48)	**NA**	0.0003 ↓SM
#8	PRP-3 diphos, desArg106	3.1 ± 2.7 (6/6)	2.8 ± 5.9 (26/35)	3.8 ± 6.2 (40/48)	•	•
#9	P-C peptide	4.1 ± 4.4 (6/6)	7.0 ± 10.0 (33/35)	20.4 ± 16.2 (48/48)	0.001 ↓CM	< 0.0001 ↓SM
#10	Statherin diphos.	5.4 ± 6.6 (5/6)	6.0 ± 6.6 (34/35)	13.4 ± 10.1 (47/47)	0.02 ↓CM	0.0001 ↓SM
#11	Statherin monophos.	0.2 ± 0.2 (3/6)	0.1 ± 0.2 (17/35)	0.3 ± 0.4 (32/47)	•	0.03 ↓SM
#12	Statherin SV-1 (des-F43)	1.1 ± 1.0 (6/6)	1.0 ± 0.9 (34/35)	1.9 ± 1.8 (46/47)	•	0.006 ↓SM
#13	Statherin desT42-F43	0.7 ± 0.4 (6/6)	0.4 ± 0.4 (25/35)	0.7 ± 0.7 (46/47)	•	•
#14	Statherin desAsp1	0.7 ± 0.9 (5/6)	0.5 ± 0.5 (31/35)	0.8 ± 0.8 (46/47)	•	0.02 ↓SM
#15	Statherin des1-9	0.1 ± 0.2 (5/6)	0.2 ± 0.4 (18/35)	0.7 ± 0.7 (40/47)	•	< 0.0001 ↓SM
#16	Statherin des1-10	0.2 ± 0.2 (5/6)	0.3 ± 0.4 (24/35)	0.5 ± 0.5 (44/47)	•	0.001 ↓SM
#17	Statherin des1-13	0.1 ± 0.1 (4/6)	0.1 ± 0.2 (22/35)	0.3 ± 0.2 (45/47)	0.04 ↓CM	< 0.0001 ↓SM
#18	P-B peptide	7.6 ± 7.9 (6/6)	9.2 ± 7.8 (34/35)	21.9 ± 14.3 (47/47)	0.004 ↓CM	< 0.0001 ↓SM
#19	P-B des1-4	0.9 ± 0.6 (6/6)	1.0 ± 0.8 (31/35)	1.5 ± 1.7 (39/48)	•	•
#20	P-B des1-5	3.0 ± 3.6 (6/6)	1.8 ± 2.4 (30/35)	2.3 ± 2.8 (42/48)	•	•

PROTEIN/ PEPTIDE		XIC PEAK AREA X 10^8 (MEAN ± SD) AND FREQUENCY			*p-values*	
		CM	SM	Ctrls	CM vs Ctrls	SM vs Ctrls
#21	P-B des1-7	1.7 ± 0.9 (6/6)	1.7 ± 1.5 (32/35)	4.4 ± 2.8 (47/48)	0.007 ↓CM	< 0.0001 ↓SM
#22	P-B des1-12	0.8 ± 0.4 (6/6)	0.6 ± 0.6 (31/35)	1.0 ± 1.4 (46/48)	•	•
#23	Histatin-1 monophos.	1.1 ± 1.6 (4/6)	1.1 ± 1.5 (24/35)	2.9 ± 2.9 (34/48)	•	0.02 ↓SM
#24	Histatin-1 nonphos.	**0 (0/6)**	0.2 ± 0.3 (15/35)	0.4 ± 0.7 (23/48)	**NA**	•
#25	Histatin-3	**0.01 ± 0.03 (1/6)**	0.3 ± 0.6 (12/35)	1.2 ± 1.7 (26/48)	**NA**	0.008 ↓SM
#26	Histatin-3 1/25	**0 (0/6)**	0.04 ± 0.1 (6/35)	0.9 ± 1.2 (25/48)	**NA**	0.0001 ↓SM
#27	Histatin-3 1/24	**0 (0/6)**	0.8 ± 1.7 (15/27)	3.0 ± 3.4 (40/48)	**NA**	< 0.0001 ↓SM
#28	Cystatin A	2.0 ± 1.4 (6/6)	1.2 ± 1.2 (35/35)	2.2 ± 1.9 (47/48)	•	0.001 ↓SM
#29	Cystatin B (all proteoforms)	1.3 ± 1.1 (6/6)	1.2 ± 1.2 (34/35)	1.4 ± 1.3 (43/48)	•	•
#30	Cystatin C (all proteoforms)	**0.03 ± 0.07 (1/6)**	0.05 ± 0.2 (3/35)	0.6 ± 0.7 (32/48)	**NA**	< 0.0001 ↓SM
#31	Cystatin D-R_{26} des1-5	**0.4 ± 0.9 (1/6)**	0.5 ± 0.8 (19/35)	0.8 ± 1.1 (28/48)	**NA**	•
#32	Cystatin D-R_{26} des1-8	**0 (0/6)**	0.06 ± 0.1 (11/35)	0.1 ± 0.2 (15/48)	**NA**	•
#33	Cystatin S	**0 (0/6)**	0.3 ± 0.7 (11/35)	2.5 ± 0.5 (29/48)	**NA**	0.0007 ↓SM
#34	Cystatin S1 (all proteoforms)	11.3 ± 12.7 (6/6)	4.5 ± 3.8 (33/35)	10.3 ± 10.6 (45/48)	•	0.002 ↓SM
#35	Cystatin S2 (all proteoforms)	1.9 ± 3.0 (2/6)	2.2 ± 3.6 (22/35)	3.8 ± 3.9 (45/48)	•	0.002 ↓SM
#36	Cystatin SN (all proteoforms)	10.2 ± 13.8 (4/6)	9.2 ± 11.1 (31/35)	20.7 ± 17.1 (47/48)	•	0.0005 ↓SM
#37	Cystatin SA	0.9 ± 1.9 (2/6)	0.8 ± 2.1 (9/35)	2.9 ± 3.8 (35/48)	•	<0.0001 ↓SM
#38	Thymosin β-4	0.4 ± 0.3 (5/6)	0.4 ± 0.5 (24/35)	0.8 ± 1.0 (30/48)	•	•
#39	α-defensin 1	1.1 ± 0.7 (6/6)	1.4 ± 1.9 (33/35)	1.9 ± 2.9 (41/48)	•	•
#40	α-defensin 2	0.8 ± 0.6 (6/6)	0.7 ± 0.9 (31/35)	1.3 ± 1.7 (37/48)	•	•
#41	α-defensin 3	0.6 ± 0.7 (5/6)	0.3 ± 0.4 (21/35)	0.8 ± 1.3 (30/48)	•	•
#42	α-defensin 4	0.2 ± 0.1 (3/6)	0.1 ± 0.2 (15/35)	0.3 ± 0.4 (23/48)	•	•
#43	S100A9(S) + (S) M-ox	1.2 ± 1.6 (4/6)	0.6 ± 1.0 (20/35)	2.4 ± 2.6 (42/48)	•	<0.0001 ↓SM
#44	S100A9 (S) monophos + (S)monophos./M-ox	0.3 ± 0.6 (2/6)	0.6 ± 1.1 (20/35)	0.6 ± 1.0 (17/48)	•	•
#45	S100A9(L) SSG	0.6 ± 1.1 (2/6)	0.7 ± 1.3 (14/35)	0.7 ± 1.5 (22/48)	•	•

PROTEIN/ PEPTIDE		XIC PEAK AREA X 10^8 (MEAN ± SD) AND FREQUENCY			*p-values*	
		CM	SM	Ctrls	CM vs Ctrls	SM vs Ctrls
#46	SLPI	0.2 ± 0.3 (3/6)	0.2 ± 0.8 (18/35)	0.1 ± 0.20 (10/48)	•	0.005 ↑SM
#47	S100A8 (all proteoforms)	**0.5 ± 1.1 (1/6)**	0.6 ± 1.1 (14/35)	0.2 ± 0.6 (9/48)	**NA**	0.02 ↑SM
#48	S100A8-SNO	**0.5 ± 1.1 (1/6)**	0.3 ± 0.7 (8/35)	0.1 ± 0.5 (6/48)	**NA**	•

Quantitative comparison between controls and patients

All the aPRP proteoforms, with the exception of PRP-3 desArg106, showed a significant lower abundance in all the patient groups with respect to the controls. Among the investigated statherin proteoforms, only statherin desT42-F43 did not show variations between the three groups, whereas, the di- and mono-phosphorylated forms, and the SV-1, desAsp1, des1-9, des1-10 and des1-13 truncated forms were significantly less concentrated in saliva of SM patients than in controls. The statherin des1-13 resulted less concentrated also in the CM group with respect to healthy subjects. The same results were obtained for P-B peptide and it truncated form P-B des1-7.

Significant decreased levels of histatins in both SM and CM patients with respect to the controls were also observed. Histatin 1 non-phosphorylated, such as Histatin 3 and their most abundant fragments, Hst-3 1/24 and 1/25, also named Hst-6 and Hst-5[116] were almost undetectable in CM patients (Table 6). Cystatin C was usually detected in 32 out of 48 healthy controls but sporadically in the patients: in 3 out of 35 SM patients, and in 1 out of 6 CM patients. In this last case, only the Met-oxidized derivative was observed. Obviously, for these proteins/peptides the statistical comparison between the CM and the other groups was inapplicable. As it concerns the proteoforms of S100A9 detectable in human adult saliva (Table 5), only the not phosphorylated short forms of S100A9, called S100A9(S), not-oxidized and mono-oxidized on a methionine residue (M59 or M77 or M79, or M90), exhibited a significant lower level in SM patients than in controls (Table 6). Similar results were obtained for the S-type cystatins, S1, S2, SN and SA.

Two truncated proteoforms of the variant R26 of cystatin D lacking the first five or eight residues and with N-terminal pyroglutamination, called cystatin D-R26 des1-5 and the des1-8[125], were sporadically detected, the latter showing a lower frequency and being always present in the samples containing the other truncated form. In CM patients they were undetectable with the exception of one subject showing only the des1-5 proteoform. Due to the high dispersion of the XIC peak area values of both cystatin D proteoforms in the SM and controls groups, it was not possible to highlight significant differences. The lowest levels of cystatin A was found in the SM patients, while CM patients exhibited levels comparable with controls.

S100A8 was sporadically detected as unmodified proteoform: in 4 SM patients, in 4 controls, never in CM patients. Analogously, the hyperoxidized form of S100A8 was found just in 2 SM patients and in 3 healthy subjects, at very low levels. Instead, nitrosylated S100A8 was found with higher frequency: in 1 CM patient, 8 SM patients and in 3 controls. By considering the total S100A8 the statistical comparison between the three groups highlighted a significant higher abundance in SM patients with respect to controls. Similar results were obtained for the antileukoproteinase (SLPI). SLPI has been found in 10 out of 48 healthy controls, whereas the patient groups showed a higher frequency, 18 out of 35 SM, and 3 out of 6 CM patients (Table 6). The SM group exhibited a significant higher abundance of SLPI than healthy controls, the significant difference was confirmed also by excluding the two outlier points that belonged to SM patients without cutaneous manifestations (p value = 0.01).

Differentiation between subgroups of SM patients

8 out of the 35 SM patients manifested only systemic symptoms, differently from the remaining 27 patients with additional cutaneous symptoms. We compared these two subgroups, called SM-C and SM+C, and identified some significant differences. The di-phosphorylated PRP-3, its C-terminal truncated form (desArg106), P-C peptide, the fragment statherin SV-1 and the fragment P-B des1-5 were more abundant in SM-C than in SM+C patients (Figure 6, panels A-E). Moreover, both cystatin D-R26 des1-5 and des1-8 resulted more abundant in SM-C patients than in those ones with cutaneous symptoms, both SM+C and CM (Fig. 7, panels A-B). The levels of α-defensins 1 and 4, and Tβ-4 were significantly different between the SM-C and SM+C groups. Indeed, the SM patients without cutaneous symptoms exhibited lower levels of these peptides with respect to the SM+C patients. α-defensin 4 has never been detected in samples from SM-C patients, whereas Tβ-4 was detected in only 3 out of 8 SM-C patients (Fig. 7 panels C-E). The significant high level of total S100A8 in SM patients was a consequence of the higher abundance of S100A8-SNO in the subgroup of SM+C with respect to the Ctrls (p = 0.03); this proteoform of S100A8 was detected just in 1 SM-C patient. Some of the differences highlighted by statistical test based on two-group comparison, Mann-Whitney or Welch-corrected t-test, were confirmed by a Volcano plot test (Figure 8). In the plot each number corresponds to a protein/peptide as listed in Table 6. The test evidenced different salivary protein profiles of the healthy controls and the groups of patients, CM, SM+C and SM-C (Figure 8 panel A-C). Higher levels of di-phosphorylated PRP-

3 and lower levels of cystatin D-R26 des1-5 in CM patients differentiated them from the SM-C (Figure 8 panel D), while higher levels of PRP-3 desArg106 in CM were distinctive with respect to the SM+C patients (Figure 8 panel E). SM+C differentiated to SM-C patients for the higher levels of α-defensin 1 and Tβ-4, and lower levels of di-phosphorylated PRP-3 and PRP-3 desArg106, P-C peptide, statherin SV-1, P-B peptide des1-5 and cystatin D-R26 des1-5 and des1-8.

Correlation analysis with tryptase

The XIC peak area of the proteins/peptides analyzed in mastocytosis patients were subjected to a correlation analysis with tryptase measured in the same patients. Tryptase is an enzyme secreted by MCs whose concentration in the blood, expressed in μg/L, is considered an indicator of the amount and the activity of MCs and may be a marker of MCs-mediated diseases. CM patients showed a lower tryptase concentration (6.3 ± 1.9μg/L) than both groups of SM patients, where tryptase corresponded to 28.3 ±23.1, and 28.2 ± 19.1 μg/L, in SM-C and SM+C, respectively. The analysis performed by considering the entire SM group resulted in a significant positive correlation of statherin des1-10, and des1-13, P-B des1-4, and Cystatin D-R26 des1-5 (Table 7). Different results were obtained by considering the two SM subgroups, SM+C and SM-C. Indeed, not any correlation was obtained in the SM+C group, whereas in the SM-C group the levels of mono-phosphorylated PRP-1 and PRP-3, PRP-3 desArg106, statherin truncated forms, desAsp1, des1-10 and des1-13, all the four truncated forms of P-B peptide, and cystatin D-R26 des1-5 (Figure 9) correlated positively with tryptase concentration, whereas a negative correlation was found with Tβ-4 level (Figure 9).

Figure 6: XIC peak area distribution of PRP-3 di-phosphorylated, PRP-3 desArg106, P-C peptide, statherin SV-1, P-B des1-5 in the three groups of patients. *** = p<0.001, ** = p<0.01, * = p <0.05.

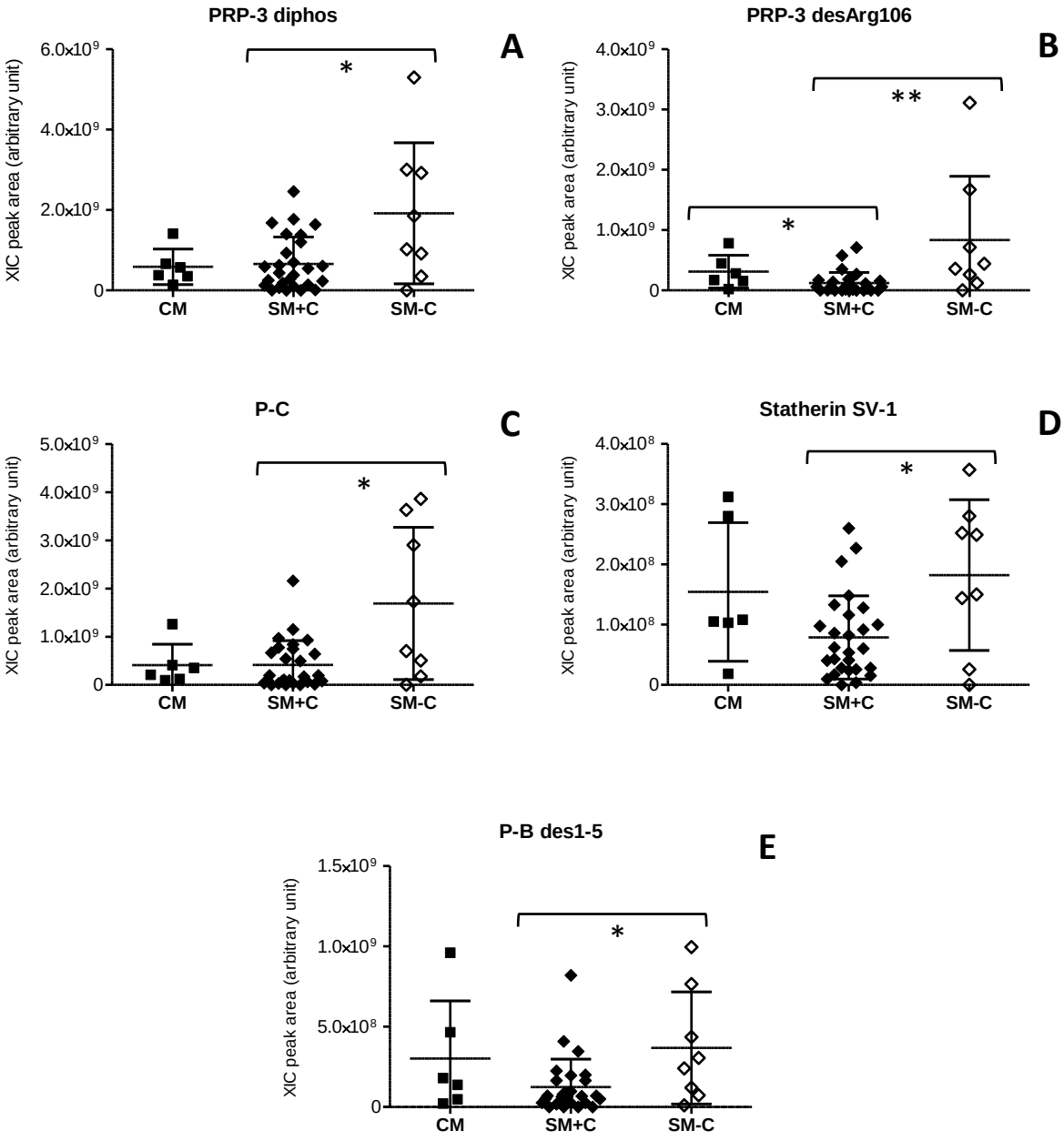


Figure 7: XIC peak area distribution of cystatin D-R26 des1-5 and des1-8, α-defensin 1 and 4, and thymosin β-4 in the three groups. *** = p<0.001, ** = p<0.01, * = p <0.05.

