Università degli studi di Cagliari

Ph.D. DEGREE

Life, Environmental and Drugs Sciences

Cycle XXXIV

TITLE OF THE Ph.D. THESIS

Proteomic investigation and characterization of cystatin B interactome in saliva of patients with Alzheimer's disease

Scientific Disciplinary Sector(s)

BIO/10

Ph.D. Student: Doctor Cristina Contini

Supervisor: Professor Tiziana Cabras

Final exam. Academic Year 2020/2021
Thesis defense: April 2022 Session

Contents

ABREVIATIONS' LIST

2-DE (two-dimensional gel electrophoresis); **Aβ** (amyloid beta); **Abs** (antibodies); **ACN** (acetonitrile); **AD** (Alzheimer's disease); **aHC** (adult healthy controls); **aPRPs** (acidic proline-rich proteins); **APP** (amyloid precursor protein); **Arps** (Actin-related proteins); **BACE-1** (beta-site amyloid precursor protein cleaving enzyme 1); **BBB** (blood brain barrier); **BCA** (bicinchoninic acid); **BP** (biological process); **CC** (cellular compartment); **ChEI** (cholinesterase inhibitors); **CID** (collision-induced dissociation); **CICs** (circulating immune-complexomes); **CSF** (cerebrospinal fluid); **CNS** (central nervous system); **Co-IP** (Co-Immunoprecipitation); **CRAPome** (Contaminant Repository for Affinity Purification); **CysB** (Cystatin B); **DAVID** (Database for Annotation, Visualization and Integrated Discovery); **ELISA** (enzyme-linked immunosorbent assay); **EO-FAD** (Early Onset-Familial Alzheimer Disease); **ESI** (electrospray ionization); **FA** (formic acid); **FDR** (false discovery rate); **GO** (Gene Ontology); **HC** (healthy controls); **aHC** (adult healthy controls); **eHC** (elderly healthy controls); **HCA** (Hierarchical Cluster Analysis); **HCD** (higher-energy collisional dissociation); **HPLC-(LR or HR)-ESI-IT-MS** (high-pressure liquid chromatography coupled to low resolution or high resolution electrospray ionization ion trap mass spectrometry); **Hst** (histatin); **ICs** (immune-complexomes); **IP** (immunoprecipitation); **kDa** (kilo Daltons); **LFQ** (label free quantification); **LOAD** (Late Onset Alzheimer Disease); **(L)SSG** (long glutathionylated); **LR-MS** (low resolution-mass spectrometry); **Mav** (average mass values); **MCI** (mild cognitive impairment); **MDG** (mean decrease of the Gini index); **MDS** (Multidimensional Scaling); **MF** (molecular function); **MS** (mass spectrometry); **MS/MS** (tandem mass spectrometry); **MUC** (mucin); **MW** (molecular weight); **NEG** (negative control); **NFTs** (neurofibrillary tangles); **NMDA** (N-methyl-D-aspartate); **NR-** (non-reducing condition); **PD** (Proteome Discoverer software); **PTMs** (post-translational modifications); **R-** (reducing confition); **RAGE** (receptor for advanced glycosylation end products); **RF** (Random Forest); **RNS** (reactive nitric oxide species); **ROS** (reactive oxygen species); **RP-HPLC** (reverse phase high performance liquid chromatography); **SAD** (sporadic Alzheimer's disease); **SD** (standard deviation); **SDS-PAGE** (Sodium dodecyl sulfate polyacrylamide gel electrophoresis); **SLPI** (Secretory Leukocyte Protease Inhibitor); **SNO** (S-Nitrosylated); **SS** (disulfide bridge); **Tβ4** (Thymosin beta 4); **TFA** (trifluoroacetic acid); **TIC** (Total Ion Current); **TPC** (total protein concentration); **TPI** (triosephosphate isomerase); **XIC** (eXtracted Ion Current).

ABSTRACT

Alzheimer's disease is the most prevalent neurodegenerative disease in the elderly, characterized by accumulation in the brain of misfolded proteins, inflammation, and oxidative damage leading to neuronal cell death. By considering its diagnosis is still heavily based on the analysis of cerebrospinal fluid, the need of finding new non-invasive biomarkers has been pointed. In the context of biomarker discovery proteomics-based, mass spectrometry techniques represent one of the most used and powerful applications to study complex protein mixtures and to define protein profiles in different tissues and body fluids.

According to literature, the possible use of saliva as a diagnostic tool has been explored in several oral and systemic diseases. Indeed, saliva comes across as one of the less invasive collectable biofluids suitable for the research of diseases' biomarkers, especially due to the possibility to detect and measure many proteins and peptides which are expressed in tissues, cell-types and biofluids different from the oral cavity, including the brain.

Thus, the aim of the present thesis has been to investigate the salivary proteome profile of Alzheimer's disease patients in comparison with healthy controls, to evidence possible qualitative/quantitative variations associated with the disease (Part I). Having a cohort of healthy subjects age and sex matched was extremely important because salivary proteome composition is influenced by the process of aging as it has been proven from childhood to adulthood. However, little is known about the changes in saliva protein composition in advanced age. For this reason, a statistical comparison between a cohort of young adult controls and a cohort of old adult controls in relation with Alzheimer's disease patients has been performed (Part II). To reach these goals, a top-down proteomic approach through HPLC-ESI-IT-MS and HPLC-ESI-high-resolution-MS/MS on the acid-soluble fraction of saliva has been chosen. Among the results obtained, the potential role of cystatins in neurodegeneration diseases appeared to agree with previous findings and we also observed the evidence of

aggregates of cystatin B in saliva. Thus, we investigated and characterized the interactome of cystatin B in whole saliva for the first time through a Co-Immunoprecipitation assay followed by bottom-up proteomic approach using a nanoHPLC-ESI-high-resolution-MS/MS system (Part III).