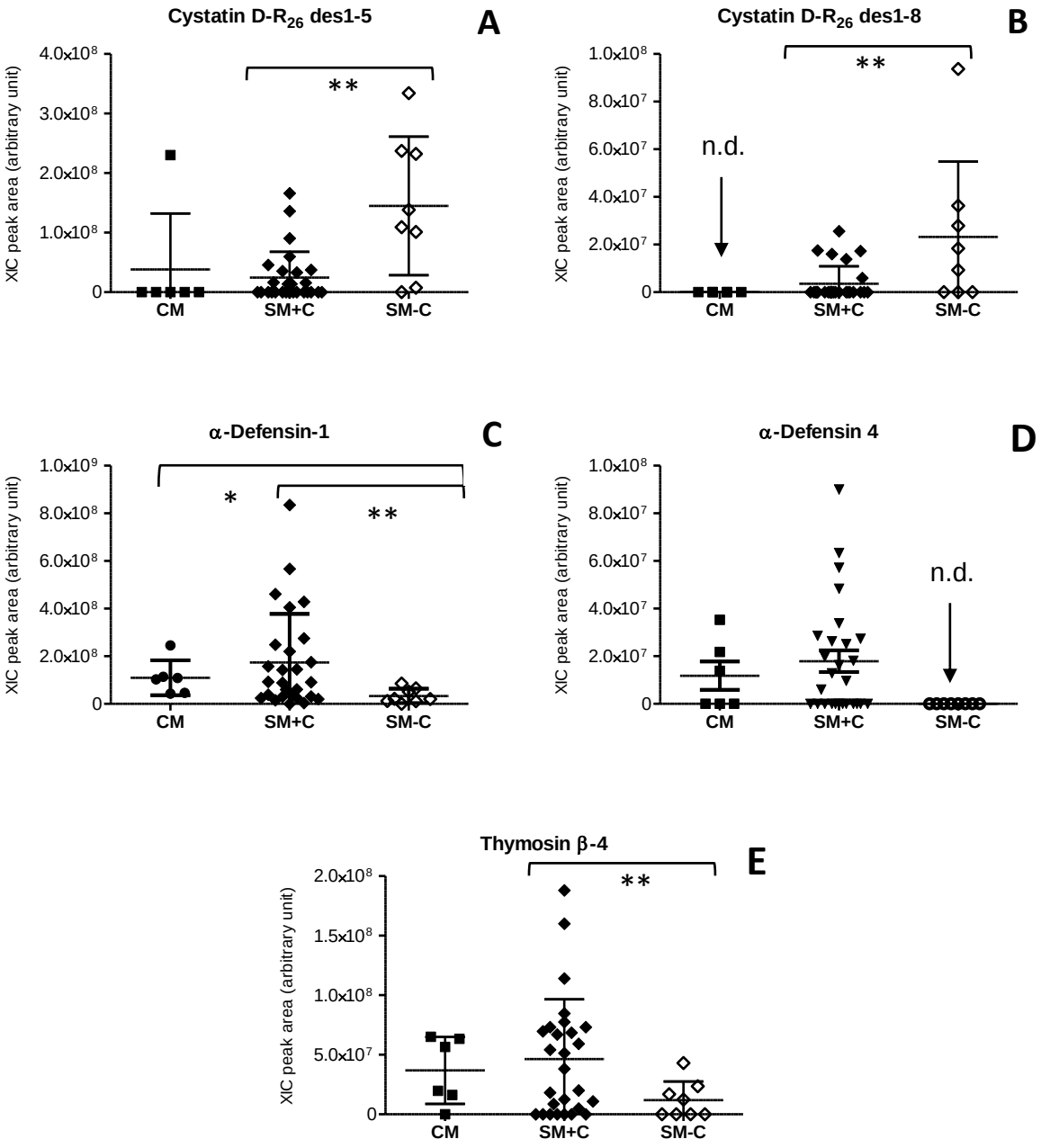


Figure 8: Volcano plot images from patient/control (panels A, B and C) and patients/patients (panels D, E and F) ratios. Points above horizonal line represent proteins with p-values <0.05; point to the left and to the right of the two vertical lines represent fold changes of mean values of XIC peak area less and more than 1, respectively. Black circles indicate proteins/peptides with lower level in the different patients compared to controls or among the patients, white circles indicate those with higher levels. Labels on the dots are referred to protein numbers reported in Table 6.

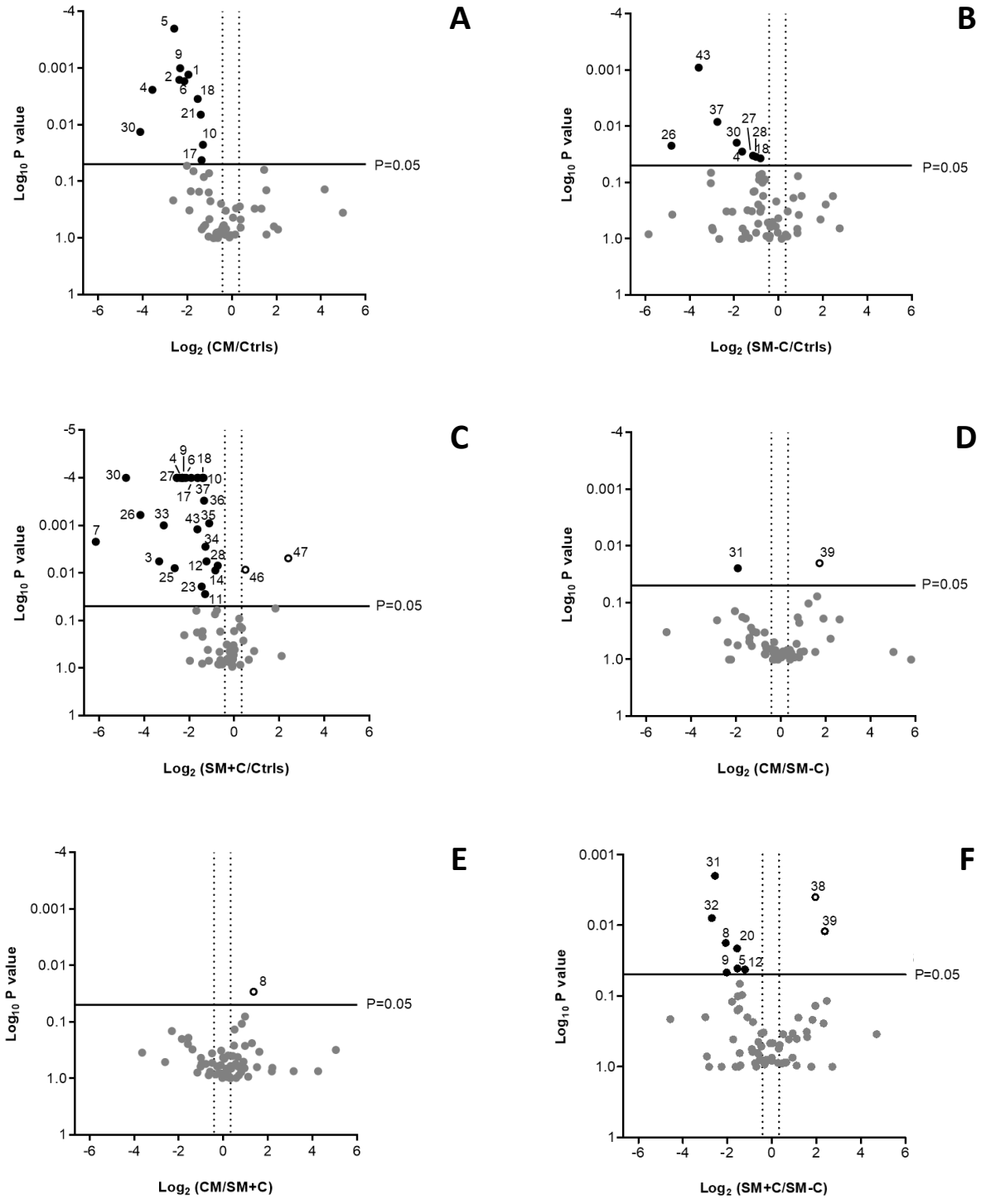
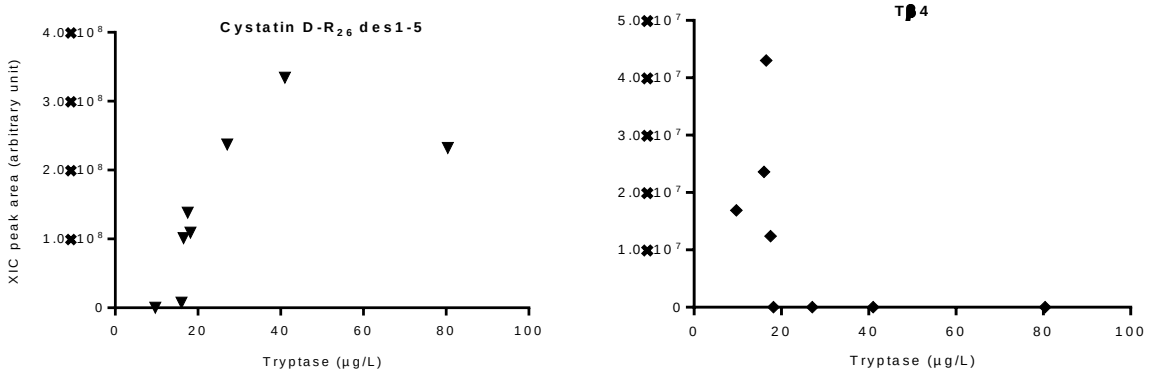


Table 7: Results of correlation analysis between salivary proteins/peptides and tryptase levels in SM patients and in the subgroup SM-C. In the Table is indicated the p-value and the Spearman R coefficient. The significant difference between SM-C and SM+C are indicated (Y).

Protein/Peptide	SM (n 35)		SM-C (n 8)		Distinctive between SM-C and SM+C
	p-value	*R*	p-value	*R*	
Prp-1 monophos.	-	-	0.04 (+)	*0.77*	
Prp-3 monophos.	-	-	0.03 (+)	*0.77*	
Prp-3 desArg106	-	-	0.007 (+)	*0.88*	Y (↑ in SM-C)
Statherin desAsp1	-	-	0.01 (+)	*0.86*	
Statherin des1-10	0.03 (+)	*0.40*	0.02 (+)	*0.79*	
Statherin des1-13	0.02 (+)	*0.40*	0.02 (+)	*0.79*	
P-B des1-4	0.02 (+)	*0.38*	0.002 (+)	*0.93*	
P-B des1-5	-	-	0.02 (+)	*0.81*	Y (↑ in SM-C)
P-B des1-7	-	-	0.01 (+)	*0.86*	
P-B des1-12	-	-	0.006 (+)	*0.89*	
Cystatin D-R_{26} des1-5	0.02 (+)	*0.40*	0.005 (+)	*0.90*	Y (↑ in SM-C)
Thymosin β-4	-	-	0.005 (-)	*-0.84*	Y (↓ in SM-C)

Figure 9: Correlation graphs of cystatin D-R26 des1-5 and thymosin β-4.



Discussion

The first part of the thesis was focused on the analysis and comparison of the protein profiles detectable by HPLC-ESI-IT-MS. The aim of this study was to investigate about possible significant variation of the salivary proteome of patients affected by mastocytosis with respect to a gender- and age-matched healthy control group valuable as potential clues of the disease. In order to analyze the intact peptides/proteins present in the salivary samples, we analyzed the protein fraction soluble in acid solution by HPLC-ESI-IT-MS. It is noticeable to underline that the lower number of patients in the groups with cutaneous mastocytosis, which is a rare disease variant in adults, has often made difficult to perform statistical analyzed, and thus the results present in this thesis are to be considered preliminary and need to be confirmed after collection of new cohort of patients. However, despite the limitation indicated, the study of salivary proteome of patients affected by mastocytosis showed important quantitative differences between patients and controls and between the three subgroups of patients. Mastocytosis is characterized by infiltration of clonally derived MCs in one or more different organs, which define the different variants: CM, in which only skin is involved, SM-C, in which is involved at least one extracutaneous organ, and SM+C, a systemic form in which are present also skin lesions[15]. Mastocytosis' diagnosis is based on clinical manifestation, histopathology analysis and laboratory evaluation. In particular, CM diagnosis is obtained with visual analysis of lesions followed by lesioned skin biopsy, while SM one is more complicated, and it is based on bone marrow biopsy, evaluation of serum tryptase level and genetic analysis to look for c-KIT receptor mutations[43]. This receptor seems to play an important role in SM onset; in fact, mutations in juxta-membrane and kinase domains lead to uncontrolled proliferation and survival of MCs, which will accumulate in different organs. Since to the difficulty diagnosis and the distinction of the different forms of mastocytosis, and since the invasiveness of the samples' collections[41], the characterization and validation of new disease-specific biomarkers is needed.

The salivary profiling of the patients showed significant low abundance of aPRPs, statherin, P-B peptide, and histatins, which are specific of the oral cavity and are secreted by salivary glands. The levels of some naturally occurring fragments of aPRPs, statherin, and P-B peptide, allowed us to distinguish, into the SM group, the patients with only systemic symptoms, SM-C, from those ones with additional cutaneous symptoms, SM+C, which showed levels

significantly lower of PRP-3 and PRP-3 desArg106, P-C peptide, statherin SV-1, also called desF43, and P-B peptide des1-5. Apart from for PRP-3 and P-C peptide, which origin from a pre-secretory proteolytic event occurring on PRP-1 proteoform, the others probably origin from post-secretory proteolysis by action of several proteases acting in the glandular ducts or in the oral cavity[100]. These salivary proteins and peptides are involved in the homeostasis and in the protection of the oral cavity, with the exception of P-B peptide, whose biological role is still undefined; consequently, their down-regulation can result in impaired homeostasis control of the oral cavity in mastocytosis patients. Indeed, statherin and aPRPs, participate to the formation of the protein pellicle covering the oral surfaces, important to modulate the colonization of microbial hosts, and to protect the enamel and the mucosa epithelium[193,194]. Histatin-1 plays an important role in the wound healing of the oral mucosa[114] and in the modulation of cell migration[115]. Moreover, mastocytosis patients could show a frailer defense against oral infections, because of the very low levels, especially in CM patients, of histatin-3 and their truncated forms histatin-3 1/24, and 1/25. They are peptides with powerful antibacterial and antifungal activities, belonging to the host defense antimicrobial peptides (AMPs), and thus key components of the innate immune system[195]. Histatin-3 1/25 is the stronger salivary peptide with anti-Candida albicans[196] and anti-ESCAPE pathogen activity[197]. All the cystatins detectable in saliva, but not cystatin B, showed a low abundance in saliva of the patients, S-type cystatins and cystatin A were down-regulated especially in SM patients, cystatin D-R26 was down-regulated in all the patients especially in CM patients, cystatin C was almost absent. Cystatins are important inhibitors of endogenous and exogenous proteases, and are involved in the inflammatory processes and in the innate immune response[198]; S-type cystatins play this function in the oral cavity, suppressing some viral, bacteria, and fungal infections of the oral cavity by inhibiting exogenous cysteine proteinases[118,144,199]. All the cystatins are able to regulate the cathepsin activity. Cystatin C is able to inhibit lysosomal cathepsins B, H, K, L, and S[198,200] in various tissues and bodily fluids[201]. A role in the modulation of immune cells was demonstrated due to its ability to inhibit migration of monocytes and T cells[202]. Cystatin D acts as inhibitor of the cathepsins B, H, L, and S[203]. Cystatin A is an inhibitor of cathepsins B, L and H acting mostly in epidermises, lymphoid tissue and oral mucosa[118,204]. S-type cystatins act as inhibitors of cathepsin C[204]. The anti-cathepsin activity of SN is implicated in the destruction of periodontal tissues[205], and associated to the processes of tumor invasion[206]. It is noticeable to underline that activated mast cells release cathepsins, in

particular cathepsin C[207]. An impaired cathepsin regulation in mastocytosis patients might led to a major inclination through oral inflammation and mucosal lesions[208–211]. Moreover, because of the salivary cystatins can be expressed in several tissues and secreted in other biological fluids[118], the condition evidenced in saliva might reflect a systemic state, where the altered suppression of cathepsin activity result in a disproportionate auto-inflammatory and allergic responses.

The SM patients exhibited, in addition, very low levels of some proteoforms of S100A9, specifically, the two not phosphorylated S100A9 short forms with and without methionine oxidation. S100A9, together S100A8, are the most abundant neutrophils proteins[167,212], but their expression can be induced in several cell types in acute and chronic inflammation, and oxidative stress conditions[212,213]. S100A9 is able to induce the chemotaxis and the activity of the neutrophils, to promote chemokine/cytokine production by leukocytes[214]. Sroussi et al. reported that expression of S100A9 may inhibit leukocyte recruitment in the healthy mucosal tissues, and that oxidized S100A9 abolished the chemo-repulsive effect on peripheral neutrophils, suggesting that the oxidation of S100A9 may serve as a molecular switch for oxidative control of inflammation[177]. On the basis of our results we could not postulate a predominant anti-inflammatory or pro-inflammatory role of S100A9(S) in the patients enrolled in the study, because of the oxidized and the not oxidized proteoforms showed the same variation trend; however, the results are suggestive for an impairment of the activity of these S100A9 proteoforms.

Interesting results have been found about thymosin β-4, and α-defensins 1 and 4 when the SM patients were divided on the base of the presence of additional cutaneous symptoms. Indeed, the lowest levels of these peptides was measured in SM-C patients. We did not observe variations for the α-defensins 2 and 3, even if α-defensins 2 derives from a proteolytic cleavage of both α-defensins 1 and 3[215]. Probably the loss of α-defensins 2 deriving from α-defensins 1 was compensated by α-defensins 2 deriving from α-defensins 3. As well as histatin 3, and their fragments, α-defensins are antimicrobial peptides belonging to the AMPs[195], and are implicated in host defense and homeostasis of tissues and biological fluids by recruiting immune cells in the infection site[216]. It is interesting the double role of α-defensins that act as antimicrobial peptides and also as modulators of inflammatory response through regulation of the cytokine production[149,217]. Their decreased levels in SM-C patients, which added to the down-regulation of other salivary AMPs, could contribute to weaken the innate system

defense of this subjects.

Tβ-4 stimulates cell growth and proliferation, especially for lymphocytes T[218], and acts as antibacterial, anti-inflammatory and antiapoptotic agent in saliva[162]. A study performed on mast cells in vitro demonstrated that Tβ-4 stimulated mast cells to release mediators involved in angiogenesis and wound healing by a process that possibly involved the actin-binding motif[219]. An immunohistochemical study proposed that Tβ-4 could be considered as a marker mast cells showing a degree of immunoreactivity comparable, in sensitivity, to chymase and tryptase[220]. Indeed, it was demonstrated that mast cells infiltrating normal dermal and mucosal tissues, such as the tumoral ones, exhibit strong expression of Tβ-4[220]. Our data did not provide information on the Tβ-4 expression into mast cells of the patients but demonstrated a less release of Tβ-4 in saliva of patients with the SM-C variant. Despite the low number of samples, these results were suggestive for a possible use of salivary Tβ-4 as marker for systemic mastocytosis variant.

SLPI and the total S100A8 exhibited an opposite trend with respect to the other proteins and peptides, being up-regulated in SM patients, but not in CM, nor in healthy controls. SLPI is an anti-inflammatory and anti-microbial protein produced by neutrophils and macrophages associated to respiratory tract mucosa, parotid and submandibular glands[143]. SLPI acts as serine-protease inhibitor of cathepsin G, elastase, and trypsin released from neutrophils, chymotrypsin and trypsin from pancreatic acinar cells, and chymase and tryptase from mast cells[221]. The result obtained on SLPI in this study was suggestive for a self-protective response to the excessive mast cell activity in our patients.

S100A8, such as S100A9, is considered as ROS scavenger able to decrease the concentration of several redox intermediates or products, like H_2O_2[166]. It was demonstrated that S100A8, being particularly sensitive to oxidative cross-linking and massive oxidation, acquires capacity to reduce oxidative damages[167]. The increased levels of S100A8 particularly as nitrosylated proteoform in the subgroup of SM+C could suggest a protective reaction against ROS injuries caused by the inflammatory status associated to the disease in these patients. The result appeared remarkable by considering that S100A8-SNO is able to suppress mast cell-mediated inflammation by reducing leukocyte adhesion and extravasation, as demonstrated in a study performed in the rat mesenteric microcirculation[175].

It was interesting to observe the strong correlation between tryptase and the levels of some salivary peptides/proteins especially in the subgroup of SM-C patients, where increasing of

tryptase level correlated with decreasing level of Tβ-4, and with increasing levels of statherin fragments desAsp1, des1-10 and des1-13, of all the fragments of P-B peptide, of cystatin D-R26 des1-5, of PRP-3 desArg106, and the monophosphorylated forms of the aPRPs. Tryptase is a trypsin-like enzyme used as marker of mast cell activity. Mast cells release tryptase also into saliva, where it was proposed as diagnostic tool to test food allergies[222]. It is released from activated mast cells together several other proteases and peptidases, as chymase, cathepsin C and probably others, carboxypeptidase A3, and metalloproteinases, these mediators in turn influence non-mast cells proteases in response to infections and in order to resolve tissue injuries[207]. An unbalanced mast cell release of proteases converts their role from protective to damaging. It would be interesting to evaluate how the release of mediators by activated mast cells in the oral cavity could be associated to the variation of the levels of specific salivary proteins/peptides. Moreover, it was intriguing to have find that Tβ-4, PRP-3 desArg106, cystatin D-R26 des1-5, and P-B des1-5 not only correlated with tryptase but also resulted distinguishing elements between the two subgroups of SM patients. These results highlighted the possibility to distinguish the several forms of the disease based on the salivary protein profiling and thus can be considered promising for further study on a larger cohort.

The data collected on cystatin D-R26 were particularly questioning; indeed, among the three patients' groups the SM-C patients showed the highest levels of the two truncated proteoforms of this protein, but in any case, similar to that one of the healthy controls and thus considerable normal. On the other hand, it was possible to observe a very scattered distribution of the measured XIC peak area values, which likely may be linked to the existence of two cystatin D variant in human adult saliva, D-R26 and C26 carrying the substitution Arg>Cys at 26 position[137]. The expression of these two variants in the occidental population was established to be 55% for cysteine variant and 45% for arginine one[137]. In CM patients, the R26 variant was almost absent, while it was founded in low percentage in SM+C samples (33%). Interestingly, SM-C patients exhibited a higher percentage of cystatin D-R26 variant (75%), meanwhile the control group showed almost the same percentage reported in literature (44%). Moreover, a preliminary study performed on salivary samples from adult healthy people[142] demonstrated that the cystatin D-C26 variant, undetectable in the acid soluble fraction of saliva, was detectable in whole saliva associated to several other proteins in a multi-protein complex, which precipitate after the acid treatment of the saliva. Based on these considerations and since cystatin D appeared one of the possible candidates to be a marker of differentiation among the different forms of mastocytosis, the study continued, in the second part, with the analysis of whole saliva collected from the same patients and controls with the aim to characterize the cystatin D-C26 multi-protein complex.

PART II: IMMUNO-PROTEOMIC PLATFORM TO HIGHLIGHT QUALITATIVE AND QUANTITATIVE VARIATION OF SALIVARY PROTEIN COMPLEXES FOR THE CLASSIFICATION OF THE DIFFERENT FORM OF MASTOCYTOSIS

Material and methods

Immunoblotting

Whole saliva from 4 CM patients (2 males, 2 females, mean age ± SD: 39.3 ± 15.7), 3 SM-C patients (3 females, mean age ± SD: 69.3 ± 5.4) and 14 SM+C patients (7 male, 7 females, mean age ± SD: 42.8 ± 10.0) and from 20 sex/age matched healthy controls (11 males, 9 females, mean age ± SD: 38.7 ± 9.8) was treated, with 1:1 v/v ratio, with a solution of PBS2X (270mM NaCl, 5mM KCl, 20mM $NaHPO_4$, 4mM KH_2PO_4) containing mini-cOMPLETE (Sigma-Aldrich/Merck, Darmstadt, Germany), which inhibits proteases naturally present in saliva. Whole saliva total protein concentration was determined by bicinchoninic acid assay (Sigma-Aldrich/Merck, Darmstadt, Germany). Western blotting and immunodetection analysis were performed to screening cystatin D-C26 presence in the patients and controls samples, by using a mouse monoclonal Ab (Santa Cruz Biotechnology, Dallas, Texas, USA). 6µg of total proteins for each sample were used for the analysis. SDS-page was performed in non-reducing and reducing conditions using 4-15% T mini precast gels (Bio-Rad, Hercules, CA, USA). Electrophoresis separation was performed at 180V and Bio-Rad precision Plus Protein ™ WesternC ™ Blotting Standard (Hercules, CA, USA) was used as molecular weight standards. Then, the SDS separated proteins were transferred to a 0.2µm polyvinylidene difluoride membranes according to the instructions provided with the Trans-Blot Turbo system (Bio-Rad, Hercules, CA, USA). After the transfer, PDVF membranes were let under stirring for an hour with blocking solution (5% BSA in Tris-buffered saline containing 0.05% tween-20, TBS-T) at room temperature. Then, the membranes were incubated with mouse monoclonal primary antibodies, diluted for both cystatin D-C26 and PIgR antibodies, 1:1000 with blocking solution. After 6 washes of 5 minutes with TBS-T, membranes were incubated 1h with the HRP conjugated anti-mouse secondary antibody (Santa Cruz Biotechnology, Dallas, Texas, USA) diluted 1:50000 with blocking solution. After 6 washes of 5 minutes with TBS-T, membranes were incubated with the detection solution, ECL western clarity (Bio-Rad, Hercules, CA, USA), for 5 minutes. Then the images were acquired using a Bio-Rad ChemiDoc MP Imaging System (Bio-Rad, Hercules, CA, USA) and analyzed with Image Lab 4.0.1.

Immunoprecipitation assay

To perform the immunoprecipitation with cystatin D-C26 or PIgR Abs, pools of whole saliva were prepared for each group (CM, SM-C, SM+C, Ctrls). Each sample was diluted with PBS1X to obtain the right concentration, and equal volumes were used to prepare 1µg/µl protein concentration for CM, SM-C and control pools and 0.66µg/µl for SM+C pool. Aliquots containing 400µg of total proteins from each pool (400µl from CM, SM-C and controls pools, 606µl from SM+C pool) were treated with Preclear solution, which contain a nonspecific primary antibody (Santa Cruz Biotechnology, Dallas, Texas, USA) and BioRad AG plus Agarose protein (Hercules, California, USA), to remove contaminants and nonspecific complexes. Afterwards, specific primary antibody was added to the solution and let it incubate under stirring 4h at 4°C. Next, BioRad AG plus Agarose protein (Hercules, California, USA), which binds to the specific antibody, was added to the solution. The solution was let under stirring over night at 4°C. To break the interaction between AG protein and antibody and between antibody and proteins was used a cracking solution (0.125 M Tris/HCl pH 6.8, 4% SDS, 20% glycerol, 0.02 Bromophenol Blue). The samples were then boiled and centrifugated; the supernatant was analyzed by SDS-page in reducing (R) and non-reducing conditions (NR).

SDS-page

Eluates from immunoprecipitation assay of the three pools of patients and the pool of controls were analyzed by SDS-page in reducing and non-reducing using the condition previously indicated. The analysis was carried out using two gels, the former underwent to western blot and immunodetection with specific antibody (Cystatin D-C26 or PIgR) with the condition described above, while the latter was stained with Bio-Safe ^{TM} Coomassie G250 stain (Bio-Rad, Hercules, CA, USA) to be submitted to an in-gel tryptic digestion and high-resolution MS/MS analysis. Unfortunately, the amount of CM pool was enough only to perform the immunodetection analysis but not for the bottom-up proteomic characterization.

Tryptic digestion and MS/MS analysis

To perform the characterization of the immunoprecipitated proteins in SM-C, SM+C and Ctrls pools, the SDS-PAGE gel obtained in reducing and non-reducing conditions were submitted to an extensive excision along every line. Stained protein bands were manually excised from the

gel and transferred to fresh tubes. In addition, the zones of the lines that did not appear stained were excised as well. The gel pieces were washed one time with 100 mM ammonium bicarbonate, one time with 100 mM ammonium bicarbonate/50% acetonitrile v/v (ACN, Sigma-Merck, St. Louis, MO, USA) and one time with ACN (10 minutes each wash) to de-stain the gel pieces. The protein's cysteine residues were reduced with 10 mM dithiothreitol (30 minutes at 65°C, in the dark) and alkylated with 55 mM iodoacetamide (30 min at RT, in the dark). The gel pieces were washed one time with 100 mM ammonium bicarbonate and one time with ACN, dried and rehydrated in digestion buffer (40 mM ammonium bicarbonate/ultrapure ACN in 10:1 v/v ratio) containing high-specific trypsin (Trypsin Single, Proteomics grade – Sigma-Merck, St. Louis, MO, USA), with a ratio of 1:30 (w/w) of enzyme to substrate. The samples were incubated overnight at 37ºC. After the incubation, 0,1% of formic acid (FA) was used to block the reaction. Extraction of tryptic peptides was performed by two-fold addition of 100 mM ammonium bicarbonate/ACN in 1:2 v/v ratio. Tryptic peptides were lyophilized and stored at -80°C until the analysis. The samples were then resuspended in 1% FA and analyzed with a QExactive Orbitrap (Thermo Fisher Scientific, Bremen) through the EASY-spray nano ESI source (Thermo Fisher Scientific, Bremen) that was coupled to an Ultimate 3000 (Dionex, Sunnyvale, CA) HPLC (high-pressure liquid chromatography) system. The trap (5 mm × 300 µm I.d.) and the EASY-spray analytical (150 mm × 75 µm) columns used were C18 Pepmap100 (Dionex, LC Packings) having a particle size of 3 µm. Peptides were trapped at 30 µl/min in 96% solvent A (0.1 % FA). Elution was achieved with the solvent B (0.1 % formic acid/80% acetonitrile v/v) at 300 nl/min. The 92 min gradient used was as follows: 0–3 min, 96% solvent A; 3–70 min, 4–25% solvent B; 70–90 min, 25–40% solvent B; 90–92 min, 90% solvent B; 90–100 min, 90% solvent B; 101-120 min, 96% solvent A. The mass spectrometer was operated at 2.3 kV in the data dependent acquisition mode. A MS2 method was used with a FT survey scan from 400 to 1600 m/z (resolution 70,000; AGC target 1E6). The 10 most intense peaks were subjected to HCD fragmentation (resolution 17,500; AGC target 5E4, NCE 28%, max. injection time 100 ms, dynamic exclusion 35 s). Spectra were processed and analyzed using Proteome Discoverer (version 2.2, Thermo – PD), with the MS Amanda (version 2.0, University of Applied Sciences Upper Austria, Research Institute of Molecular Pathology) and Sequest HT search engines (University of Washington, in license to Thermo Electron Corp., San Jose, CA). Uniprot (TrEMBL and Swiss-Prot) protein sequence database (version of October 2017) was used for all searches under *Homo sapiens*. Database search

parameters were as follows: carbamidomethylation of cysteine, oxidation of methionine, cysteine/aspartic-acid/histidine/lysine/arginine/serine/threonine/tyrosine phosphorylation, C-terminal pyroglutamic residue, N-terminal acetylation, and the allowance for up to two missed tryptic cleavages. The peptide mass tolerance was 10 ppm and fragment ion mass tolerance was 0.02 Da. To achieve a 1% false discovery rate, the Percolator (version 2.0, Thermo) node was implemented for a decoy database search strategy and peptides were filtered for high confidence and a minimum length of 6 amino acids, and proteins were filtered for a minimum number of peptide sequences of 2. Settings of FDR were: 0.01 (strict) and 0.05 (relaxed). Every PD identification was confirmed by manual inspection of the MS/MS spectra selected by the software. All the proteins identified by PD on the base of MS/MS spectra considered by us poor of multiply-charged ions attributed to the protein sequence were discarded. PD protein quantification was based on the total abundance of tryptic peptides generated by a protein; then PD software calculated, for each comparison between the different groups, the median of ratios between abundancies of the peptides belonging to a specific protein.

Results

Immunoprecipitation assay and immunodetection with cystatin D-C26 antibody

The results from acid-soluble fraction of saliva highlighted differences between the three groups of patients regarding the concentration of cystatin D-R26, which was the only isoform detectable in acid-soluble fraction of salivary proteins. Since this result, we focused on the study of whole saliva's patients in order to highlight the formation of high molecular weight multi-protein complex associated to cystatin D-C26, SIC-D, as described by Liori et al.[142], and underline eventually qualitative/quantitative variation between patients and controls and among the patients.

2 CM (37, 45), 2 SM-C (43, 46) and 13 SM+C (13, 19, 20, 36, 47, 51, 52, 58, 60, 63, 64, 65,66) patients and some of the controls (8) whole saliva was subjected to immunodetection by western blotting analysis using the cystatin D-C26 Ab. As shown in Figure 10, the different groups of patients showed a heterogeneous trend regarding cystatin D concentration. In NR condition is possible to distinguish a signal higher than 250kDa, with variable intensity depending on the subject, that disappear in R condition, suggesting the presence of cystatin D-C26 bounded to one or more proteins. Since this preliminary data from the screening of patients and controls, we decided to perform an immunoprecipitation assay using cystatin D-C26 antibody on pools of different patients (CM, SM-C and SM+C) and on the pool of control. The result of western blot and immunodetection using cystatin D-C26 antibody are shown in Figure 11. In NR conditions, it was possible to highlight a signal higher than 250kDa with different intensity depending on the pool, suggesting that, as in whole saliva, cystatin D-C26 was part of complex (SIC-D) with to one or more partners. The most intense signal in NR condition was at 150kDa and it corresponded to the antibody used for the immunoprecipitation assay. On the contrary, in R condition the signals in the upper part of the gel disappeared and it was possible to distinguish the signal of heavy and light chain of cystatin D-C26 antibody at 50kDa and 25kDa and the signal of cystatin D between 10kDa and 15kDa.

Figure 10: Western blot and immunodetection with cystatin D-C26 antibody of whole saliva from patients and controls in NR and R conditions. Panels A-D show the immunodetection using the cystatin D-C26 Ab of patient's whole saliva, while in Panels E and F of the controls, both performed in NR and R conditions. The red square indicates the high molecular weight complex in which is present the cystatin D and its partners (Panels A, C and E), that disappears in R conditions (Panels B, D and F).

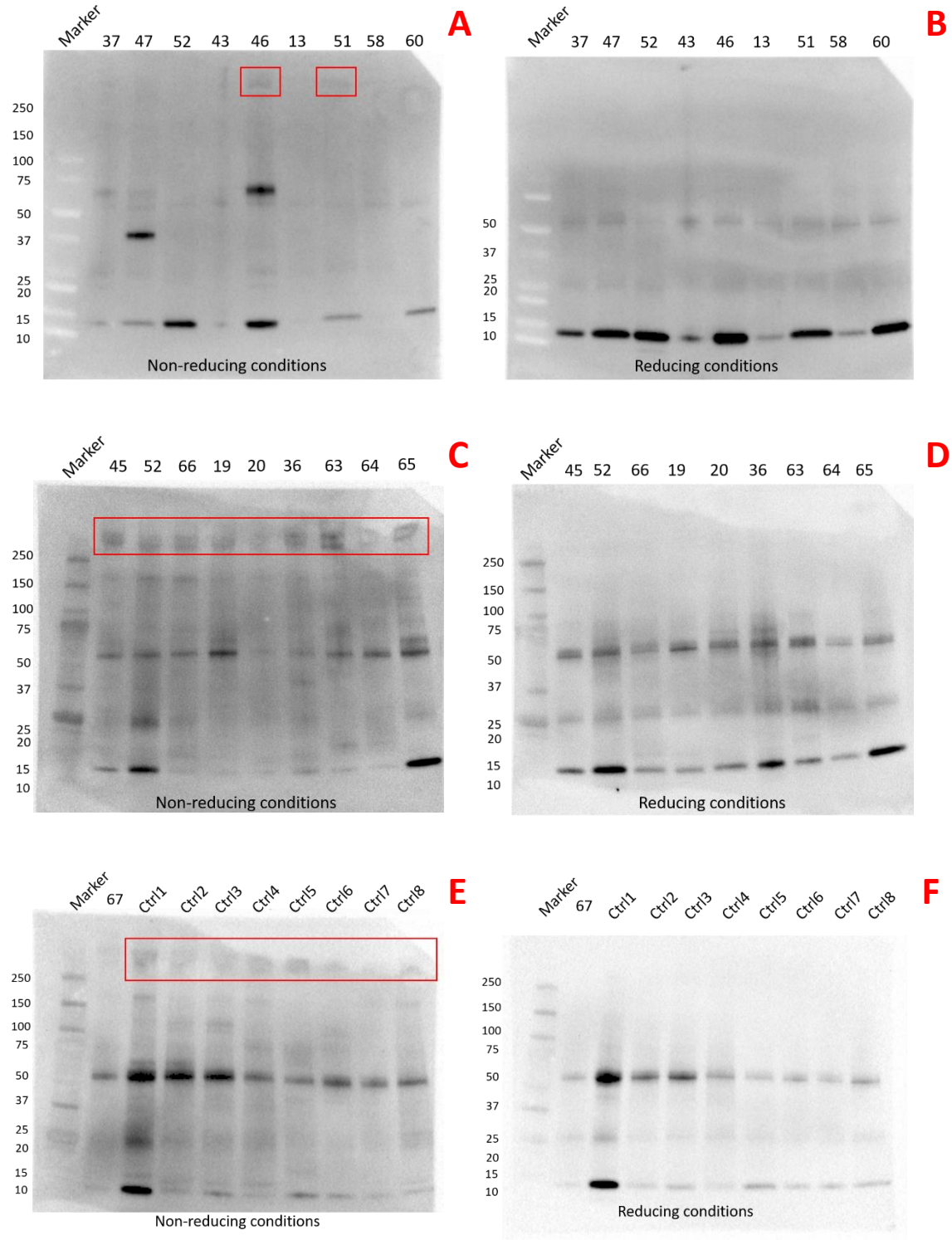
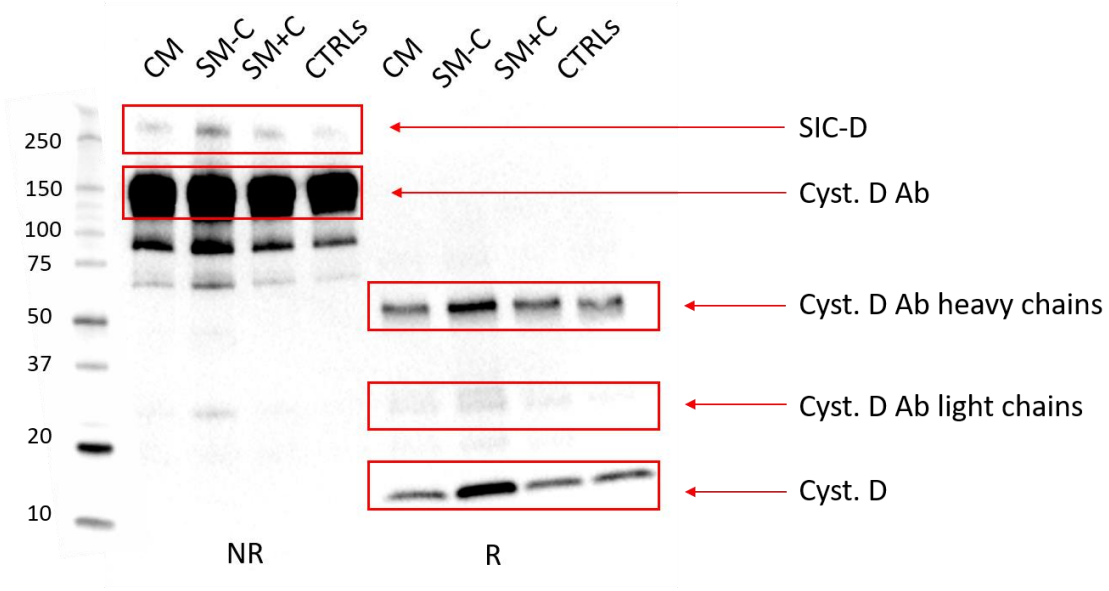


Figure 11: Western blot and immunodetection with cystatin D-C26 antibody of immunoprecipitated samples from the pools of patients and the pool of controls in NR and R conditions.

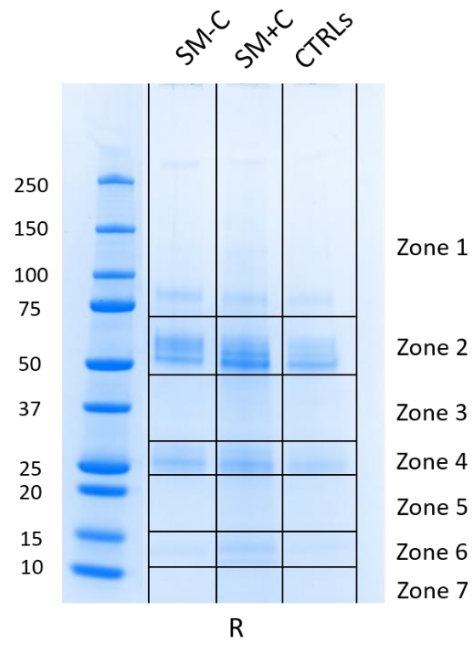


High-resolution MS/MS analysis of immunoprecipitated with cystatin D-C26 antibody

The results from western blot and immunodetection using cystatin D C-26 Ab highlighted the presence of a high molecular weight SIC-D in which cystatin D seemed to be linked to one or more partners. Thus, we decided to characterize SIC-D in order to highlight qualitative and/or quantitative variations between the groups of patients and the healthy subjects and among the patients. Since the low volume of whole saliva from CM samples, the high-resolution mass spectrometry analysis was carried out only on immunoprecipitated (IP) from SM-C, SM+C and Ctrls pools.

For the bottom-up analysis, immunoprecipitated samples from patients and control pools were submitted to SDS-PAGE in reducing condition and Comassie Brilliant Blue staining, and the lines of the three IP were excised in the same way as described in Figure 14: piece 1 (zone > 75kDa); piece 2(50 > zone < 75kDa); zone 3 (40kDa > zone < 50kDa); zone 4 (25kDa > zone < 40kDa); zone 5 (15kDa > zone < 25kDa); zone 6 (10kDa > zone < 15kDa); zone 7 (<10kDa). All the gel pieces were digested using trypsin and analyzed by nano-HPLC-high-resolution-MS/MS. Proteome discoverer (PD) software was used to perform the spectra analysis and the qualitative and quantitative.

Figure 12: Gel that has been undergone to tryptic digestion and high-resolution mass spectrometry analysis. In the image is indicated how the bands have been cut.



The proteins identified in the IPs from the patients' groups (SM-C and SM+C) and controls are shown respectively in Tables 8, 9, and 10, which report just those proteins, among the identified ones with high confidence, having a good combination of score, coverage and quality of MS/MS spectra provided by PD software. All the results obtained by high-resolution MS/MS experiments have been deposited to ProteomeXchange Consortium (http://ww.ebi.ac.uk/pride) via the PRIDE[223] partner repository with the dataset identifier PXD015866.