INTRODUCTION

Alzheimer's disease

Alzheimer's disease (AD) is a neurodegenerative pathology and it is the leading cause of progressive dementia in the elderly population[1]. AD is characterized by the accumulation of amyloid beta (Aβ) plaques and by the development in the brain of neurofibrillary tangles (NFTs) composed of tau protein[2]. Clinically, AD is a slow but unstoppable neurodegenerative process that affects cognitive functions. The progression of the disease can manifest itself with different symptoms depending on the brain region involved: memory loss, dysfunction of language, of the ability to judge, loss of motor coordination, changes in behavioral and social aspects[3]. Only a small portion of AD cases has genetic causes while the majority (95%) are sporadic and occur in people over 65 years[4]. The sporadic form of AD is more complex since it likely results from a combination of genetic and environmental factors. Details on the characteristics of the disease are discussed in the present introduction.

Epidemiology

Dementia, especially AD, is present in all the populations around the world. According to data collected in the last two decades, the prevalence of dementia doubles every 5 years in individuals between the ages of 65 and 85 and continues increasing after the age of 90. It is estimated that only in the United States 4.7 million individuals aged 65 years or older are affected by AD dementia. This includes 0.7 million between 65 and 74 years, 2.3 million aged 75–84 years, and 1.8 million aged 85 years and older. This prevalence may rise to around 14 million in the United States and 130 million worldwide by 2050[5]. More women than men died from dementia in 2016 with age-standardized death rates in women higher than in men, probably due to both the longer lifespan of women and a higher prevalence of the disease itself[6]. The average incidence of dementia is similar in China, in the United States and in Europe and

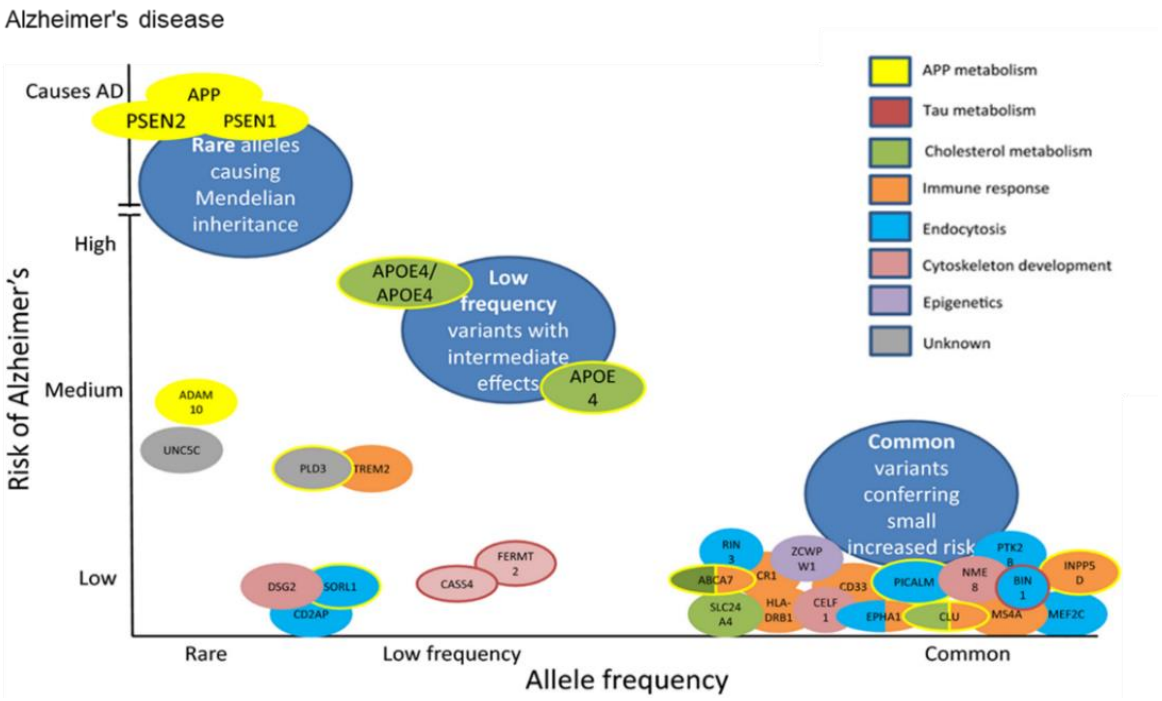
it is represented by approximately 10/1000 person-years[7]. Globally, in 2016, dementia was the fifth-largest cause of death after ischemic heart disease, chronic obstructive pulmonary disease, intracerebral hemorrhage, and ischemic stroke. More in detail, observing only the number of deaths in individuals aged more than 70 years, dementia was the second largest cause of death after ischemic heart disease[6]. It has been estimated that the number of deaths due to dementia increased by 148% between 1990 and 2016. Although traditionally mortality statistics are not a good measure of deaths attributed to dementia or AD, it is important to underline that while individuals may not die of AD *per se*, advanced dementia increases vulnerability for other disorders (*e.g.*, common infections), which ultimately leads to death, or decisions to not treat aggressively other conditions in severely demented patients[8].

Etiology

AD was first described in 1906 by a clinical psychiatrist and neuroanatomist, Alois Alzheimer, during the 37th Meeting of South-West German Psychiatrists in Tübingen. The clinical observation was about distinctive plaques and NFTs in the brain histology of a 50 years woman[9]. Unfortunately, after more than one century, the etiology of AD has not been fully elucidated yet. The disease can manifest as Early Onset-Familial Alzheimer Disease (EO-FAD) or Late Onset Alzheimer Disease (LOAD) also known as sporadic Alzheimer's disease (SAD). EO-FAD typically occurs between 30 and 50 years of age with a Mendelian autosomal dominant inheritance. This form of AD is strongly associated with the mutations of three different genes: APP (chromosome 21; coding for the amyloid precursor), PSEN1 (chromosome 14; coding for prosenilin 1) and PSEN2 (chromosome 1; coding for prosenilin 2)[10]. One of the leading genetic risk factors of EO-FAD is Down Syndrome due to the presence of extra chromosome 21 where APP gene is located, but also disfunction in autophagy, lysosomal activity and mitochondria[11]. On the other hand, LOAD or SAD is the most common

form of AD and, being most likely the result of a combination of genetic mutation associated with other risk factors such as environmental and behavioral ones, it represents a more complex condition[12]. Some risk elements have been identified in aging, genetic and environmental factors, vascular diseases, oxidative stress, neuroinflammation of microglia, smoking, depression, stress[4,12]. The first gene to be identified and whose mutations represent a considerable risk for sporadic AD is APOE[13]. However, genetic studies have reported many other genes involved in different pathways that may contribute to the development of AD with different risk-levels (Figure 1). The most common consequence of exposure to environmental factors is the overproduction of reactive oxygen species (ROS). Oxidative stress, observed in AD, is mainly due to Aβ misfolding[14], but altered metal homeostasis, which has been demonstrated in the brain and plasma/serum of AD patients[15], may contribute to oxidative stress too. Additionally, it has been recently hypothesized that prolonged exposure to several heavy metals (aluminum, arsenic, cadmium, lead, mercury), particulate air, some pesticides, metal-containing nanoparticles and the alteration of oral and gut microbiota may increase the risk of developing AD[16]. Indeed, the oral and gut microbiota-induced neuronal inflammation is a gradually emerging idea promoted by the discovery that brain infections, involving bacteria or viruses as external risk factors, can trigger Aβ deposition and AD development[17–19].

Fig.1: An overview of genes which have been implicated in AD. The internal color corresponds to their understood function. Where there are two internal colors, the gene has been implicated in more than one pathway. Genes circled in yellow are also thought to influence APP metabolism; genes circled in red are thought to influence tau metabolism[12].



Pathogenesis

AD is characterized by reactive microgliosis, dystrophic neurites and loss of synapses and neurons from the brain. The neuropathological hallmarks and the two main pathogenesis hypotheses of AD are represented by Aβ cascade and the hyperphosphorylation of tau proteins that leads to NFTs[11,20] as shown in Figure 2.

It is known that Aβ peptide (39–43 amino acids in length) is the principal component of senile plaques. The discovery of Aβ peptides in cerebrovascular amyloid deposits dates back to the mid 1980's by Glenner and Wong, while the full Aβ amino acid sequence from amyloid plaque cores has been identified in 1992[21]. Aβ peptides derives from sequential proteolysis of APP by different secretases (Figure 3). The β-secretase is responsible of the cleavage of the

extracellular component of APP, generating the N-terminus of the Aβ peptide. From the C-terminal end of the APP protein, the C-terminal fragment is initially subject to ε-cleavage either at threonine (Thr_{48}) or leucine (Leu_{49}). Then the γ-secretase complex mediates cleavage of the remaining intramembrane fragment at the C-terminal in position 38, 40, 42, and 43 (γ-sites) generating fragments that are released extracellularly. The γ-site cleavage appears to be the primary factor influencing the self-aggregation of Aβ peptide. Aβ42 is the most hydrophobic of these Aβ peptides and it is considered the primary aggregation-prone peptide in vivo[22].

Fig.2: Neuropathological hallmarks that characterize AD. As AD progresses, the brain tissue shrinks, the volume of the ventricle, which contains cerebrospinal fluid, increases markedly. At the molecular level Aβ peptides are produced by the cleavage of APP in the membrane of the neurons (1). In the space between the neurons, Aβ forms oligomers that are thought to disrupt the function of the synapses and act in receptors present in the neuron plasma membrane (2). The fibrils of the Aβ oligomers are added in plaques, which interfere with the function of the neurons (3). Tau hyperphosphorylation causes NFTs within neurons, displacing intracellular organelles and disrupting vesicular transport (4)[11].