Zone 1: high molecular weight proteins. IPs from the patients and from the control group highlighted the presence of high molecular weight proteins, such as PIgR (the secretory component, SC), deleted in malignant brain tumors 1 protein (DMBT-1), and protein-glutamine gamma-glutamyl transferase E, also known as transglutaminase E or 3. In SM-C and SM+C patients also lactotrasferrin and desmoglein-1 were identified that were not present in control groups. On the other hand, SM-C and controls SIC-D exhibited respectively of junction plakoglobin and lactoperoxidase.

Zones 2 and 3: proteins with of molecular weight from 25 to 75kDa. All the patients and the control group showed the presence of immunoglobulins, i.e. IgA1, IgA2 and IgG1, while SM+C group exhibited also IgG2, IgG4 and IgM. In addition, serum albumin and actin-1 were present in all the three groups. Annexin A1, a protein involved in innate immune response, was identified in the SIC-D of all the groups analyzed. The presence of annexin A2 was revealed only in SM-C patients, as well as α-amylase 1, while actin-2 was found only in SM-C and controls. SM+C and control complex showed the presence of carbonic anhydrase 6 (CA-6), which were not found in SM-C. Strangely, in the zones 2 and 3 of all the three groups were found immunoglobulin heavy variable (IgHV) region, the variable domain of immunoglobulin heavy chains that participates in the antigen recognition, even though their molecular weight suggest that they should be in lower zones. In SM-C, SM+C and controls ones were present IgHV 3-9, IgHV 3-23 and IgHV 3-74, while in SM-C and SM+C were detected IgHV 3-72 and IgHV 3-74. IgHV 3-7 and 3-13 were found only in SM-C.

Zones 4, 5, 6 and 7: low-molecular weight proteins (< 25kDa). All the three groups showed the presence of proteins and peptides related to inflammatory response and immunity system, such as S100A8, α-defensin 1/2, cystatin D-C26, prolactin-inducible protein (PIP), lysozyme C and Igκ. Cystatin C was identified in SIC-D of SM-C patients and control, but not in SMC+C patients. Antileukoproteinase (SLPI) and S100A14 were identified in both SM-C and SM+C SIC-D. Some proteins were characterized just in the SIC-D of one patient group, such as cystatin A in SM-C and cystatin B, S100A9 and dermcidin in SM+C.

Table 8: Qualitative results of mass spectrometry and PD analysis of SM-C immunoprecipitated with cystatin D-C26 antibody.

Zone 1 SM-C								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P01833	Polymeric immunoglobulin receptor (PIgR)	8960.23	49.46	40	26	26	83.2	4.21E+06
Q9UGM3	Deleted in malignant brain tumors 1protein (DMBT-1)	7340.85	25.36	35	19	19	260.6	4.22E+06
Q08188	Protein-glutamine gamma-glutamyltransferase E	2871.73	39.11	28	13	13	76.6	2.70E+05
P14923	Junction plakoglobin	701.05	16.19	11	6	6	81.7	7.76E+04
Q02413	Desmoglein-1	557.56	15.79	8	5	5	113.7	3.15E+04
P02788	Lactotransferrin	520.45	2.87	5	3	3	78.1	2.46E+04
Zone 2 SM-C								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
A0A0B4J1X5	Immunoglobulin heavy variable 3-74 (IgHV3-74)	714.71	9.44	41	5	3	12.8	4.24E+04
P04745	α-amylase 1	2077.02	47.84	36	15	15	57.7	5.93E+05
P0DOX2	Immunoglobulin alpha-2 heavy chain	5592.46	104.81	35	12	2	48.9	3.46E+06
P63261	Actin, cytoplasmic 2	2171.67	33.08	35	9	1	41.8	*
P60709	Actin, cytoplasmic 1	2161.67	33.08	35	9	1	41.7	4.00E+05
P01780	Immunoglobulin heavy variable 3-7 (IgHV 3-7)	868.36	11.26	31	4	2	12.9	8.39E+04
P0DOX5	Immunoglobulin gamma-1 heavy chain	2257.33	22.91	29	10	9	49.3	1.73E+06
P01833	Polymeric immunoglobulin receptor (PIgR)	984.39	*NA*	11	7	7	83.2	1.16E+05
P01764	Immunoglobulin heavy variable 3-23 (IgHV 3-23)	451.47	6.19	22	3	2	12.6	3.07E+04
A0A0B4J1Y9	Immunoglobulin heavy variable 3-72 (IgHV 3-72)	361.19	10.16	21	3	3	13.2	1.20E+05
P01782	Immunoglobulin heavy variable 3-9 (IgHV 3-9)	484.62	6.69	19	2	1	12.9	9.97E+03
P01766	Immunoglobulin heavy variable 3-13 (IgHV 3-13)	323.89	6.69	19	2	1	12.5	3.69E+03
P02768	Serum albumin	1321.8	20.07	10	5	5	69.3	1.13E+06

Zone 3 SM-C								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P04083	Annexin A1	855.06	7.49	13	5	5	38.7	1.26E+05
P01876	Immunoglobulin heavy constant alpha 1	722.34	9.56	11	4	2	37.6	6.58E+04
P07355	Annexin A2	1237.62	27.91	10	4	4	38.6	8.59E+04
P0DOX2	Immunoglobulin alpha-2 heavy chain	137.4	2.26	10	4	2	48.9	1.87E+04
Zone 4 SM-C								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P01834	Immunoglobulin kappa constant	2347.4	52.76	81	6	1	11.8	3.95E+04
P0DOX7	Immunoglobulin kappa light chain	1929.05	55.23	50	7	2	23.4	2.33E+06
P0DOX8	Immunoglobulin lambda-1 light chain	624.01	32.72	29	5	2	22.8	6.23E+04
Zone 5 SM-C								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P01591	Immunoglobulin J chain	*NA*	8.8	8	1	1	18.1	1.73E+05
Zone 6 SM-C								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P05109	Protein S100-A8	342.86	*NA*	19	2	2	10.8	1.27E+05
P12273	Prolactin-inducible protein	711.54	1.98	14	2	2	16.6	4.20E+04
P28325	Cystatin-D-C26	549.45	*NA*	11	1	1	16.1	3.16E+04
P03973	Antileukoproteinase (SLPI)	96.55	*NA*	9	1	1	14.3	2.54E+04
P01591	Immunoglobulin J chain	230.58	4.12	8	1	1	18.1	3.53E+03
P61626	Lysozyme C	61.14	*NA*	8	1	1	16.5	1.66E+03
Zone 7 SM-C								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P59665	α-defensin 1/2	598.91	*NA*	19	2	2	10.2	1.94E+05

Accession	Description	Score MS Amanda	Score Sequest HT	Coverage [%]	Peptides	Unique Peptides	MW [kDa]	Abundance
Q9HCY8	Protein S100-A14	*NA*	5.08	14	1	1	11.7	1.78E+03
P01040	Cystatin-A	*NA*	2.72	12	1	1	11	9.15E+02

** Discovered in traces*
NA: in some cases, search engine did not provide a score.

Table 9: Qualitative results of mass spectrometry and PD analysis of SM+C immunoprecipitated with cystatin D-C26 antibody.

Zone 1 SM+C								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P01833	Polymeric immunoglobulin receptor (PIgR)	3359.33	21.93	25	16	16	83.2	2.21E+06
Q9UGM3	Deleted in malignant brain tumors 1 protein (DMBT1)	751.96	7.08	22	5	5	260.6	3.40E+05
P02788	Lactotransferrin	720.51	8.97	15	9	9	78.1	5.35E+04
Q08188	Protein-glutamine gamma-glutamyltransferase E	640.05	21.8	13	5	5	76.6	3.48E+04
Q02413	Desmoglein-1	671.65	28.13	8	5	5	113.7	4.14E+04
Zone 2 SM+C								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
A0A0B4J1Y9	Immunoglobulin heavy variable 3-72 (IgHV3-72)	448.68	6.95	30	4	4	13.2	7.97E+04
A0A0B4J1X5	Immunoglobulin heavy variable 3-74 (IgHV3-74)	767.25	9.54	25	3	1	12.8	2.64E+04
P01764	Immunoglobulin heavy variable 3-23 (IgHV3-23)	476.18	6.85	22	3	2	12.6	3.72E+04
P01782	Immunoglobulin heavy variable 3-9 (IgHV3-9)	349.77	3.58	19	2	2	12.9	2.99E+04
P01859	Immunoglobulin heavy constant gamma 2	1748.92	9.32	18	6	1	35.9	1.00E+06
P0DOX5	Immunoglobulin gamma-1 heavy chain	1286.69	11.01	18	7	4	49.3	1.70E+06
P01861	Immunoglobulin heavy constant gamma 4	1503.96	2.31	18	5	1	35.9	1.17E+03
P0DOX2	Immunoglobulin alpha-2 heavy chain	2932.11	66.56	15	8	2	48.9	6.82E+05
P02768	Serum albumin	2064.06	40.25	13	9	8	69.3	2.87E+06
P60709	Actin, cytoplasmic 1	415.15	2.71	13	3	2	41.7	8.97E+03
P0DOX6	Immunoglobulin mu heavy chain	417.72	5.99	7	3	3	63.4	1.84E+04
P23280	Carbonic anhydrase 6 (CA-6)	162.08	5.12	7	2	2	35.3	5.24E+03
SM+C 3								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P04083	Annexin A1	458.79	5.33	9	3	3	38.7	3.15E+04

Accession	Description	Score MS Amanda	Score Sequest HT	Coverage [%]	Peptides	Unique Peptides	MW [kDa]	Abundance
P01876	Immunoglobulin heavy constant alpha 1	309.22	2.44	6	2	2	37.6	3.34E+04
SM+C 4								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P01834	Immunoglobulin kappa constant	771.66	20.29	66	4	1	11.8	2.35E+03
P0DOX7	Immunoglobulin kappa light chain	641.04	11.85	33	4	1	23.4	4.06E+04
P59665	α-defensin 1/2	73.83	*NA*	10	1	1	10.2	9.64E+02
P01861	Immunoglobulin heavy constant gamma 4	190.59	*NA*	8	2	2	35.9	3.02E+04
SM+C 6								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P12273	Prolactin-inducible protein	390.95	2.5	21	3	3	16.6	2.64E+04
P28325	Cystatin-D-C26	265.59	2.81	20	2	2	16.1	7.71E+04
Q9HCY8	Protein S100-A14	*NA.*	4.58	14	1	1	11.7	1.35E+03
P05109	Protein S100-A8	285.48	*NA*	12	1	1	10.8	1.04E+05
P04080	Cystatin-B	73	*NA*	12	1	1	11.1	5.74E+03
P06702	Protein S100-A9	120.58	4.05	11	1	1	13.2	1.39E+05
P81605	Dermcidin	63.82	*NA*	10	1	1	11.3	2.48E+03
P03973	Antileukoproteinase (SLPI)	101.3	*NA*	9	1	1	14.3	1.41E+04
P61626	Lysozyme C	98.38	6.44	8	1	1	16.5	2.22E+04
P01034	Cystatin-C	91.86	*NA*	8	1	1	15.8	9.07E+03

NA: in some cases, search engine did not provide a score.

Table 10: Qualitative results of mass spectrometry and PD analysis of controls immunoprecipitated with cystatin D-C26 antibody.

Zone 1 Ctrls								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
Q9UGM3	Deleted in malignant brain tumors 1 protein (DMBT-1)	1398.06	11.85	11	2	2	260.6	4.38E+06
P01833	Polymeric immunoglobulin receptor (PIgR)	3027.45	3.37	8	4	4	83.2	2.15E+06
Q08188	Protein-glutamine gamma-glutamyltransferase E	600.66	3.52	7	3	3	76.6	2.42E+05
P22079	Lactoperoxidase	209.97	2.55	4	2	2	80.2	7.52E+05
Zone 2 Ctrls								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P01876	Immunoglobulin heavy constant alpha 1	2013.84	41.54	20	6	4	37.6	2.08E+06
P01764	Immunoglobulin heavy variable 3-23 (IgHV3-23)	277.21	2.96	19	2	2	12.6	3.89E+04
P01782	Immunoglobulin heavy variable 3-9 (IgHV3-9)	236.39	4.57	19	2	2	12.9	2.44E+04
P01859	Immunoglobulin heavy constant gamma 2	986.96	5.75	18	6	3	35.9	2.31E+05
P0DOX5	Immunoglobulin gamma-1 heavy chain	995.02	13.72	16	6	3	49.3	6.18E+05
P0DOX2	Immunoglobulin alpha-2 heavy chain	1141.93	23.45	15	6	4	48.9	6.12E+05
P02768	Serum albumin	1006.8	8.49	8	6	6	69.3	4.36E+05
Zone 3 Ctrls								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P04083	Annexin A1	1780.88	66.15	25	6	6	38.7	9.46E+04
P63261	Actin, cytoplasmic 2	378.07	13.92	16	4	1	41.8	*
P60709	Actin, cytoplasmic 1	370.38	13.92	16	4	1	41.7	1.03E+05
P23280	Carbonic anhydrase 6 (CA-6)	1195	43.41	7	2	2	35.3	3.64E+04
Zone 4 Ctrls								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P01834	Immunoglobulin kappa constant	1461.94	4.13	42	3	1	11.8	1.66E+04

Accession	Description	Score MS Amanda	Score Sequest HT	Coverage [%]	Peptides	Unique Peptides	MW [kDa]	Abundance
P0DOX7	Immunoglobulin kappa light chain	1384.01	6.88	21	3	1	23.4	3.33E+04
Zone 5 Ctrls								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P01591	Immunoglobulin J chain	148.4		8	1	1	18.1	*
Zone 6 Ctrls								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P12273	Prolactin-inducible protein	164.75	*NA*	17	2	2	16.6	7.55E+03
P05109	Protein S100-A8	1199.2	*NA*	12	1	1	10.8	2.66E+06
P28325	Cystatin-D-C26	518.4	*NA*	11	1	1	16.1	*
P61626	Lysozyme C	289.2	3.34	8	1	1	16.5	1.06E+06
P01064	Cystatin C	60.52	2.14	8	1	1	15.8	5.76E+03
Zone 7 Ctrls								
Accession	**Description**	**Score MS Amanda**	**Score Sequest HT**	**Coverage [%]**	**Peptides**	**Unique Peptides**	**MW [kDa]**	**Abundance**
P59665	α-defensin 1/2	432.27	*NA*	10	1	1	10.2	*

** Discovered in traces*
NA: in some cases, search engine did not provide a score.

Quantitative comparison of the immunoprecipitated with cystatin D-C26 antibody

PD software was used to perform the comparison of SIC-D protein abundances between patients and controls and among the patients' groups. The quantification was based on the total abundance of tryptic peptides generated by a specific protein, the PD calculated for each comparison between patients and controls and patients' groups the median of ratio between abundancies of the tryptic peptides belonging to a specific protein. Despite 44 proteins/peptides were identified in the different groups, significant differences were obtained for the 16 proteins/peptides, that are shown in table 11.

Immunoglobulins turned out to be the most expressed proteins in SIC-D, with different levels between the different groups. SM-C and SM+C complexes revealed respectively minor levels of IgA1 and both IgA1 and IgA2 compared to the controls one, while not considerable differences emerged comparing the patients (Table 11 and Figure 13A). On the contrary, IgG1, IgG2 and Igλ turned out to be more concentrated in SM-C and SM+C complexes than in the healthy subjects' complex, but no differences have been found comparing the two patient groups. SM-C SIC-D complex showed decreased level of immunoglobulin J chain when compared to both SM+C and controls. The different HV regions of immunoglobulins resulted to have different levels in the three groups. IgHV 3-15 abundance resulted lower in SM-C than in SM+C and controls, while IgHV 3-23 levels turned out to be lower in SM+C with respect to SM-C ad healthy controls. No significant differences emerged for IgHV 3-7, IgHV 3-9, IgHV 3-13, IgHV 3-72 and IgHV 3-74.

The levels of PIgR resulted lower in SM-C SIC-D with respect to controls and SM+C ones (Table 11 and Figure 13B). Many other proteins related to immune system such as DMTB-1, α-defensin 1/2 and S100A8 co-immunoprecipitated with cystatin D-C26. The abundance of DMTB-1 resulted to be lower in both SM-C and SM+C complexes than in healthy controls. Among the patients, SM-C showed the lowest amount (Figure 13C). The levels of the α-defensin 1/2 turned out to be higher in SM-C and SM+C complexes compared to the controls one. Comparing the patients, SM-C showed the lowest level. Similarly, CA-6 abundance resulted to be higher in patients' complex compared to the healthy controls (Figure 14A). Cystatin D-C26 revealed to be more abundant in SM+C group with respect to both SM-C and controls groups, with SM-C group showed the lowest abundance of cystatin D-C26 (Table 11 and Figure 14B). SM-C and SM+C SIC-D showed a lower abundance of S100A8 compared to

the controls. None differences have been found comparing the two patients' complexes (Figure 14C).

The data on the 16 proteins reported in Table 11 were elaborated by PD into PCA plot, which is displayed in Figure 15. Panel A highlighted that the three different groups are different one to each other regarding the levels of the different proteins/peptides that are co-immunoprecipitated with cystatin D-C26 and thus are part of SIC-D. Panel B showed the distribution of the different proteins and peptides of SIC-D based on their concentration. It was possible to discriminate in the four quadrants proteins at equal level in SM-C and SM+C (Quadrant II), proteins distinguishing SM-C and SM+C (Quadrant III) and proteins distinguishing controls and patients (Quadrant IV).

Table 11: Quantitative comparison of SIC-D between patients and controls and among patients' groups. In the Table are indicated score, coverage and ratios of each protein identified.

Accession	Description	Score MS Amanda	Score Sequest HT	Coverage [%]	SM-C/Ctrls	SM+C/Ctrls	SM-C/SM+C
P01876 (#1)	Immunoglobulin heavy constant alpha 1	22915.9	415.0	41	0.38 ↓SM-C	0.63 ↓SM+C	0.64 ↓SM-C
P01834 (#2)	Immunoglobulin kappa constant	8485.2	123.3	79	0.44 ↓SM-C	1.47 ↑SM+C	0.30 ↓SM-C
P0DOX7 (#3)	Immunoglobulin kappa light chain	7877.2	93.1	40	0.09 ↓SM-C	0.60	0.15 ↓SM-C
P0DOX2 (#4)	Immunoglobulin alpha-2 heavy chain	10906.3	142.2	24	0.37 ↓SM-C	0.46 ↓SM+C	0.62 ↓SM-C
P01833 (#5)	Polymeric immunoglobulin receptor (PIgR)	7005.2	45.5	12	0.06 ↓SM-C	1.12	0.06 ↓SM-C
P0DOX5 (#6)	Immunoglobulin gamma-1 heavy chain	4953.2	52.4	28	2.36 ↑SM-C	4.48 ↑SM+C	0.54 ↓SM-C
P01859 (#7)	Immunoglobulin heavy constant gamma 2	3771.5	34.7	24	1.74 ↑SM-C	1.69 ↑SM+C	0.95
Q9UGM3 (#8)	Deleted in malignant brain tumors 1 protein (DMBT-1)	2138.2	32.0	12	0.05 ↓SM-C	0.22 ↓SM+C	0.22 ↓SM-C
P0DOY2 (#9)	Immunoglobulin lambda constant 2	2221.1	24.3	44	1.27	2.43 ↑SM+C	0.52 ↓SM-C
P0DOX8 (#10)	Immunoglobulin lambda-1 light chain	1610.6	14.4	18	1.62 ↑SM-C	3.13 ↑SM+C	0.52 ↓SM-C
P23280 (#11)	Carbonic anhydrase 6 (CA-6)	970.0	7.9	7	8.49 ↑SM-C	11.85 ↑SM+C	0.72
P59665 (#12)	α-defensin 1/2	570.5	*NA*	19	4.53 ↑SM-C	100 ↑SM+C	0.02 ↓SM-C
P01764 (#13)	Immunoglobulin heavy variable 3-23 (IgHV 3-23)	774.6	*NA*	9	0.57 ↓SM-C	0.37 ↓SM+C	1.56 ↑SM-C
P01591 (#14)	Immunoglobulin J chain	868.8	11.7	8	0.40 ↓SM-C	1.26	0.32 ↓SM-C
P28325 (#15)	Cystatin-D	341.2	*NA*	11	0.42 ↓SM-C	1.95 ↑SM+C	0.21 ↓SM-C

Accession	Description	Score MS Amanda	Score Sequest HT	Coverage [%]	SM-C/Ctrls	SM+C/Ctrls	SM-C/SM+C
P05109 (#16)	Protein S100A8	68.8	*NA*	12	0.24 ↓SM-C	0.22 ↓SM+C	1.11

NA: in some cases, search engine did not provide a score.

Figure 13: Abundancies and ratios of IgA1 (A), PIgR (B) and DMBT-1 (C) in the different subjects. Yellow, green and blue columns indicate respectively SM-C, SM+C and Ctrls abundancies.