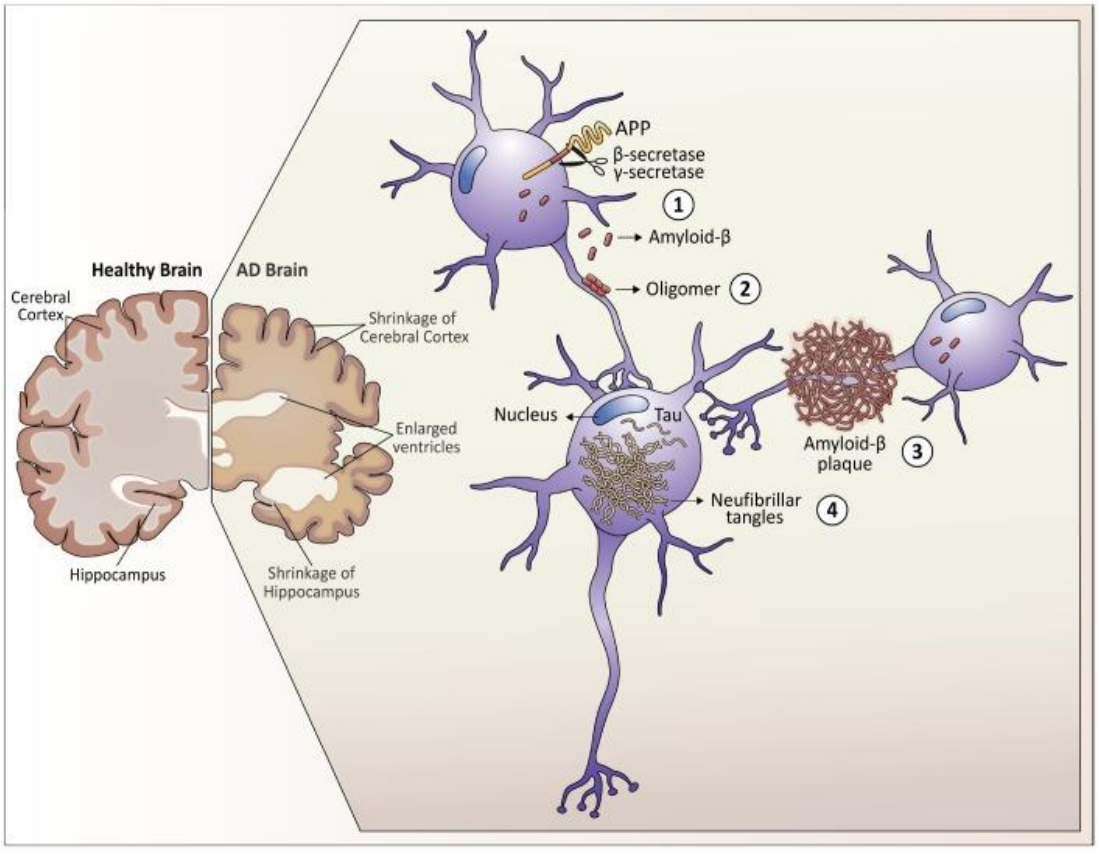
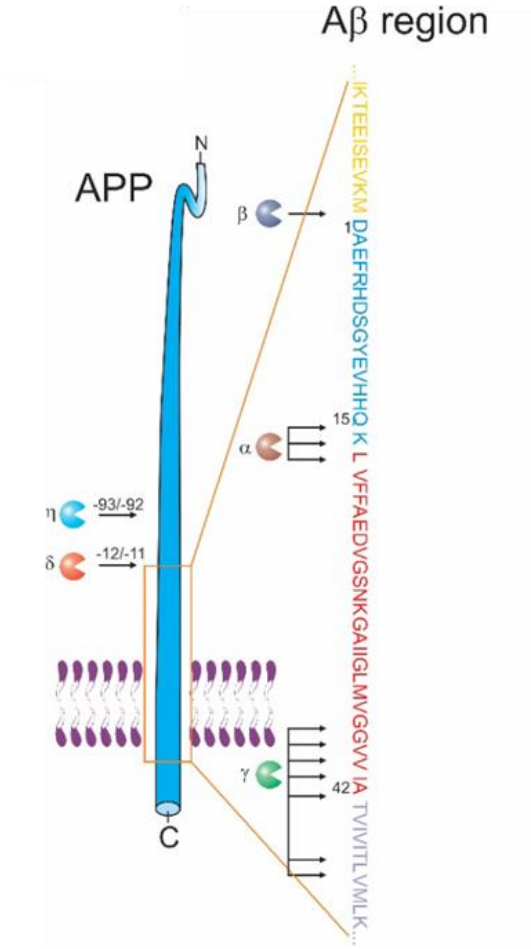


Fig.3: A schematic of APP indicating the most common cleavage sites, and an expanded illustration of the Aβ sequence and flanking amino acids. The amino acid sequence of Aβ is shown in blue and red flanked by residues N- and C-terminal (gold and purple, respectively) of APP[22].



Alternative processing, referred to as non-amyloidogenic pathways, give rise to additional peptides. Initial cleavage by α-secretase followed by γ-secretase cleavage produces the p3 fragment (Aβ17-40/42) and a combination of cleavage by β-secretase and α-secretase produces short N-terminal Aβ peptides (Aβ1-15/16). In addition, combinations of α- or β-secretase with δ- and η-secretase produce different peptides extending N-terminally of the β-site[22].

Most Aβ studies have focused on the neurotoxic role of Aβ peptides (particularly Aβ42) due to their central role in AD. However, it seems peptides may also have multiple physiological functions, playing a role in neurogenesis, calcium homeostasis, modulation of synaptic activity,

and plasticity[23–25]. Their production is normally balanced by clearance due to proteases enzymatic degradation[26], while aggregated forms of Aβ are degraded by matrix metalloproteinase-2 and -9, the cysteine protease cathepsin B, and the aspartyl protease cathepsin D[26,27]. Disruption of the Aβ homeostasis was initially proposed to underlie the symptoms of SAD in the amyloid cascade hypothesis, but over the course of the years, even if Aβ plaques correlate with neuroinflammation and neuritic dystrophy, no correlation was found with the cognitive symptoms of AD, indeed the majority of anti-Aβ therapies failed and it became apparent that Aβ pathology could be present also in non-demented subjects[22].

Tau is a protein capable of binding microtubules through multiple sites of phosphorylation located in its N-terminal region. These sites are phosphorylated by different kinases, but when hyperphosphorylated, mainly because of dysregulation of cyclin-dependent kinase-5 and glycogen synthase kinase 3β, the result is an abnormal increase of cytoskeletal proteins, axoplasmic transport disorders and neuronal degeneration[28].

Aβ and tau proteins are visible in post-mortem AD brains as amyloid plaques and NTFs, respectively. They interact with each other at different levels: aggregated Aβ induces tau hyperphosphorylation and interferes with its oligomerization and aggregation, but at the same time it has been proposed that Aβ toxicity is dependent on tau with the ultimate result of neuronal loss, synaptic damage and cognitive decline of AD patients[29]. Together, they can interact with astrocytes, lead to damage of mitochondria and have a combined effect on microglia. However, treatments that target Aβ or tau alone have not achieved good clinical results and the existence of other proteins, different from Aβ, which can act on tau hyperphosphorylation or the phosphorylation by kinases has been suggested[29].

Finally, the “cellular phase” of AD has been proposed as an extension of the amyloid cascade[30]. According to this hypothesis, AD starts with a “biochemical phase” characterized by the accumulation of cerebral amyloid and tau pathology in a slow, gradual process which can be

tolerated by central nervous system (CNS) cells in the early stages of disease. The biochemical phase serves as a risk factor for development of clinical symptoms, but the disease is only manifest clinically when cellular homeostatic mechanisms fail, leading to impaired clearance of aggregated pathologic protein, increased cellular stress, and a complex breakdown of finely tuned intercellular physiologic functions that ultimately lead to neurodegeneration. Specifically, this "cellular phase" appears characterized by dysfunction of the neurovascular unit, aberrant neuronal network activity, and impaired astrocyte and microglia homeostatic functions with possible gain-of-toxic functions[30].

Clinical features

Clinically, both familial and sporadic AD manifest similar conditions with the majority of cases showing an insidious onset of episodic memory difficulties followed by inexorable progression of cortical cognitive deficits. The most relevant difference between the two form of AD relates to subjects' age[10]. Typical amnestic cases are characterized by early impairment in learning and memory, followed by later impairments in complex attention, executive function, language, visuospatial function, praxis, gnosis and behavior/social comportment[1]. As represented in Figure 2, the brain of AD patients appears atrophic compared to healthy individuals and it has also been demonstrated that the synaptic and neuronal loss in the entorhinal cortex generally correlates well with onset of cognitive impairment[19].

Diagnosis and Biomarkers

Patients first go through an intermediate stage recognized as mild cognitive impairment (MCI) before they are diagnosed with AD. MCI is a heterogeneous syndrome that is considered a prodromal form of AD when accompanied with amnestic clinical evidence[31]. AD diagnosis is then heavily based on detection of Aβ and tau proteins in the CNS, either imaged using positron

emission tomography or measured in the cerebrospinal fluid (CSF); moreover, magnetic resonance imaging can be used to measure function and reveal brain atrophy[32]. Some specific biomarkers in the CSF were shown to have excellent diagnostic accuracy[33] such as Aβ42, which is related to the extracellular senile plaques[34], total tau protein (T-tau), which reflects the level of neuronal damage[35], and phosphorylated tau protein on threonine 181 (P-tau181), which correlates with tangle pathology[36]. Unfortunately, the invasive nature of CSF collection limits its widespread use in routine primary care practice[37]. Furthermore, the growing complexity of AD pathogenesis highlighted the need for additional biomarkers for early preclinical diagnosis. Blood biomarkers are more easily detectable than those from the CSF for population screening; however, standardization and validation is still required[38]. Peripheral biomarkers have also been studied in other tissues, including the eyes and the skin[39] or other body fluids as saliva[40] mainly focusing on Aβ or tau, but without resolutive outcomes.

Therapy

Despite enormous efforts by the pharmaceutical industry, to date, AD is still un uncurable disease mainly treated acting on the symptoms. During the course of AD pathogenesis, cholinergic neurons are lost, causing a general cholinergic deficit. This loss is thought to contribute to early attention and memory dysfunction in AD, for this reason one therapeutic strategy is to use cholinesterase inhibitors (ChEI) as donepezil, rivastigmine and galantamine to reverse this deficiency by increasing synaptic levels of acetylcholine. ChEI are currently approved for use in mild-to-moderate AD[19]. Another drug, memantine, is used to improve the cognition in moderate-to-severe AD patients. Memantine is a uncompetitive NMDA (N-methyl-D-aspartate) receptor modulator that may act to inhibit glutamate- mediated neurotoxicity that develops as neurons die during AD progression[41]. Unfortunately, both strategies have no effect on long-term disease progression[19].