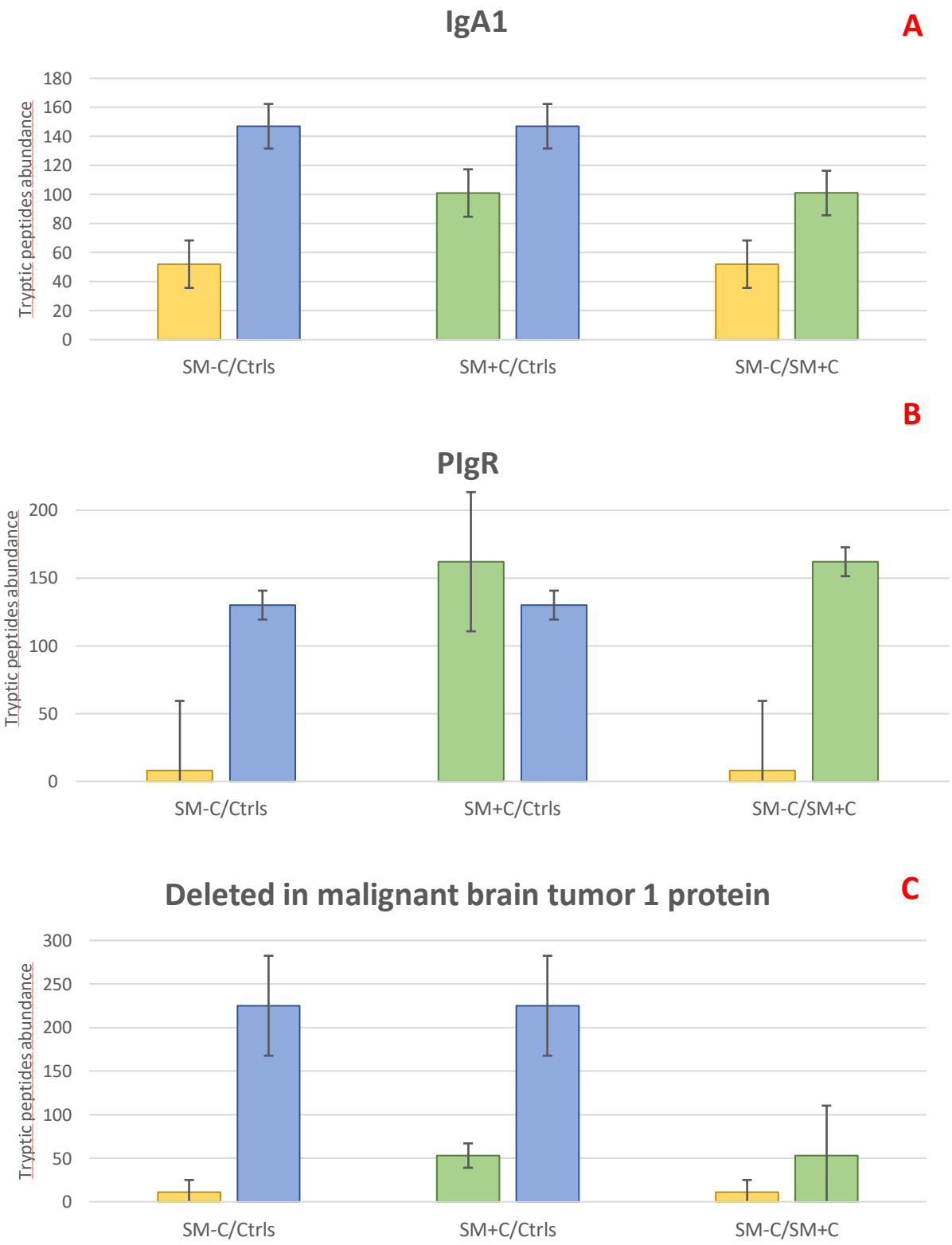


Figure 14: Abundancies and ratios of CA-6 (A), cystatin D-C26 (B) and S100A8 (C) in the different subjects. Yellow, green and blue columns indicate respectively SM-C, SM+C and Ctrls abundancies.

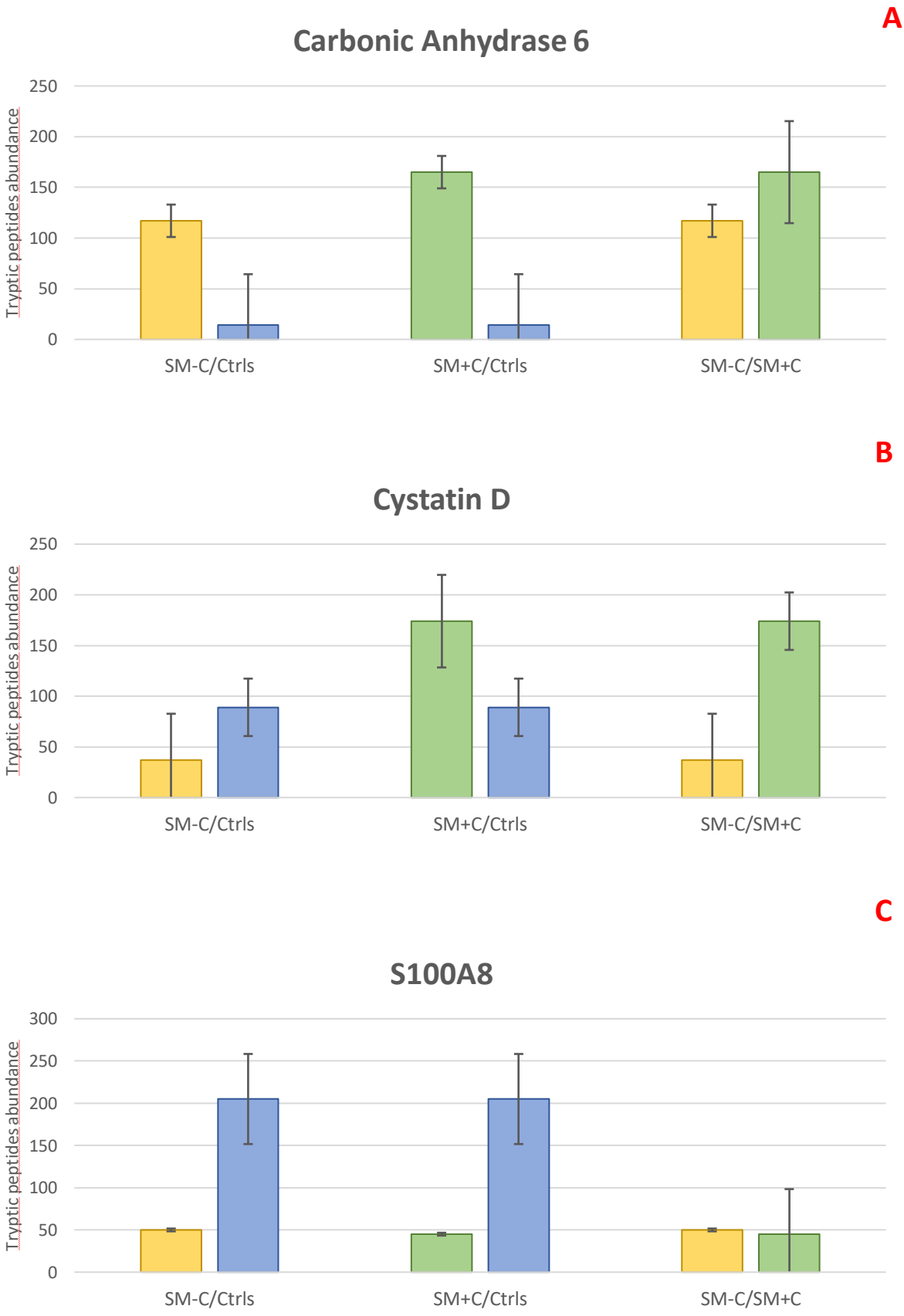
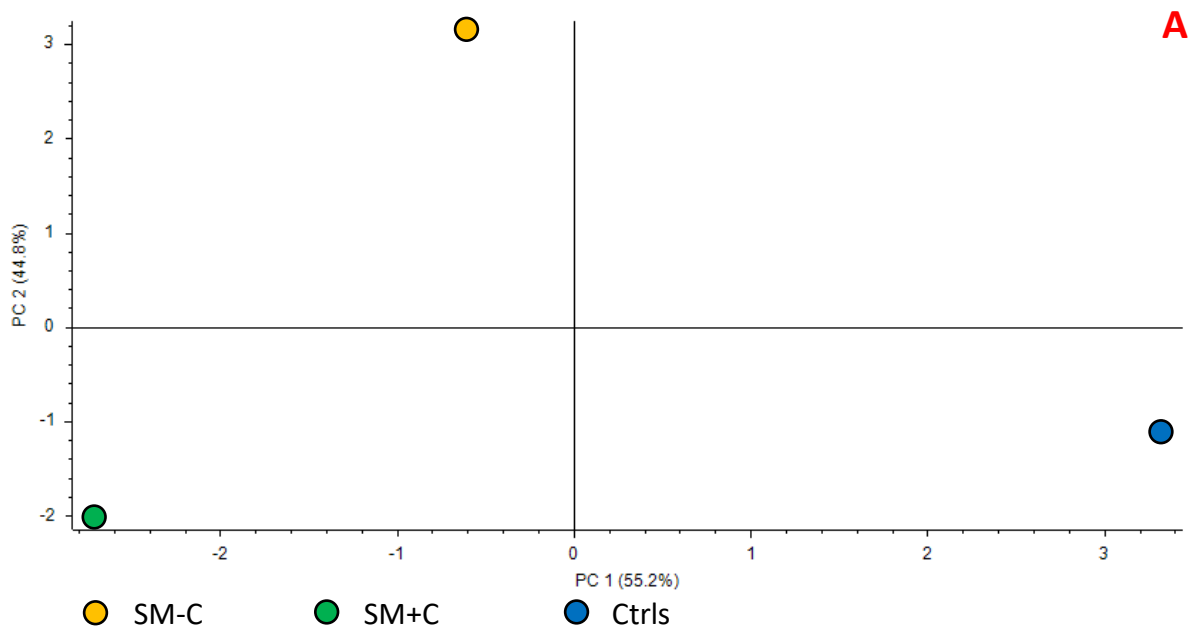
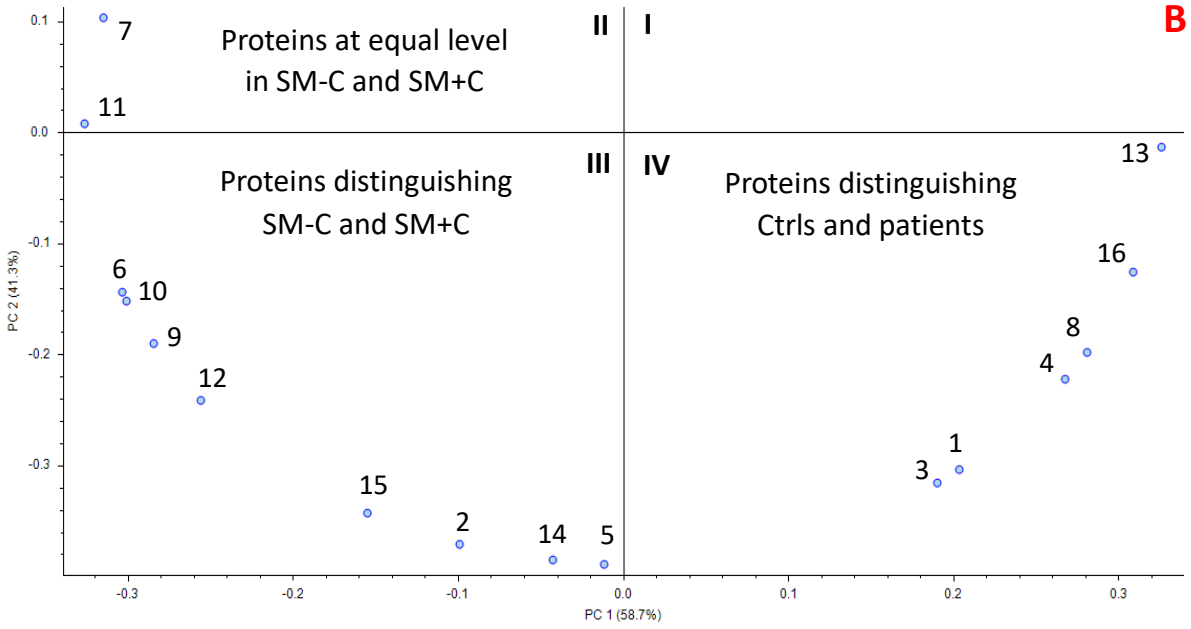


Figure 15: PCA plots of quantitative analysis. Panel A highlights that the four groups differ for the concentration of the different proteins. In panel B proteins is shown the distribution of the proteins based on their concentration (numbers in panel B are referred to proteins in Table 11).





Immunoprecipitation assay and immunodetection with PIgR antibody

The results from qualitative and quantitative analysis of immunoprecipitated assay with cystatin D-C26 antibody indicated that PIgR was one of cystatin D partners. Thus, pools of whole saliva from the different patients (CM, SM-C and SM+C) and from healthy subjects were subjected to immunoprecipitation assay using PIgR antibody, western blot and immunodetection with PIgR antibody. The results are shown in Figure 16. Similar to cystatin D-C26 immunoprecipitation and detection, in NR conditions was possible to highlight a signal higher than 250kDa with different intensity depending on the pool and with a different band shape, suggesting that, as cystatin D-C26, PIgR tends to associate to one or more partners, forming a complex (SIPIgR). The signal at 150kDa in NR corresponds to the antibody used for the immunoprecipitation assay. In R condition, the signal at > 250kDa higher signal disappears, and it is possible to discriminate the PIgR signal at 80kDa, and the signal of PIgR antibody at 50kDa and 25kDa.

Figure 17 shows the western blot and immunodetection from the two high molecular weight salivary immunocomplexes. SIC-D appears as a single band, while SIPIgR shows two or more bands depending on the sample. The western blot and immunodetection indicate that probably two salivary immunocomplexes are present, with differences from both qualitative and quantitative point of view.

Figure 16: Western blot and immunodetection with PIgR antibody of immunoprecipitated samples from the pools of patients and the pool of controls in non-reducing (NR) and reducing (R) conditions.

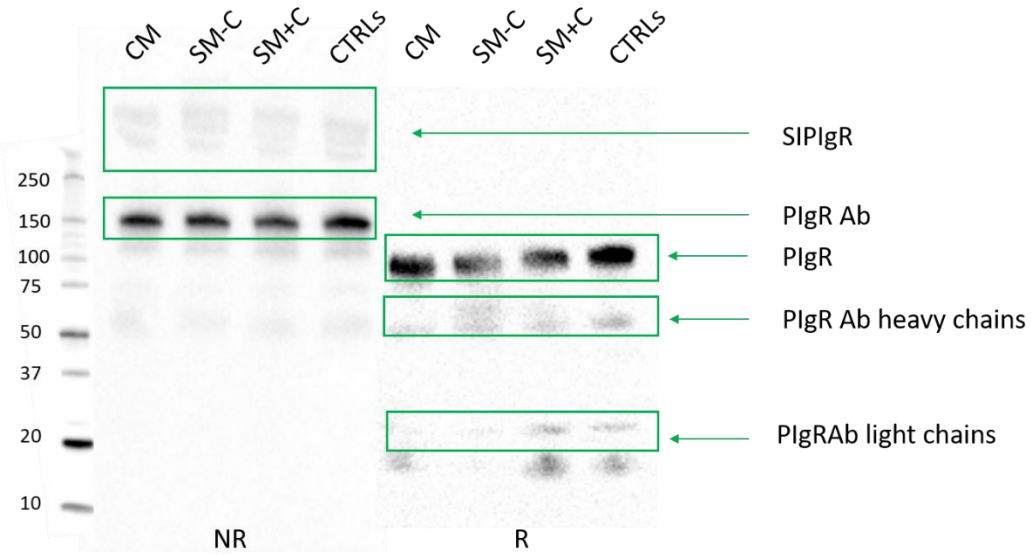
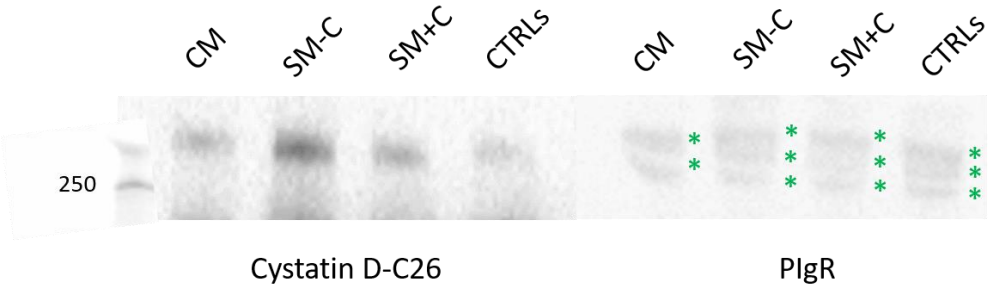


Figure 17: Comparison of high molecular weight salivary immunocomplexes in the four samples. On the left is shown the western blot of immunoprecipitated with cystatin D-C26 antibody detected with cystatin D-C26 antibody, while on the right is displayed the western blot of immunoprecipitated with PIgR antibody detected with PIgR antibody.



Next step would be the characterization of the SIPIgR using a bottom-up approach in order to highlight qualitative and/or quantitative variations between the groups of patients and the healthy subjects and among the patients. Moreover, since the SIPIgR appears to have two or more bands, it would be interesting separate them using a gel with a lower T (7.5% or 4%) and analyze them separately.

Discussion

This study allowed to confirm the presence of a high molecular weight salivary immuno-complex aggregating to cystatin D-C26 (SIC-D) and permitted to highlight qualitative and quantitative variation in protein composition between the patients' complex and the control one and among the two groups of patients, SM-C and SM+C. As mentioned above, mastocytosis is an abnormal infiltration of clonally derived mast cells in one or more different organs whose onset is difficult to establish. Moreover, diagnosis and distinction of the different subgroup of mastocytosis is usually tricky because of the difficulties in samples collection. Thus, qualitative and quantitative variations in salivary multi-protein could be helpful on the diagnosis and distinction of mastocytosis disease.

The immuno-blotting detection of the proteins co-immunoprecipitated with cystatin D-C26 was made in reducing and non-reducing conditions, and, based on the results obtained, the aggregation of the protein partners seemed to happen through disulfide bridges. However, another kind of non-covalent interaction have not been excluded by the applied experimental procedure. The SIC-D multi-protein complex could be defined an immunocomplex since the abundance of Igs and PIgR among its components. Some proteins have been identified in SIC-D of all the compared groups, such as IgA, IgG, annexin A1, S100A8, cystatin D-C26, lysozyme C and others. Desmoglein-1, lactotransferrin, S100A14 and SLPI were identified in patients but not in the control SIC-D immunocomplex. The result of SLPI found in SIC-D accorded to the data collected by HPLC-ESI-MS top-down analysis of acid soluble fraction of salivary samples. Some proteins were specifically found in just one of the three compared groups: lactoperoxidase was present only in control SIC-D immunocomplex; α-amylase 1, annexin A2 and cystatin A only in SM-C patients; cystatin B, S100A9, dermicidin and cystatin C only in SM+C patients.

A quantitative comparison was made by PD software elaboration on proteins found in all the groups, controls, SM-C and SM+C, and among these ones, 16 proteins were evidenced as changed in their abundance when the groups were compared, and they will be discussed on following.

The antibody anti-cystatin D used allowed immunodetecting and immunoprecipitating specifically the variant C-26 since recognized as epitope the portion of sequence containing the cysteine residue. This variant is not detectable in the acid fraction of salivary samples and

thus not directly analyzable by the HPLC-ESI-IT-MS top-down approach used to analyze, instead, the variant R-26. The immunodetection approach coupled to MS/MS analysis of the IPs obtained from the patients and healthy controls allowed revealing that cystatin D-C26 is present in whole saliva associated to protein partners in the formation of a complex (named SIC-D). Data obtained on C-26 variant in whole saliva were complementary to those ones obtained on R-26 variant on acid soluble fraction of salivary samples. Indeed, the variant R-26 was low abundant in SM+C patients, and almost absent in CM, but with a very similar level in SM-C and control groups. While the variant C-26 was found highly abundant in the SM+C patients with respect to both controls and SM-C groups, and the last one was the group with the lowest level of cystatin D-C26. PCA analysis evidenced that cystatin D-C26 was one of those proteins distinctive for the SM+C group. Unfortunately, data on whole saliva of CM patients could not be collected. From the HPLC-MS top-down analysis of the cystatin D-R26, it was possible to distinguish two N-terminally truncated proteoforms, naturally occurring in saliva, and highlight a different correlation with tryptase concentration in SM-C and SM+C patients. The bottom-up approach used to study the SIC-D complex did not produce information about an analogue PTM eventually occurring also on the C-26 variant. It is not known if the two variants are able to act differently against cathepsins, nor if they show different biological roles. As already discussed in the first part of the thesis, cystatin D is a potent cathepsin inhibitor, especially against cathepsin S, which is known to control MHC class II-mediated antigen presentation[224]. From this point of view, it is intriguing the ability of cystatin D to interact with the antigen presenting cells[140] and probably to regulate their maturation by blocking cathepsin S. The association of the only C-26 variant to the immunocomplex, characterized in this study, suggests that this protein could have a specific role in the immune response and that role may be altered in the mastocytosis patients.