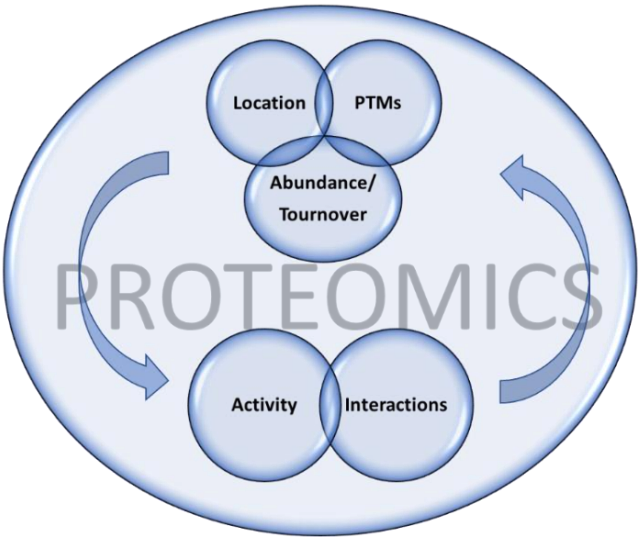
On June 2021, the Food and Drug Administration (FDA) in the United States approved the use of another drug called Aducanumab or, commercially, Aduhelm with the following indication: "*Aduhelm is an amyloid beta-directed antibody indicated to treat Alzheimer's disease. Aduhelm is approved under the accelerated approval pathway, which provides patients with a serious disease earlier access to drugs when there is an expectation of clinical benefit despite some uncertainty about the clinical benefit*" (www.fda.gov). Aduhelm is the first disease-modifying drug to be approved for AD but with clear controversy regarding its efficacy[19,42,43].

Proteomics and Mass Spectrometry

Proteomics studies the complete set of proteins expressed by a given genome at a given time in the life of a cell, tissue or organism, including all isoforms and post-translational modifications (PTMs). Since the expression of a protein, as well as its modifications, vary over time depending on the cell type and in response to different endogenous and exogenous signals (pathological, environmental, pharmacological, toxicological), proteomics studies can help to understand how these changes occur, how proteins are involved in the different pathways that regulate the life of a cell, a tissue or an organism and weather they interact between each other (Figure 4). A proteomic study allows to simultaneously analyze hundreds or even thousands of different proteins and peptides expressed in the biological system under exam (tissue, cell culture, serum, plasma, saliva, urine, CSF and so on). It allows the structural characterization of proteins, their quantification, and the qualitative and quantitative comparison of the same protein profile in distinct populations subjected to different stresses (physiological, environmental, pharmacological, pathological).

Proteomic studies exploit the high yield, sensitivity, accuracy and selectivity of technologically advanced analytical methodologies as Mass Spectrometry (MS) with the help of bioinformatic tools and proteomic and genomic databases.

Fig.4: Schematic of proteomics' areas of interest.



MS is used to determine the molecular weight of proteins through their mass to charge ratio (m/z). Obtaining gas-phase ions is the first step, then the ions are separated on the basis of m/z values in the presence of electric or magnetic fields in a compartment known as mass analyzer. Finally, the amount of each species with a particular m/z value is measured. This technique was originally applied to study the atomic structure and to characterize small molecules, but then became the most powerful in analyzing proteomes[44]. The sequence and characterization of proteins through tandem MS or MS/MS can be obtained using different approaches named top-down, bottom-up or middle-down as illustrated in Figure 5. Despite the approach chosen to analyze a proteome, developing a suitable separation method prior to MS analysis is a critical step depending on the complexity of proteins/peptides mixture in the sample. To do so, there are two main strategies: gel-free and gel-based which can be further coupled with MS with on-line or off-line workflows. The conventional techniques for proteins' separation are chromatographic, such as ion exchange, size exclusion and affinity chromatography. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), and two-dimensional gel electrophoresis (2-DE) can also be used for separation of complex protein samples. These two

so called “gel-based” approaches have been successful in many proteomic investigations using bottom-up MS strategy where proteins are firstly separated based on their molecular mass or both their charge and mass, respectively. On the other hand, the most used gel-free approach for proteomic purposes is reverse phase liquid chromatography (RPLC) which can be easily coupled with MS using a electrospray ionization (ESI) source and thus known as LC-MS or LC-ESI-MS[45,46].

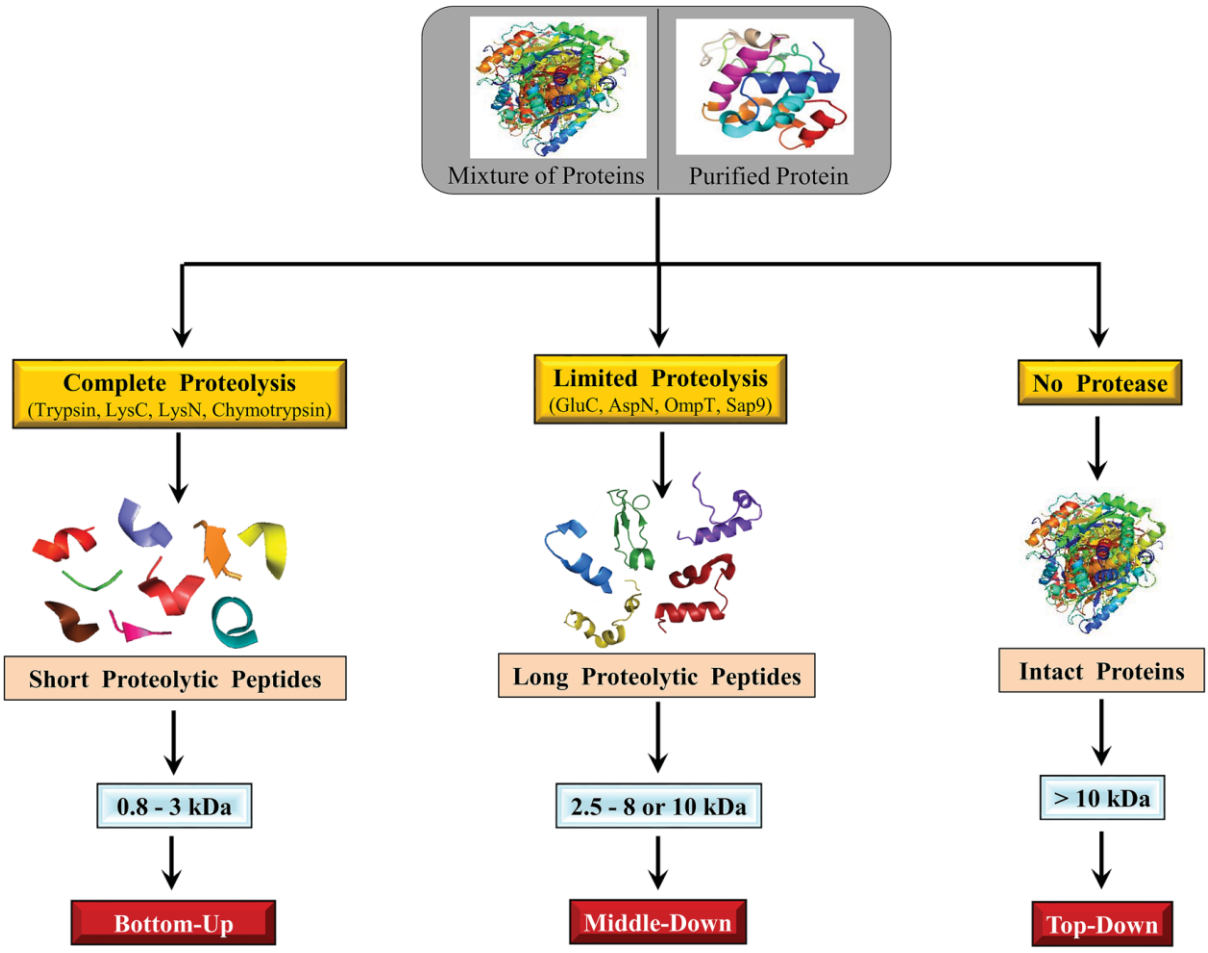
Top-down proteomics allows to study proteins in their naturally occurring form, which is extremely useful to detect PTMs and isoforms according to precise m/z values. In this strategy particular attention is given in avoiding, as much as possible, any sample alteration and proteins/peptides are detected in their intact form[47]. Indeed, the majority of proteins are subjected to PTMs, including endogenous cleavages, which can deeply affect their functions and thus, the ability of detecting proteoforms and isoforms is a key in understanding the biological system under exam[46]. On the other hand, bottom-up approach works as a peptides-based identification. It requires complete enzymatic digestion of proteins prior to tandem MS analysis which is usually performed with trypsin. In this case, the presence of a specific protein in the sample mixture is adjudged from the MS detection and sequencing of its unique tryptic peptides. If the sample is directly digested by enzymatic cleavage skipping the fractionation of proteins through chromatographic or gel-based techniques, the chosen strategy is no longer a bottom-up approach, but it is called shot-gun proteomics.

The digestion process causes losses in the sequence information and prevents the identification of PTMs, but at the same time peptides offer increased separation efficiency, limited number of charge, uniform molecular weight and their mass spectra are easier to analyze due to a simpler isotope distribution, which is beneficial for MS detection[46,48].

Top-down proteomics offers unique advantages over bottom-up or shot-gun proteomics due to the ability of proteoforms identification, yet it shows many analytical challenges mainly

because of MS data complexity, the need of solubilize proteins in MS-compatible buffers and sensitivity issues[46]. Middle-down strategy is an emerging approach that aim to combine top-down and bottom-up through incomplete or partial digestion of proteins. Obtaining longer peptides using different proteases allows to reduce losses in the sequence information and to conserve some PTMs into the proteolytic products. Moreover, the number of proteolytic peptides in samples are relatively lesser compared to bottom-up, resulting in less complex samples and therefore in the enhanced probability of detecting unique peptides[45].

Fig.5: Illustration of the fundamental criteria of three different approaches for analysis of proteins or proteomes[45].



Regardless the approach or technique chosen, one of the most important application of proteomics is the identification and quantification of biomarkers which can reflect a specific condition, like a pathological state, and thus it can be useful for prognostic and/or diagnostic aims. Indeed, the search for biomarkers is a challenge that aims to satisfy the need of detecting and discriminating early stages of the diseases, improve the understanding of diseases' etiology and pathogenicity, discovering new therapies and prevention targets, improve the existing diagnostic assays, particularly in the context of personalized medicine, predict the ability of patients to respond to specific drugs or therapeutic programs as well as following the progression of patients to treatments[49]. In this context, following the principles of precision, specificity, sensitivity and stability in sample collection, analysis and validation, the proteomic investigation of biological samples represent a promising approach for discovering new biomarkers and advancing knowledge of disease pathology, prevention, diagnostics and therapeutics strategies[49].