Igs represent the first barrier against the pathogens in the oral cavity. They are produced by specific cells that reside in the oral mucosa[180] and are exported in saliva thanks to an epithelial transport in which PIgR and its secretory component play an important role[144]. Their main function is to block the microorganism through binding and/or agglutination, preventing mucosal adhesion[225,226]. Igs resulted to be one of the most abundant components in SIC-D, with differences in proteins' concentration between patients and control and among the patients. Both SM-C and SM+C immunocomplexes exhibited a low abundance of Ig component with respect to healthy controls except for Igγ1 heavy chain and heavy constant γ2, and Igγ1

light chain, which were very abundant in the patients. The immunocomplex of the SM-C group exhibited the lowest levels of all the Ig components with respect to the SM+C group. These results are in accord with the impaired status of the immune systems in mastocytosis patients. On the other hand, the increased levels of IgG1 and IgG2 in the immunocomplex of the patients could be related to the predisposition of the mastocytosis patients with systemic involvements to develop monoclonal gammopathy, an increased concentration of IgGs in serum[227,228].

PIgR is a transmembrane glycoprotein responsible of the epithelial transport of dimeric IgA and pentameric IgM from interstitial space to the saliva, thanks to the affinity of its secretory component (SC) to the J chain of IgA and IgM[182]. The process of transport is coupled to the proteolytic cleavage of SC from PIgR, which could be fused to the Igs[182] or stay unbound in the oral cavity[185,186]. The SIC-D of SM-C patients showed decreased levels of SC of PIgR with respect to both compared to SM+C patients and controls. The lower levels of SC may be reflect also the decreased levels in SIC-D of IgA 1 and 2.

DMBT-1, also known as salivary agglutinin or lung glycoprotein-340, is a scavenger-receptor glycoprotein rich in cysteine[229] belonging to a superfamily of secreted or membrane-bound proteins with SRCR domains. The protein is found in saliva and is expressed in salivary glands, ocular mucosal tissue, lacrimal glands, lung, trachea, gastrointestinal tract and macrophages[229]. It may play roles in mucosal defense system, cellular immune defense and epithelial differentiation. DMBT-1 acts as a scavenger receptor that recognizes and binds a wide range of oral pathogens, including bacteria and viruses[229–231]. The bound of pathogens increases their clearance from the oral cavity toward the stomach, leading to their acidic digestion[232]. DMBT-1 is able to interact also with mucosal defense proteins, such as IgA, surfactant proteins and MUC5B, and participate to the inflammation regulation by activating the complement cascade, stimulating the alveolar macrophage migration, and suppressing neutrophil oxidative stress[229]. The SIC-D of both SM-C and SM+C showed lower levels of DMBT-1 compared to the healthy controls, and SM-C patients exhibited the lowest level of this proteins. These results could be correlated to an increased susceptibility of these subjects to develop infections[233] and to their deficiency in the regulation of the inflammatory response.

Differently from the acid soluble fraction of salivary samples, where a significant decreasing of α-defensin 1 was observed in the patients, in the SIC-D immunocomplex α-defensin 1/2 was more abundant in patients, both SM-C and SM+C, than in healthy controls. It resulted, from

PCA analysis, that α-defensin 1/2, as well as cystatin D-C26, was one of the SIC-D partners distinctive for the SM+C patients. From the MS/MS analysis, it was not possible to discriminate between α-defensin 1 or 2, since the covered sequence was common to both, being the α-defensin 2 derived from proteolysis of α-defensin 1. Defensins are cationic peptides that possess antibacterial, antifungal and antiviral properties[234]. Their positive net charge permits to bind and to integrate in the microorganism cell membrane leading to the formation of ion channels, transmembrane pores and membrane rupture, resulting in the destruction of the microorganism[150,151]. Based on the results presented in this thesis, it appears that the loss of α-defensin 1/2 as free peptide active in saliva for the oral cavity defense may be associated to the recruitment of the peptide in a bound form as partner of the other proteins composing the SIC-D immunocomplex. The reason of this α-defensin 1/2 movement could not be explained in this study. A biological role similar to α-defensin 1/2 is exhibited by dermcidin, an anionic antimicrobial peptide expressed largely in epithelial cells but also in keratinocytes, dendric cells and in immune cells such as neutrophils, intestinal Paneth cells, NK cells, and B cells[235]. Dermcidin, such as other antimicrobial peptides, can act indirectly by augmenting innate and adaptive immune responses by regulating functions of the host's immune cells. An important component of these actions is the induction of inflammatory reactions by activating mast cells and thus recruitment of other immune cells[236]. Interestingly, the SIC-D of SM+C patients, but not the SM-C patients nor the controls, contained as partner this peptide.

Annexin A1, member of annexins family, is a calcium-dependent phospholipid binding protein[237]. Initially it was thought having a role in inflammatory processes since its inhibitor activity against phospholipase A2[237], recently it has been found that it can have a role in epithelial wound repair[238]. Even though this function has been highlighted in other tissues[239], this protein could have wound repair function also in the oral cavity, since it is often exposed to traumatic events that can lead to damages and inflammation[238,240]. Interestingly, only SM-C SIC-D complex showed the presence of annexin A2, which was not found in SM+C nor controls. Annexin A2 is a calcium regulated binding protein, which is involved in stress responses, mainly localized in the cytoplasm and plasma membrane[241] but revealed also in the nucleus[242,243]. Annexin A2 can be found as a monomer or as a heterotetrametric form, in which two molecules of annexin A2 are linked by a dimer of S100A10[244]. The cytoplasmatic hetero-tetramer has been suggested to be involved in the endocytic and exocytic vesicular transport, interaction with cell adhesion molecules regulation, and to function as a scaffold

protein in lipid raft organization[245]. The monomer form possess an N-terminal reactive cysteine residue (Cys-8) that is oxidized by H_2O_2 and subsequently reduced by the thioredoxin system, allowing annexin A2 to participate in multiple redox cycles[246]. Thus, its increased concentration in the SIC-D of SM-C of patients could be connected to its protective role against the ROS that typically increase in inflammatory processes.

S100A8 plays, together with S100A9, an important role in the immune response and in the regulation of inflammatory responses. It can also form a heterodimer with S100A9 and in recent study it was found to have a powerful chemotactic activity against neutrophils[214]. S100A8 is a ROS scavenger and it can decrease the concentration of several redox intermediates[166]. Both SM-C and SM+C SIC-D showed decreased abundance of S100A8 compared to control one. Also in this case, it was highlighted a complementary trend to that one observed in the acid soluble fraction of saliva, where S100A8 was abundant in patients with systemic symptoms, both SM-C and SM+C, but not in CM patients and healthy controls, while the protein was found at very low level as aggregated form in the SIC-D immunocomplex of the patients but not in healthy controls. S100A9 was detected only as a partner in SIC-D immunocomplex of SM+C patients, a result not well accorded with the one obtained in the salivary acid soluble fraction, where the protein was found at very low level in both SM-C and SM+C patients with respect to the controls and the CM patients.

Carbonic anhydrase is a zinc metalloenzyme that catalyze the formation of carbon dioxide and water from bicarbonate and hydrogen ions. It plays an important role in the regulation of acid-base balance in various tissue and biological fluids[247–249]. Isoenzyme 6 (CA-6) is secreted from parotid and submandibular glands, and it is important in the buffering system of saliva. It was demonstrated that CA-6 can contribute to neutralize plaque acids and to reduce enamel demineralization[250]. Moreover, CA-6 is present in the enamel pellicle and it was demonstrated that it can prevent caries accelerating the clearance of hydrogen ions and acid metabolic products in the local microenvironment of tooth surfaces[251]. Carbonic anhydrase has also been associated with type 2 cytokine response[252,253]. The results of this study demonstrated that CA-6 participated to the SIC-D formation and it was found particularly abundant in the immunocomplex of the patients when compared to the controls, mainly the SM+C group. It is interesting correlate these results with those one obtained by other research groups that demonstrated that carbonic anhydrase enzymes are involved in the regulation of mast cells

activity[254] and that they may be used as targets in potential therapeutic implications in the management of mastocytosis and type-2 inflammation[255].

The study carried out in this second part of the thesis allowed to characterize the interaction of several salivary proteins to associate in a multi-protein complex. The immunocomplex seemed to act as a fulcrum of aggregation of proteins and peptides playing roles as pro-inflammatory and regulatory of inflammation, in the innate and adaptive defenses, present in the oral cavity and ready to release or exchange components as needed from the conditions of the oral cavity.

The immunocomplex composition was partially different between the two groups of patients compared, those with systemic symptoms. Further studies are required to enlarge the number of patients enrolled for each group, in order to confirm the present results and to extend the characterization of the immunocomplex to the adult patients manifesting only the cutaneous symptoms. However, despite the small number of subjects under study, it was possible to highlight qualitative and quantitative variation concerning proteins and peptides involved in both the innate and acquired immune response, such as in the inflammation. These are all processes that undergo to dysregulation in mastocytosis, and probably the situation delineated in saliva at proteomic level may reflect a systemic status. The obtained results are promising and intriguing in the perspective to profiling the salivary proteome in different forms of the disease.

Acknowledgments

I would like to gratefully acknowledge Sardinian Regional Government for the financial support of my PhD scholarship (P.O.R. Sardegna F.S.E. - Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2014-2020 - Axis III Education and training, Thematic goal 10, Investment Priority 10ii), Specific goal 10.5., and the European Eramus+ Program for further financial support during the mobility period at Chemistry Department of Universidade de Aveiro (Portugal).

Me and my research group acknowledge the CeSAR (Centro Servizi d'Ateneo per la Ricerca) of the University of Cagliari, Italy for the High Resolution Mass Spectrometry experiments performed with Orbitrap Elite instrument.

Bibliography

1. Gri, G. *et al.* Mast Cell: An Emerging Partner in Immune Interaction. *Front. Immunol.* **3**, (2012).

2. Theoharides, T. C. *et al.* Mast cells and inflammation. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1822**, 21–33 (2012).

3. Puri, N. & Roche, P. A. Mast cells possess distinct secretory granule subsets whose exocytosis is regulated by different SNARE isoforms. *Proc. Natl. Acad. Sci.* **105**, 2580–2585 (2008).

4. Wasiuk, A., de Vries, V. C., Hartmann, K., Roers, A. & Noelle, R. J. Mast cells as regulators of adaptive immunity to tumours. *Clin. Exp. Immunol.* **155**, 140–146 (2009).

5. Kirshenbaum, a S. *et al.* Demonstration that human mast cells arise from a progenitor cell population that is CD34(+), c-kit(+), and expresses aminopeptidase N (CD13). *Blood* **94**, 2333–42 (1999).

6. Kirshenbaum, A. S., Kessler, S. W., Goff, J. P. & Metcalfe, D. D. Demonstration of the origin of human mast cells from CD34+ bone marrow progenitor cells. *J. Immunol.* **146**, 1410 LP – 1415 (1991).

7. Metcalfe, D. D., Mekori, J. A. & Rottem, M. Mast cell ontogeny and apoptosis. *Exp. Dermatol.* **4**, 227–230 (2018).

8. Komi, D. E. A., Rambasek, T. & Wöhrl, S. Mastocytosis: from a Molecular Point of View. *Clin. Rev. Allergy Immunol.* **54**, 397–411 (2018).

9. Lennartsson, J. & Rönnstrand, L. Stem Cell Factor Receptor/c-Kit: From Basic Science to Clinical Implications. *Physiol. Rev.* **92**, 1619–1649 (2012).

10. Heissig, B. *et al.* Recruitment of stem and progenitor cells from the bone marrow niche requires MMP-9 mediated release of kit-ligand. *Cell* **109**, 625–37 (2002).

11. Longley, B. J. *et al.* Chymase cleavage of stem cell factor yields a bioactive, soluble product. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 9017–21 (1997).

12. Zou, J. *et al.* Catalytic Activity of Human ADAM33. *J. Biol. Chem.* **279**, 9818–9830 (2004).

13. Zhang, Z., Zhang, R., Joachimiak, A., Schlessinger, J. & Kong, X. P. Crystal structure of human stem cell factor: implication for stem cell factor receptor dimerization and activation. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7732–7 (2000).

14. Morales, J. K., Falanga, Y. T., Depcrynski, A., Fernando, J. & Ryan, J. J. Mast cell homeostasis and the JAK–STAT pathway. *Genes Immun.* **11**, 599–608 (2010).

15. Theoharides, T. C., Valent, P. & Akin, C. Mast Cells, Mastocytosis, and Related Disorders. *N. Engl. J. Med.* **373**, 163–172 (2015).

16. Cohen, S. S. *et al.* Epidemiology of systemic mastocytosis in Denmark. *Br. J. Haematol.* **166**, 521–528 (2014).

17. Brockow, K. Epidemiology, Prognosis, and Risk Factors in Mastocytosis. *Immunol. Allergy Clin. North Am.* **34**, 283–295 (2014).

18. Akin, C. & Valent, P. Diagnostic criteria and classification of mastocytosis in 2014. *Immunology and Allergy Clinics of North America* **34**, 207–218 (2014).

19. Wolff, K., Komar, M. & Petzelbauer, P. Clinical and histopathological aspects of cutaneous mastocytosis. *Leuk. Res.* **25**, 519–528 (2001).

20. Valent, P., Akin, C. & Metcalfe, D. D. Mastocytosis: 2016 updated WHO classification and novel emerging treatment concepts. *Blood* **129**, 1420–1427 (2017).

21. Hartmann, K. *et al.* Cutaneous manifestations in patients with mastocytosis: Consensus report of the European Competence Network on Mastocytosis; the American Academy of Allergy, Asthma & Immunology; and the European Academy of Allergology and Clinical Immunology. *J. Allergy Clin. Immunol.* **137**, 35–45 (2016).

22. Carter, M. C., Metcalfe, D. D. & Komarow, H. D. Mastocytosis. *Immunol. Allergy Clin. North Am.* **34**, 10.1016/j.iac.2013.09.001 (2014).

23. Sperr, W. R., Horny, H.-P. & Valent, P. Spectrum of Associated Clonal Hematologic Non-Mast Cell Lineage Disorders Occurring in Patients with Systemic Mastocytosis. *Int. Arch. Allergy Immunol.* **127**, 140–142 (2002).

24. Brockow, K. & Metcalfe, D. D. Mastocytosis. in *Anaphylaxis* 110–124 (KARGER, 2010). doi:10.1159/000315946

25. Kettelhut, B. V. & Metcalfe, D. D. Pediatric Mastocytosis. *J. Invest. Dermatol.* **96**, S15–S18 (1991).

26. Kuint, J. *et al.* c-kit activating mutation in a neonate with in-utero presentation of systemic mastocytosis associated with myeloproliferative disorder. *Br. J. Haematol.* **106**, 838–839 (1999).

27. Longley, B. J. *et al.* Activating and dominant inactivating c-KIT catalytic domain mutations in distinct clinical forms of human mastocytosis. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 1609–1614 (1999).

28. Laine, E., Chauvot de Beauchêne, I., Perahia, D., Auclair, C. & Tchertanov, L. Mutation D816V Alters the Internal Structure and Dynamics of c-KIT Receptor Cytoplasmic Region: Implications for Dimerization and Activation Mechanisms. *PLoS Comput. Biol.* **7**, e1002068 (2011).

29. Frost, M. J., Ferrao, P. T., Hughes, T. P. & Ashman, L. K. Juxtamembrane mutant V560GKit is more sensitive to Imatinib (STI571) compared with wild-type c-kit whereas the kinase domain mutant D816VKit is resistant. *Mol. Cancer Ther.* **1**, 1115–24 (2002).

30. Castells, M., Metcalfe, D. D. & Escribano, L. Diagnosis and Treatment of Cutaneous Mastocytosis in Children. *Am. J. Clin. Dermatol.* **12**, 259–270 (2011).

31. Tefferi, A. & Pardanani, A. Systemic mastocytosis: current concepts and treatment advances. *Curr. Hematol. Rep.* **3**, 197–202 (2004).

32. Escribano, L., Akin, C., Castells, M. & Schwartz, L. B. Current options in the treatment of mast cell mediator-related symptoms in mastocytosis. *Inflamm. Allergy Drug Targets* **5**, 61–77 (2006).

33. Brockow, K., Akin, C., Huber, M. & Metcalfe, D. D. Assessment of the extent of cutaneous involvement in children and adults with mastocytosis: relationship to symptomatology, tryptase levels, and bone marrow pathology. *J. Am. Acad. Dermatol.* **48**, 508–16 (2003).

34. Brockow, K. *et al.* Regression of urticaria pigmentosa in adult patients with systemic mastocytosis: correlation with clinical patterns of disease. *Arch. Dermatol.* **138**, 785–90 (2002).

35. Cherner, J. A. *et al.* Gastrointestinal Dysfunction in Systemic Mastocytosis. *Gastroenterology* **95**, 657–667 (1988).

36. Sokol, H. *et al.* Gastrointestinal manifestations in mastocytosis: A study of 83 patients. *J. Allergy Clin. Immunol.* **132**, 866-873.e3 (2013).

37. Horan, R. F. & Austen, K. F. Systemic Mastocytosis: Retrospective Review of a Decade's Clinical Experience at the Brigham and Women's Hospital. *J. Invest. Dermatol.* **96**, S5–S14 (1991).

38. Gupta, R., Bain, B. J. & Knight, C. L. Cytogenetic and Molecular Genetic Abnormalities in Systemic Mastocytosis. *Acta Haematol.* **107**, 123–128 (2002).

39. Travis, W. D. & Li, C. Y. Pathology of the lymph node and spleen in systemic mast cell disease. *Mod. Pathol.* **1**, 4–14 (1988).

40. Mican, J. Hepatic involvement in mastocytosis: Clinicopathologic correlations in 41 cases. *Hepatology* **22**, 1163–1170 (1995).

41. Valent, P. Diagnostic Evaluation and Classification of Mastocytosis. *Immunol. Allergy Clin. North Am.* **26**, 515–534 (2006).

42. Garriga, M. M., Friedman, M. M. & Metcalfe, D. D. A survey of the number and distribution of mast cells in the skin of patients with mast cell disorders. *J. Allergy Clin. Immunol.* **82**, 425–432 (1988).

43. Morgado, J. M. T. *et al.* Immunophenotyping in systemic mastocytosis diagnosis: 'CD25 positive' alone is more informative than the 'CD25 and/or CD2' WHO criterion. *Mod. Pathol.* **25**, 516–521 (2012).

44. Hahn, H. P. & Hornick, J. L. Immunoreactivity for CD25 in Gastrointestinal Mucosal Mast Cells is Specific for Systemic Mastocytosis. *Am. J. Surg. Pathol.* **31**, 1669–1676 (2007).

45. Gschwandtner, M. *et al.* Proteome analysis identifies L1CAM/CD171 and DPP4/CD26 as novel markers of human skin mast cells. *Allergy* **72**, 85–97 (2017).

46. Valent, P. *et al.* Mastocytosis: Pathology, genetics, and current options for therapy. *Leuk. Lymphoma* **46**, 35–48 (2005).

47. Valent, P. *et al.* Aggressive systemic mastocytosis and related mast cell disorders:

current treatment options and proposed response criteria. *Leuk. Res.* **27**, 635–41 (2003).

48. Söderholm, J. D. Mast Cells and Mastocytosis. *Dig. Dis.* **27**, 129–136 (2009).

49. Niedoszytko, M., de Monchy, J., van Doormaal, J. J., Jassem, E. & Oude Elberink, J. N. G. Mastocytosis and insect venom allergy: diagnosis, safety and efficacy of venom immunotherapy. *Allergy* **64**, 1237–1245 (2009).

50. de Olano, D. G. *et al.* Safety and effectiveness of immunotherapy in patients with indolent systemic mastocytosis presenting with Hymenoptera venom anaphylaxis. *J. Allergy Clin. Immunol.* **121**, 519–526 (2008).

51. Valent, P. *et al.* Standards and standardization in mastocytosis: Consensus Statements on Diagnostics, Treatment Recommendations and Response Criteria. *Eur. J. Clin. Invest.* **37**, 435–453 (2007).

52. Lasota, J. Not all c-kit mutations can be corrected by imatinib. *Lab. Invest.* **87**, 317 (2007).

53. Dubreuil, P. *et al.* Masitinib (AB1010), a Potent and Selective Tyrosine Kinase Inhibitor Targeting KIT. *PLoS One* **4**, e7258 (2009).

54. Pan, J. *et al.* Sensitivity of human cells bearing oncogenic mutant kit isoforms to the novel tyrosine kinase inhibitor INNO-406. *Cancer Sci.* **98**, 1223–1225 (2007).

55. Shah, N. P. Dasatinib (BMS-354825) inhibits KITD816V, an imatinib-resistant activating mutation that triggers neoplastic growth in most patients with systemic mastocytosis. *Blood* **108**, 286–291 (2006).

56. Gleixner, K. V. *et al.* Synergistic growth-inhibitory effects of two tyrosine kinase inhibitors, dasatinib and PKC412, on neoplastic mast cells expressing the D816V-mutated oncogenic variant of KIT. *Haematologica* **92**, 1451–1459 (2007).

57. Akin, C. *et al.* Effects of tyrosine kinase inhibitor STI571 on human mast cells bearing wild-type or mutated c-kit. *Exp. Hematol.* **31**, 686–92 (2003).

58. Gleixner, K. V *et al.* PKC412 inhibits in vitro growth of neoplastic human mast cells expressing the D816V-mutated variant of KIT: comparison with AMN107, imatinib, and

cladribine (2CdA) and evaluation of cooperative drug effects. *Blood* **107**, 752 LP – 759 (2006).

59. Purtill, D. *et al.* Dasatinib therapy for systemic mastocytosis: four cases. *Eur. J. Haematol.* **80**, 456–458 (2008).

60. Gotlib, J. Activity of the tyrosine kinase inhibitor PKC412 in a patient with mast cell leukemia with the D816V KIT mutation. *Blood* **106**, 2865–2870 (2005).

61. Gleixner, K. V. *et al.* KIT-D816V-independent oncogenic signaling in neoplastic cells in systemic mastocytosis: role of Lyn and Btk activation and disruption by dasatinib and bosutinib. *Blood* **118**, 1885–1898 (2011).

62. Caplan, R. M. The Natural Course of Urticaria Pigmentosa. *Arch. Dermatol.* **87**, 146 (1963).

63. Kiszewski, A., Duran-Mckinster, C., Orozco-Covarrubias, L., Gutierrez-Castrellon, P. & Ruiz-Maldonado, R. Cutaneous mastocytosis in children: a clinical analysis of 71 cases. *J. Eur. Acad. Dermatology Venereol.* **18**, 285–290 (2004).

64. Lanternier, F. *et al.* Phenotypic and Genotypic Characteristics of Mastocytosis According to the Age of Onset. *PLoS One* **3**, e1906 (2008).

65. Ben-Amitai, D., Metzker, A. & Cohen, H. A. Pediatric cutaneous mastocytosis: a review of 180 patients. *Isr. Med. Assoc. J.* **7**, 320–2 (2005).

66. Cruse, G., Metcalfe, D. D. & Olivera, A. Functional Deregulation of KIT. *Immunol. Allergy Clin. North Am.* **34**, 219–237 (2014).

67. Addona, T. A. *et al.* A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease. *Nat. Biotechnol.* **29**, 635–643 (2011).

68. Whiteaker, J. R. *et al.* A targeted proteomics–based pipeline for verification of biomarkers in plasma. *Nat. Biotechnol.* **29**, 625–634 (2011).

69. Hedl, T. J. *et al.* Proteomics Approaches for Biomarker and Drug Target Discovery in ALS and FTD. *Front. Neurosci.* **13**, (2019).

70. Cui, W., Rohrs, H. W. & Gross, M. L. Top-down mass spectrometry: Recent

developments, applications and perspectives. *Analyst* **136**, 3854 (2011).

71. Pandeswari, P. B. & Sabareesh, V. Middle-down approach: a choice to sequence and characterize proteins/proteomes by mass spectrometry. *RSC Adv.* **9**, 313–344 (2019).

72. Meyer, K. Primary sensory cortices, top-down projections and conscious experience. *Prog. Neurobiol.* **94**, 408–417 (2011).

73. Zhang, C.-Z. *et al.* Saliva in the diagnosis of diseases. *Int. J. Oral Sci.* **8**, 133–137 (2016).

74. Khurshid, Z. *et al.* Human Saliva Collection Devices for Proteomics: An Update. *Int. J. Mol. Sci.* **17**, 846 (2016).

75. Cabras, T. *et al.* Top-down analytical platforms for the characterization of the human salivary proteome. *Bioanalysis* **6**, 563–581 (2014).

76. Bandhakavi, S., Stone, M. D., Onsongo, G., Van Riper, S. K. & Griffin, T. J. A Dynamic Range Compression and Three-Dimensional Peptide Fractionation Analysis Platform Expands Proteome Coverage and the Diagnostic Potential of Whole Saliva. *J. Proteome Res.* **8**, 5590–5600 (2009).

77. Siqueira, W. L., Salih, E., Wan, D. L., Helmerhorst, E. J. & Oppenheim, F. G. Proteome of Human Minor Salivary Gland Secretion. *J. Dent. Res.* **87**, 445–450 (2008).

78. Ngo, L. H. *et al.* Mass Spectrometric Analyses of Peptides and Proteins in Human Gingival Crevicular Fluid. *J. Proteome Res.* **9**, 1683–1693 (2010).

79. Castagnola, M. *et al.* Top-down platform for deciphering the human salivary proteome. *J. Matern. Neonatal Med.* **25**, 27–43 (2012).

80. Marsh, P. D., Do, T., Beighton, D. & Devine, D. A. Influence of saliva on the oral microbiota. *Periodontol. 2000* **70**, 80–92 (2016).

81. Carpenter, G. H. The Secretion, Components, and Properties of Saliva. *Annu. Rev. Food Sci. Technol.* **4**, 267–276 (2013).

82. Castagnola, M., Cabras, T., Vitali, A., Sanna, M. T. & Messana, I. Biotechnological implications of the salivary proteome. *Trends Biotechnol.* **29**, 409–418 (2011).

83. Castagnola, M. *et al.* The Surprising Composition of the Salivary Proteome of Preterm

Human Newborn. *Mol. Cell. Proteomics* **10**, M110.003467 (2011).

84. Amado, F., Lobo, M. J. C., Domingues, P., Duarte, J. A. & Vitorino, R. Salivary peptidomics. *Expert Rev. Proteomics* **7**, 709–721 (2010).

85. Zhang, A., Sun, H., Wang, P. & Wang, X. Salivary proteomics in biomedical research. *Clin. Chim. Acta* **415**, 261–265 (2013).

86. Cabras, T. *et al.* Age-Dependent Modifications of the Human Salivary Secretory Protein Complex. *J. Proteome Res.* **8**, 4126–4134 (2009).

87. Messana, I. *et al.* Chrono-Proteomics of Human Saliva: Variations of the Salivary Proteome during Human Development. *J. Proteome Res.* **14**, 1666–1677 (2015).

88. Wang, Q., Yu, Q., Lin, Q. & Duan, Y. Emerging salivary biomarkers by mass spectrometry. *Clin. Chim. Acta* **438**, 214–221 (2015).

89. Schafer, C. A. *et al.* Saliva Diagnostics: Utilizing Oral Fluids to Determine Health Status. in 88–98 (2014). doi:10.1159/000358791

90. Cuevas-Córdoba, B. & Santiago-García, J. Saliva: A Fluid of Study for OMICS. *Omi. A J. Integr. Biol.* **18**, 87–97 (2014).

91. Jou, Y.-J. *et al.* Proteomic identification of salivary transferrin as a biomarker for early detection of oral cancer. *Anal. Chim. Acta* **681**, 41–48 (2010).

92. Shintani, S., Hamakawa, H., Ueyama, Y., Hatori, M. & Toyoshima, T. Identification of a truncated cystatin SA-I as a saliva biomarker for oral squamous cell carcinoma using the SELDI ProteinChip platform. *Int. J. Oral Maxillofac. Surg.* **39**, 68–74 (2010).

93. Hu, S. & Wong, D. T. Oral cancer proteomics. *Curr. Opin. Mol. Ther.* **9**, 467–76 (2007).

94. Cabras, T. *et al.* Significant Modifications of the Salivary Proteome Potentially Associated with Complications of Down Syndrome Revealed by Top-down Proteomics. *Mol. Cell. Proteomics* **12**, 1844–1852 (2013).

95. Castagnola, M. *et al.* Hypo-Phosphorylation of Salivary Peptidome as a Clue to the Molecular Pathogenesis of Autism Spectrum Disorders. *J. Proteome Res.* **7**, 5327–5332 (2008).

96. Sanna, M. *et al.* The salivary proteome profile in patients affected by SAPHO syndrome characterized by a top-down RP-HPLC-ESI-MS platform. *Mol. Biosyst.* **11**, 1552–1562 (2015).

97. Cabras, T. *et al.* Proteomic investigation of whole saliva in Wilson's disease. *J. Proteomics* **128**, 154–163 (2015).

98. Beeley, J. Basic proline-rich proteins: multifunctional defence molecules? *Oral Dis.* **7**, 69–70 (2001).

99. Manconi, B. *et al.* The intriguing heterogeneity of human salivary proline-rich proteins. *J. Proteomics* **134**, 47–56 (2016).

100. Messana, I., Inzitari, R., Fanali, C., Cabras, T. & Castagnola, M. Facts and artifacts in proteomics of body fluids. What proteomics of saliva is telling us? *J. Sep. Sci.* **31**, 1948–1963 (2008).

101. Cabras, T. *et al.* HPLC–MS characterization of cyclo-statherin Q-37, a specific cyclization product of human salivary statherin generated by transglutaminase 2. *J. Sep. Sci.* **29**, 2600–2608 (2006).

102. Lyons, K. M., Stein, J. H. & Smithies, O. Length polymorphisms in human proline-rich protein genes generated by intragenic unequal crossing over. *Genetics* **120**, 267–278 (1988).

103. Azen, E. A., Amberger, E., Fisher, S., Prakobphol, A. & Niece, R. L. PRB1, PRB2, and PRB4 coded polymorphisms among human salivary concanavalin-A binding, II-1, and Po proline-rich proteins. *Am. J. Hum. Genet.* **58**, 143–53 (1996).

104. Lyons, K. M., Azen, E. A., Goodman, P. A. & Smithies, O. Many protein products from a few loci: assignment of human salivary proline-rich proteins to specific loci. *Genetics* **120**, 255–65 (1988).

105. Bennick, A. Interaction of Plant Polyphenols with Salivary Proteins. *Crit. Rev. Oral Biol. Med.* **13**, 184–196 (2002).

106. Cabras, T. *et al.* Responsiveness to 6-n-Propylthiouracil (PROP) Is Associated with Salivary Levels of Two Specific Basic Proline-Rich Proteins in Humans. *PLoS One* **7**,

e30962 (2012).

107. Melis, M. *et al.* 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**, e0131104 (2015).

108. Messana, I. *et al.* Trafficking and Postsecretory Events Responsible for the Formation of Secreted Human Salivary Peptides: A Proteomics Approach. *Mol. Cell. Proteomics* **7**, 911–926 (2008).

109. Jensen, J. L., Lamkin, M. S., Troxler, R. F. & Oppenheim, F. G. Multiple forms of statherin in human salivary secretions. *Arch. Oral Biol.* **36**, 529–534 (1991).

110. Inzitari, R. *et al.* Detection in human saliva of different statherin and P-B fragments and derivatives. *Proteomics* **6**, 6370–6379 (2006).

111. Schlesinger, D. H. & Hay, D. I. Complete covalent structure of statherin, a tyrosine rich acidic peptide which inhibits calcium phosphate precipitation from human parotid saliva. *J. Biol. Chem.* **252**, 1689–1695 (1977).

112. Hay, D. I., Smith, D. J., Schluckebier, S. K. & Moreno, E. C. Basic Biological Sciences Relationship between Concentration of Human Salivary Statherin and Inhibition of Calcium Phosphate Precipitation in Stimulated Human Parotid Saliva. *J. Dent. Res.* **63**, 857–863 (1984).

113. Cabras, T. *et al.* Tyrosine Polysulfation of Human Salivary Histatin 1. A Post-Translational Modification Specific of the Submandibular Gland. *J. Proteome Res.* **6**, 2472–2480 (2007).

114. Oudhoff, M. J. *et al.* The role of salivary histatin and the human cathelicidin LL-37 in wound healing and innate immunity. *Biol. Chem.* **391**, (2010).

115. Torres, P., Castro, M., Reyes, M. & Torres, V. Histatins, wound healing, and cell migration. *Oral Dis.* **24**, 1150–1160 (2018).

116. Castagnola, M. *et al.* A Cascade of 24 Histatins (Histatin 3 Fragments) in Human Saliva: Suggestion for a Pre-Secretory Sequential Cleavage Pathway. *J. Biol. Chem.* **279**, 41436–41443 (2004).

117. Zhou, A., Webb, G., Zhu, X. & Steiner, D. F. Proteolytic Processing in the Secretory Pathway. *J. Biol. Chem.* **274**, 20745–20748 (1999).

118. Dickinson, D. P. Cysteine peptidases of mammals: Their biological roles and potential effects in the oral cavity and other tissues in health and disease. *Crit. Rev. Oral Biol. Med.* **13**, 238–275 (2002).

119. Räsänen, O., Järvinen, M. & Rinne, A. Localization of the human SH-protease inhibitor in the epidermis. *Acta Histochem.* **63**, 193-IN2 (1978).

120. Järvinen, M., Pernu, H., Rinne, A., Hopsu-Havu, V. K. & Altonen, M. Localization of three inhibitors of cysteineproteinases in the human oral mucosa. *Acta Histochem.* **73**, 279-IN2 (1983).

121. Davies, M. E. & Barrett, A. J. Immunolocalization of human cystatins in neutrophils and lymphocytes. *Histochemistry* **80**, 373–377 (1984).

122. Rinne, A. *et al.* Demonstration of immunoreactive acid cysteine-proteinase inhibitor in reticulum cells of lymph node germinal centres. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **43**, 121–126 (1983).

123. Järvinen, M. & Rinne, A. Human spleen cysteineproteinase inhibitor. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **708**, 210–217 (1982).

124. Hopsu-Havu, V. K., Joronen, I. A., ärvinen, M. J., Rinne, A. & Aalto, M. Cysteine proteinase inhibitors produced by mononuclear phagocytes. *Cell Tissue Res.* **236**, 161–164 (1984).

125. Manconi, B. *et al.* Salivary Cystatins: Exploring New Post-Translational Modifications and Polymorphisms by Top-Down High-Resolution Mass Spectrometry. *J. Proteome Res.* **16**, 4196–4207 (2017).

126. Vasilopoulos, Y. *et al.* A nonsynonymous substitution of cystatin A, a cysteine protease inhibitor of house dust mite protease, leads to decreased mRNA stability and shows a significant association with atopic dermatitis. *Allergy* **62**, 514–519 (2007).

127. Lin, Y.-Y. *et al.* Tissue Levels of Stefin A and Stefin B in Hepatocellular Carcinoma. *Anat. Rec.* **299**, 428–438 (2016).

128. Blaydon, D. C. *et al.* Mutations in CSTA, Encoding Cystatin A, Underlie Exfoliative Ichthyosis and Reveal a Role for This Protease Inhibitor in Cell-Cell Adhesion. *Am. J. Hum. Genet.* **89**, 564–571 (2011).

129. Muttardi, K., Nitoiu, D., Kelsell, D. P., O'Toole, E. A. & Batta, K. Acral peeling skin syndrome associated with a novel CSTA gene mutation. *Clin. Exp. Dermatol.* **41**, 394–398 (2016).

130. Cabras, T. *et al.* 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. *J. Proteomics* **75**, 908–913 (2012).

131. Yang, F. *et al.* Cystatin B inhibition of TRAIL-induced apoptosis is associated with the protection of FLIPL from degradation by the E3 ligase itch in human melanoma cells. *Cell Death Differ.* **17**, 1354–1367 (2010).

132. Lehtinen, M. K. *et al.* Cystatin B Deficiency Sensitizes Neurons to Oxidative Stress in Progressive Myoclonus Epilepsy, EPM1. *J. Neurosci.* **29**, 5910–5915 (2009).

133. Škerget, K. *et al.* Interaction between Oligomers of Stefin B and Amyloid-β in Vitro and in Cells. *J. Biol. Chem.* **285**, 3201–3210 (2010).

134. Shah, A. & Bano, B. Cystatins in health and diseases. *Int. J. Pept. Res. Ther.* **15**, 43–48 (2009).

135. Hall, A., Håkansson, K., Mason, R. W., Grubb, A. & Abrahamson, M. Structural Basis for the Biological Specificity of Cystatin C. *J. Biol. Chem.* **270**, 5115–5121 (1995).

136. Jurczak, P., Groves, P., Szymanska, A. & Rodziewicz-Motowidlo, S. Human cystatin C monomer, dimer, oligomer, and amyloid structures are related to health and disease. *FEBS Lett.* **590**, 4192–4201 (2016).

137. Balbín, M. *et al.* A sequence variation in the human cystatin D gene resulting in an amino acid (Cys/Arg) polymorphism at the protein level. *Hum. Genet.* **90**, 668–669 (1993).

138. Alvarez-Fernandez, M., Liang, Y.-H., Abrahamson, M. & Su, X.-D. Crystal Structure of Human Cystatin D, a Cysteine Peptidase Inhibitor with Restricted Inhibition Profile. *J. Biol. Chem.* **280**, 18221–18228 (2005).

139. Álvarez-Díaz, S. *et al.* Cystatin D is a candidate tumor suppressor gene induced by vitamin D in human colon cancer cells. *J. Clin. Invest.* **119**, 2343–2358 (2009).

140. Nashida, T. *et al.* Antigen-presenting cells in parotid glands contain cystatin D originating from acinar cells. *Arch. Biochem. Biophys.* **530**, 32–39 (2013).

141. Ferrer-Mayorga, G. *et al.* Cystatin D Locates in the Nucleus at Sites of Active Transcription and Modulates Gene and Protein Expression. *J. Biol. Chem.* **290**, 26533–26548 (2015).

142. Liori, B. *et al.* Interactors of human salivary Cystatin D explored by Immunoprecipitation coupled to Mass Spectrometry. in *XII ItPA National Congress (Lecce, June 2017)* Page 90 (Abstract Book) (2017).

143. Doumas, S., Kolokotronis, A. & Stefanopoulos, P. Anti-Inflammatory and Antimicrobial Roles of Secretory Leukocyte Protease Inhibitor. *Infect. Immun.* **73**, 1271–1274 (2005).

144. Fábián, T. K., Hermann, P., Beck, A., Fejérdy, P. & Fábián, G. Salivary Defense Proteins: Their Network and Role in Innate and Acquired Oral Immunity. *Int. J. Mol. Sci.* **13**, 4295–4320 (2012).

145. Swain, N., Pathak, J. & Hosalkar, R. M. Defensin. in *Encyclopedia of Signaling Molecules* 1339–1345 (Springer International Publishing, 2018). doi:10.1007/978-3-319-67199-4_102004

146. Goebel, C., Mackay, L. ., Vickers, E. . & Mather, L. . Determination of defensin HNP-1, HNP-2, and HNP-3 in human saliva by using LC/MS. *Peptides* **21**, 757–765 (2000).

147. Pisano, E. *et al.* Peptides of human gingival crevicular fluid determined by HPLC-ESI-MS. *Eur. J. Oral Sci.* **113**, 462–468 (2005).

148. Ganz, T. & Lehrer, R. I. Defensins. *Curr. Opin. Immunol.* **6**, 584–589 (1994).

149. Lehrer, R. I. & Lu, W. α-Defensins in human innate immunity. *Immunol. Rev.* **245**, 84–112 (2012).

150. Wiesner, J. & Vilcinskas, A. Antimicrobial peptides: The ancient arm of the human immune system. *Virulence* **1**, 440–464 (2010).

151. Brogden, K. A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?

Nat. Rev. Microbiol. **3**, 238–250 (2005).

152. Grant, D. S. *et al.* Matrigel induces thymosin beta 4 gene in differentiating endothelial cells. *J. Cell Sci.* **108 (Pt 1**, 3685–3694 (1995).

153. Malinda, K. M., Goldstein, A. L. & Kleinman, H. K. Thymosin beta 4 stimulates directional migration of human umbilical vein endothelial cells. *FASEB J.* **11**, 474–481 (1997).

154. Goldstein, A. L., Hannappel, E., Sosne, G. & Kleinman, H. K. Thymosin β4: a multi-functional regenerative peptide. Basic properties and clinical applications. *Expert Opin. Biol. Ther.* **12**, 37–51 (2012).

155. Bock-Marquette, I., Saxena, A., White, M. D., Michael DiMaio, J. & Srivastava, D. Thymosin β4 activates integrin-linked kinase and promotes cardiac cell migration, survival and cardiac repair. *Nature* **432**, 466–472 (2004).

156. Smart, N. *et al.* Thymosin β4 induces adult epicardial progenitor mobilization and neovascularization. *Nature* **445**, 177–182 (2007).

157. Sosne, G. *et al.* Thymosin Beta 4 Promotes Corneal Wound Healing and Modulates Inflammatory Mediators in vivo. *Exp. Eye Res.* **72**, 605–608 (2001).

158. Smart, N., Rossdeutsch, A. & Riley, P. R. Thymosin β4 and angiogenesis: modes of action and therapeutic potential. *Angiogenesis* **10**, 229–241 (2007).

159. Badamchian, M. *et al.* Identification and Quantification of Thymosin beta4 in Human Saliva and Tears. *Ann. N. Y. Acad. Sci.* **1112**, 458–465 (2007).

160. Inzitari, R. *et al.* HPLC-ESI-MS analysis of oral human fluids reveals that gingival crevicular fluid is the main source of oral thymosins β4 and β10. *J. Sep. Sci.* **32**, 57–63 (2009).

161. Nemolato, S. *et al.* Thymosin β4 and β10 Levels in Pre-Term Newborn Oral Cavity and Foetal Salivary Glands Evidence a Switch of Secretion during Foetal Development. *PLoS One* **4**, e5109 (2009).

162. Reti, R., Kwon, E., Qiu, P., Wheater, M. & Sosne, G. Thymosin beta4 is cytoprotective in human gingival fibroblasts. *Eur. J. Oral Sci.* **116**, 424–430 (2008).

163. Ji, Y. F., Huang, H., Jiang, F., Ni, R. Z. & Xiao, M. B. S100 family signaling network and

related proteins in pancreatic cancer (review). *Int. J. Mol. Med.* **33**, 769–776 (2014).

164. Marenholz, I., Lovering, R. C. & Heizmann, C. W. An update of the S100 nomenclature. *Biochim. Biophys. Acta - Mol. Cell Res.* **1763**, 1282–1283 (2006).

165. Donato, R. S100: a multigenic family of calcium-modulated proteins of the EF-hand type with intracellular and extracellular functional roles. *Int. J. Biochem. Cell Biol.* **33**, 637–668 (2001).

166. Lim, S. Y., Raftery, M. J., Goyette, J., Hsu, K. & Geczy, C. L. Oxidative modifications of S100 proteins: functional regulation by redox. *J. Leukoc. Biol.* **86**, 577–587 (2009).

167. Goyette, J. & Geczy, C. L. Inflammation-associated S100 proteins: new mechanisms that regulate function. *Amino Acids* **41**, 821–842 (2011).

168. Santamaria-Kisiel, L., Rintala-Dempsey, A. C. & Shaw, G. S. Calcium-dependent and -independent interactions of the S100 protein family. *Biochem. J.* **396**, 201–214 (2006).

169. Andrassy, M. *et al.* Posttranslationally Modified Proteins as Mediators of Sustained Intestinal Inflammation. *Am. J. Pathol.* **169**, 1223–1237 (2006).

170. Zimmer, D. B., Wright Sadosky, P. & Weber, D. J. Molecular mechanisms of S100-target protein interactions. *Microsc. Res. Tech.* **60**, 552–559 (2003).

171. Alowami, S., Qing, G., Emberley, E., Snell, L. & Watson, P. H. Psoriasin (S100A7) expression is altered during skin tumorigenesis. *BMC Dermatol.* **3**, 1 (2003).

172. Eckert, R. L. *et al.* S100 Proteins in the Epidermis. *J. Invest. Dermatol.* **123**, 23–33 (2004).

173. Strupat, K., Rogniaux, H., Dorsselaer, A., Roth, J. & Vogl, T. Calcium-induced noncovalently linked tetramers of MRP8 and MRP14 are confirmed by electrospray ionization-mass analysis. *J. Am. Soc. Mass Spectrom.* **11**, 780–788 (2000).

174. Harrison, C. A. *et al.* Oxidation Regulates the Inflammatory Properties of the Murine S100 Protein S100A8. *J. Biol. Chem.* **274**, 8561–8569 (1999).

175. Raftery, M. J., Yang, Z., Valenzuela, S. M. & Geczy, C. L. Novel Intra- and Inter-molecular Sulfinamide Bonds in S100A8 Produced by Hypochlorite Oxidation. *J. Biol. Chem.* **276**, 33393–33401 (2001).

176. Sroussi, H. Y., Berline, J., Dazin, P., Green, P. & Palefsky, J. M. S100A8 Triggers Oxidation-sensitive Repulsion of Neutrophils. *J. Dent. Res.* **85**, 829–833 (2006).

177. Sroussi, H. Y., Berline, J. & Palefsky, J. M. Oxidation of methionine 63 and 83 regulates the effect of S100A9 on the migration of neutrophils in vitro. *J. Leukoc. Biol.* **81**, 818–824 (2007).

178. Lim, S. Y. *et al.* S-Nitrosylated S100A8: Novel Anti-Inflammatory Properties. *J. Immunol.* **181**, 5627–5636 (2008).

179. Brenner, S. & Miller, J. *Encyclopedia of Genetics*. *Encyclopedia of Genetics* (2001). doi:10.1006/rwgn.2001.0563

180. Fabian, T. K., Fejérdy, P. & Csermely, P. Saliva in Health and Disease, Chemical Biology of. in *Wiley Encyclopedia of Chemical Biology* (John Wiley & Sons, Inc., 2008). doi:10.1002/9780470048672.wecb643

181. Fabian, T., Fejerdy, P. & Csermely, P. Salivary Genomics, Transcriptomics and Proteomics: The Emerging Concept of the Oral Ecosystem and their Use in the Early Diagnosis of Cancer and other Diseases. *Curr. Genomics* **9**, 11–21 (2008).

182. Brandtzaeg, P. Do Salivary Antibodies Reliably Reflect Both Mucosal and Systemic Immunity? *Ann. N. Y. Acad. Sci.* **1098**, 288–311 (2007).

183. Brandtzaeg, P. & Korsrud, F. R. Significance of different J chain profiles in human tissues: generation of IgA and IgM with binding site for secretory component is related to the J chain expressing capacity of the total local immunocyte population, including IgG and IgD producing cells,. *Clin. Exp. Immunol.* **58**, 709–18 (1984).

184. Kinane, Lappin, F., Koulouri & Buckley. Humoral immune responses in periodontal disease may have mucosal and systemic immune features. *Clin. Exp. Immunol.* **115**, 534–541 (1999).

185. Braathen, R., Hohman, V. S., Brandtzaeg, P. & Johansen, F.-E. Secretory Antibody Formation: Conserved Binding Interactions between J Chain and Polymeric Ig Receptor from Humans and Amphibians. *J. Immunol.* **178**, 1589–1597 (2007).

186. Phalipon, A. & Corthésy, B. Novel functions of the polymeric Ig receptor: well beyond

transport of immunoglobulins. *Trends Immunol.* **24**, 55–8 (2003).

187. Dallas, S. D. & Rolfe, R. D. Binding of Clostridium difficile toxin A to human milk secretory component. *J. Med. Microbiol.* **47**, 879–888 (1998).

188. Zhang, Z. & Marshall, A. G. A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. *J. Am. Soc. Mass Spectrom.* **9**, 225–233 (1998).

189. Inzitari, R. *et al.* Different isoforms and post-translational modifications of human salivary acidic proline-rich proteins. *Proteomics* **5**, 805–815 (2005).

190. Manconi, B. *et al.* Top-down proteomic profiling of human saliva in multiple sclerosis patients. *J. Proteomics* **187**, 212–222 (2018).

191. Ong, S.-E. & Mann, M. Mass spectrometry–based proteomics turns quantitative. *Nat. Chem. Biol.* **1**, 252–262 (2005).

192. Castagnola, M. *et al.* The human salivary proteome: a critical overview of the results obtained by different proteomic platforms. *Expert Rev Proteomics* **9**, 33–46 (2012).

193. Yao, Y., Lamkin, M. S. & Oppenheim, E. Pellicle Precursor Proteins: Acidic Proline-rich Proteins, Statherin, and Histatins, and their Crosslinking Reaction by Oral Transglutaminase. *J. Dent. Res.* **78**, 1696–1703 (1999).

194. Schupbach, P. *et al.* Electron-microscopic demonstration of proline-rich proteins, statherin, and histatins in acquired enamel pelliclesin vitro. *Eur. J. Oral Sci.* **109**, 60–68 (2001).

195. Wang, G. Human Antimicrobial Peptides and Proteins. *Pharmaceuticals* **7**, 545–594 (2014).

196. Xu, T., Levitz, S. M., Diamond, R. D. & Oppenheim, F. G. Anticandidal activity of major human salivary histatins. *Infect. Immun.* **59**, 2549–54 (1991).

197. Du, H. *et al.* Human Salivary Protein Histatin 5 Has Potent Bactericidal Activity against ESKAPE Pathogens. *Front. Cell. Infect. Microbiol.* **7**, (2017).

198. Magister, Š. & Kos, J. Cystatins in Immune System. *J. Cancer* **4**, 45–56 (2013).

199. Zavasnik-Bergant, T. Cystatin protease inhibitors and immune functions. *Front. Biosci.* **13**, 4625–37 (2008).

200. Grzonka, Z. *et al.* Structural studies of cysteine proteases and their inhibitors. *Acta Biochim. Pol.* **48**, 1–20 (2001).

201. Abrahamson, M., Barrett, A. J., Salvesen, G. & Grubb, A. Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids. *J. Biol. Chem.* **261**, 11282–9 (1986).

202. Haves-Zburof, D. *et al.* Cathepsins and their endogenous inhibitors cystatins: expression and modulation in multiple sclerosis. *J. Cell. Mol. Med.* **15**, 2421–2429 (2011).

203. Balbín, M. *et al.* Structural and functional characterization of two allelic variants of human cystatin D sharing a characteristic inhibition spectrum against mammalian cysteine proteinases. *J. Biol. Chem.* **269**, 23156–62 (1994).

204. Soond, S. M., Kozhevnikova, M. V., Townsend, P. A. & Zamyatnin, A. A. Cysteine Cathepsin Protease Inhibition: An update on its Diagnostic, Prognostic and Therapeutic Potential in Cancer. *Pharmaceuticals* **12**, 87 (2019).

205. Baron, A. C., Gansky, S. A., Ryder, M. I. & Featherstone, J. D. B. Cysteine protease inhibitory activity and levels of salivary cystatins in whole saliva of periodontally diseased patients. *J. Periodontal Res.* **34**, 437–444 (1999).

206. Oh, S.-S. *et al.* Extracellular cystatin SN and cathepsin B prevent cellular senescence by inhibiting abnormal glycogen accumulation. *Cell Death Dis.* **8**, e2729–e2729 (2017).

207. Caughey, G. H. Mast cell proteases as pharmacological targets. *Eur. J. Pharmacol.* **778**, 44–55 (2016).

208. Walsh, L. J. Mast Cells and Oral Inflammation. *Crit. Rev. Oral Biol. Med.* **14**, 188–198 (2003).

209. Farahani, S. S., Navabazam, A. & Sadr Ashkevari, F. Comparison of mast cells count in oral reactive lesions. *Pathol. - Res. Pract.* **206**, 151–155 (2010).

210. Butterfield, J. H. Systemic Mastocytosis: Clinical Manifestations and Differential Diagnosis. *Immunology and Allergy Clinics of North America* **26**, 487–513 (2006).

211. Rama, T. A., Côrte-Real, I., Gomes, P. S., Escribano, L. & Fernandes, M. H. Mastocytosis: oral implications of a rare disease. *J. Oral Pathol. Med.* **40**, 441–450 (2010).

212. Edgeworth, J., Gorman, M., Bennett, R., Freemont, P. & Hogg, N. Identification of p8,14 as a highly abundant heterodimeric calcium binding protein complex of myeloid cells. *J. Biol. Chem.* **266**, 7706–13 (1991).

213. Carlsson, H. *et al.* Psoriasin (S100A7) and calgranulin-B (S100A9) induction is dependent on reactive oxygen species and is downregulated by Bcl-2 and antioxidants. *Cancer Biol. Ther.* **4**, 998–1005 (2005).

214. Ryckman, C., Vandal, K., Rouleau, P., Talbot, M. & Tessier, P. A. Proinflammatory Activities of S100: Proteins S100A8, S100A9, and S100A8/A9 Induce Neutrophil Chemotaxis and Adhesion. *J. Immunol.* **170**, 3233–3242 (2003).

215. Valore, E. V & Ganz, T. Posttranslational processing of defensins in immature human myeloid cells. *Blood* **79**, 1538–44 (1992).

216. Suarez-Carmona, M., Hubert, P., Delvenne, P. & Herfs, M. Defensins: “Simple” antimicrobial peptides or broad-spectrum molecules? *Cytokine Growth Factor Rev.* **26**, 361–370 (2015).

217. Chaly, Y. V *et al.* Neutrophil alpha-defensin human neutrophil peptide modulates cytokine production in human monocytes and adhesion molecule expression in endothelial cells. *Eur. Cytokine Netw.* **11**, 257–66 (2000).

218. Hannappel, E. Thymosin β4 and its posttranslational modifications. *Ann. N. Y. Acad. Sci.* **1194**, 27–35 (2010).

219. Wyczółkowska, J., Walczak-Drzewiecka, A., Wagner, W. & Dastych, J. Thymosin β4 and thymosin β4-derived peptides induce mast cell exocytosis. *Peptides* **28**, 752–759 (2007).

220. Nemolato, S. *et al.* Thymosin beta 4 expression in normal skin, colon mucosa and in tumor infiltrating mast cells. *Eur. J. Histochem.* **54**, 3 (2010).

221. He, S.-H., Chen, P. & Chen, H.-Q. Modulation of enzymatic activity of human mast cell tryptase and chymase by protease inhibitors. *Acta Pharmacol. Sin.* **24**, 923–9 (2003).

222. Ruëff, F., Friedl, T., Arnold, A., Kramer, M. & Przybilla, B. Release of Mast Cell Tryptase into Saliva: A Tool to Diagnose Food Allergy by a Mucosal Challenge Test? *Int. Arch. Allergy Immunol.* **155**, 282–288 (2011).

223. Vizcaíno, J. A. *et al.* 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456 (2016).

224. Beers, C. *et al.* Cathepsin S Controls MHC Class II-Mediated Antigen Presentation by Epithelial Cells In Vivo. *J. Immunol.* **174**, 1205–1212 (2005).

225. Shugars, D. C. & Wahl, S. M. The Role of the Oral Environment in HIV-1 Transmission. *J. Am. Dent. Assoc.* **129**, 851–858 (1998).

226. Carrero, J. C. *et al.* The role of the secretory immune response in the infection by Entamoeba histolytica. *Parasite Immunol.* **29**, 331–338 (2007).

227. Richardson, A. & Butterfield, J. H. Concurrent mastocytosis and monoclonal gammopathy of undetermined significance, a case series. *J. Allergy Clin. Immunol.* **143**, AB180 (2019).

228. Varkonyi, J. *et al.* Coexistent systemic mastocytosis and essential thrombocythemia complicated with monoclonal gammopathy and hypocomplementaemia. *Open Med.* **7**, (2012).

229. Ligtenberg, A. J. M., Veerman, E. C. I., Nieuw Amerongen, A. V. & Mollenhauer, J. Salivary agglutinin/glycoprotein-340/DMBT1: a single molecule with variable composition and with different functions in infection, inflammation and cancer. *Biol. Chem.* **388**, (2007).

230. Hartshorn, K. L. *et al.* Salivary agglutinin and lung scavenger receptor cysteine-rich glycoprotein 340 have broad anti-influenza activities and interactions with surfactant protein D that vary according to donor source and sialylation. *Biochem. J.* **393**, 545–553 (2006).

231. Edwards, A. M. *et al.* Scavenger receptor gp340 aggregates group A streptococci by binding pili. *Mol. Microbiol.* **68**, 1378–1394 (2008).

232. Malamud, D. *et al.* Antiviral Activities in Human Saliva. *Adv. Dent. Res.* **23**, 34–37 (2011).

233. Walsh, L. J., Savage, N. W., Ishii, T. & Seymour, G. J. Immunopathogenesis of oral lichen planus. *J. Oral Pathol. Med.* **19**, 389–396 (1990).

234. White, M. R. *et al.* Multiple components contribute to ability of saliva to inhibit influenza viruses. *Oral Microbiol. Immunol.* **24**, 18–24 (2009).

235. Gilfillan, A. M. & Beaven, M. A. Regulation of mast cell responses in health and disease. *Crit. Rev. Immunol.* **31**, 475–529 (2011).

236. Pundir, P. & Kulka, M. The role of G protein-coupled receptors in mast cell activation by antimicrobial peptides: is there a connection? *Immunol. Cell Biol.* **88**, 632–640 (2010).

237. Goulding, N. J., Godolphin, J. L., Sampson, M. B., Maddison, P. J. & Flower, R. J. Hydrocortisone induces lipocortin 1 production by peripheral blood mononuclear cells in vivo in man. *Biochem. Soc. Trans.* **18**, 306–307 (1990).

238. Babbin, B. A. *et al.* Annexin A1 Regulates Intestinal Mucosal Injury, Inflammation, and Repair. *J. Immunol.* **181**, 5035–5044 (2008).

239. Martin, G. R., Perretti, M., Flower, R. J. & Wallace, J. L. Annexin-1 modulates repair of gastric mucosal injury. *Am. J. Physiol. Liver Physiol.* **294**, G764–G769 (2008).

240. Leoni, G. *et al.* Annexin A1, formyl peptide receptor, and NOX1 orchestrate epithelial repair. *J. Clin. Invest.* **123**, 443–454 (2013).

241. Courtneidge, S., Ralston, R., Alitalo, K. & Bishop, J. M. Subcellular location of an abundant substrate (p36) for tyrosine-specific protein kinases. *Mol. Cell. Biol.* **3**, 340–50 (1983).

242. Arrigo, A. P., Darlix, J. L. & Spahr, P. F. A cellular protein phosphorylated by the avian sarcoma virus transforming gene product is associated with ribonucleoprotein particles. *EMBO J.* **2**, 309–15 (1983).

243. Vishwanatha, J. K., Jindal, H. K. & Davis, R. G. The role of primer recognition proteins in DNA replication: association with nuclear matrix in HeLa cells. *J. Cell Sci.* **101 (Pt 1**, 25–34 (1992).

244. Waisman, D. M. Annexin II tetramer: structure and function. *Mol. Cell. Biochem.* **149–150**, 301–322 (1995).

245. Gerke, V., Creutz, C. E. & Moss, S. E. Annexins: linking Ca2+ signalling to membrane dynamics. *Nat. Rev. Mol. Cell Biol.* **6**, 449–461 (2005).

246. Madureira, P. & Waisman, D. Annexin A2: The Importance of Being Redox Sensitive. *Int. J. Mol. Sci.* **14**, 3568–3594 (2013).

247. Tashian, R. E. The carbonic anhydrases: Widening perspectives on their evolution, expression and function. *BioEssays* **10**, 186–192 (1989).

248. Tashian, R. E. Genetics of the Mammalian Carbonic Anhydrases. in 321–356 (1992). doi:10.1016/S0065-2660(08)60323-5

249. Parkilla, S., Parkilla, A. K. & Rajaniemi, H. Circadian periodicity in salivary carbonic anhydrase VI concentration. *Acta Physiol. Scand.* **154**, 205–211 (1995).

250. Kivelä, J., Parkkila, S., Parkkila, A.-K. & Rajaniemi, H. A Low Concentration of Carbonic Anhydrase Isoenzyme VI in Whole Saliva Is Associated with Caries Prevalence. *Caries Res.* **33**, 178–184 (1999).

251. Leinonen, J., Kivelä, J., Parkkila, S., Parkkila, A.-K. & Rajaniemi, H. Salivary Carbonic Anhydrase Isoenzyme VI Is Located in the Human Enamel Pellicle. *Caries Res.* **33**, 185–190 (1999).

252. Kamsteeg, M. *et al.* Increased Expression of Carbonic Anhydrase II (CA II) in Lesional Skin of Atopic Dermatitis: Regulation by Th2 Cytokines. *J. Invest. Dermatol.* **127**, 1786–1789 (2007).

253. Wen, T. *et al.* Carbonic Anhydrase IV Is Expressed on IL-5–Activated Murine Eosinophils. *J. Immunol.* **192**, 5481–5489 (2014).

254. Henry, E. K. *et al.* Carbonic anhydrase enzymes regulate mast cell–mediated inflammation. *J. Exp. Med.* **213**, 1663–1673 (2016).

255. Winum, J.-Y. Carbonic anhydrase enzymes for regulating mast cell hematopoiesis and type-2 inflammation: a patent evaluation (WO2017/058370). *Expert Opin. Ther. Pat.* **28**, 741–743 (2018).