Saliva and biomarkers

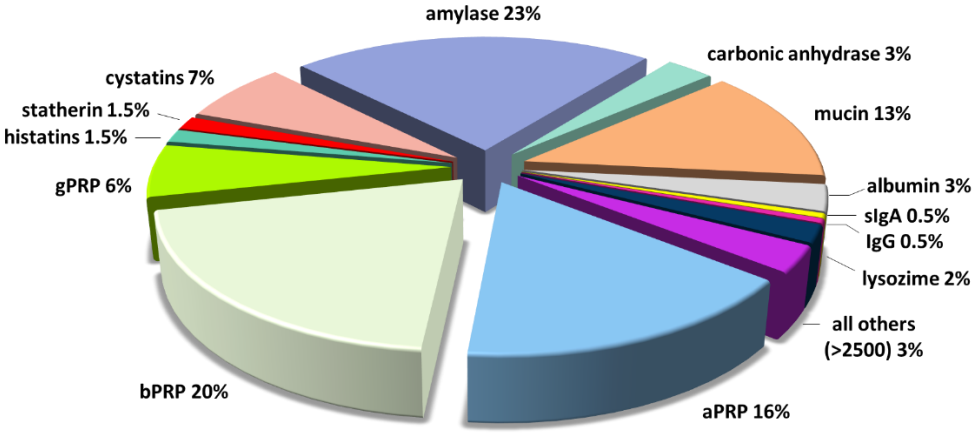
Human saliva is composed by water for approximately 99%. The remaining 1% includes electrolytes, enzymes, hormones, nucleic acids, cytokines and antibodies. Saliva also contains carbohydrates, blood-group substances, lipids, and vitamins, as well as many proteins[50]. Unstimulated saliva flow rate displays circadian variation, with a peak level in the afternoon, originating preponderantly from three pairs of major salivary glands (parotid, sublingual, and submandibular) and from numerous minor salivary glands. The autonomic nervous system regulates the activity of salivary glands, but secretion is also influenced by various centers in the brain and gastrointestinal hormones as well[51]. Among its multiple functions, which are important for the maintenance of oral and general health, saliva lubricates and cleanses the teeth and oral mucosa, maintains neutral pH through its buffering capacity, prevents tooth

demineralization, exerts antimicrobial actions, aids in taste and bolus formation, initiates enzymatic digestion of starch and is imperative for mastication, swallowing and articulation of speech[51]. Saliva participates in the formation of the acquired enamel pellicle and the mucosal pellicle, which have a protective function and determine the initial adhesion and colonization of microorganisms and the composition of the resident oral microbiota[51,52]. The whole saliva proteome is composed of around 3000 proteins and peptides under the secretions of the minor and major salivary glands as well as from mucosal transudations, gingival crevicular fluid, serum and some blood derivatives, desquamated epithelial cells, bacteria, viruses, fungi, and food debris[50,53].

The classes of proline-rich proteins (PRPs; divided in (a) acidic, (b) basic and (g) basic glycosylated), α-amylases, mucins, salivary ("S-type") cystatins, histatins, statherin, represent around 90% in weight of the components detected in saliva and derive from the secretion of the three couples of major glands[53,54]. A representation of the principal classes of salivary proteins and peptides in human whole saliva is shown in Figure 6. Among the proteins/peptides with non-glandular origin, it is possible to measure in saliva defensins, β-thymosins, S100A protein family, cystatin B, cystatin C and others, which are also expressed in other organs and tissues and involved in many different pathways and biological functions[55,56].

In the last decade human saliva has gained more attention for its potential use in biomarkers search and diagnostic field[57,58]. The developments and availability of MS techniques have improved the research in the salivary proteome and saliva has been proposed as an easily collectable source of potential biomarkers for diagnosis and risk assessment for a range of pathological conditions that occur not only orally but also systemically, including neurodegenerative disorders[40,59,60].

Fig.6: Approximate percentages in weight of the principal classes of salivary proteins and peptides in human whole saliva (readapted from Boroumand *et al.*[56]).



Several research groups have studied the salivary protein profiles characterizing neurodegenerative diseases or tried to quantify specific proteins disease-related in saliva to evidence possible significant differences between patients and controls. For instance, promising results have been obtained on salivary α-synuclein levels from Parkinson's disease patients using specific ELISA (enzyme-linked immunosorbent assay) kits[61]; Bloom and colleagues succeeded in detecting higher levels of Huntingtin protein in saliva from Huntington's disease patients by Western Blot and ELISA assay and found that salivary interleukin-6 could work as potential non-invasive biomarker for Huntington's disease symptoms[62,63]; Obayashi and colleagues found a correlation between salivary levels of the neuroendocrine secretory protein chromogranin A measured through a specific immunoassay and the disease severity of patients affected by amyotrophic lateral sclerosis[64]; in the investigation of autism spectrum disorder twelve salivary proteins involved in immune reactions were found differentially expressed between patients and controls subjects through LC-MS/MS[65], as well as the evidence of altered phosphorylation state of some salivary proteins through top-down proteomic by RP-HPLC-ESI-MS[66]; top-down proteomic approach has been also applied to investigate saliva from patients afflicted by multiple sclerosis[67] where salivary

proteins levels appeared to reflect the inflammatory condition and altered immune response typical of the pathology. All these examples reflect the urgent need of finding non-invasive biomarkers, which could lead to the early diagnosis of neurodegenerative disorders. In fact, the diagnosis of diseases affecting the CNS remains a challenge. Usually, the tests performed for the diagnosis of neurological conditions are either blood tests or lumbar puncture to collect CSF. Their invasive nature, especially for the lumbar puncture, results in discomfort, pain and disagreeable side effects for patients, which necessitates the search for accurate, more advanced and less invasive testing methods[60]. In this perspective, the use of saliva shows clear advantages with respect to other body fluids as CSF or blood: its collection does not require the presence of medical staff, it is fast, easy, inexpensive and non-invasive, which makes saliva collectable in all kinds of subjects, including infants or non-collaborative ones[56,68]. Furthermore, most blood biomarkers can also be found in saliva since proteins from the blood can pass into the saliva via passive diffusion, active transport, or microfiltration[59,69] and it was reported that proteins from CNS are also excreted into the saliva[69]. All these characteristics make saliva a body fluid with high potential in the field of biomarker discovery proteomics-based and further suitable for longitudinal studies, especially if compared to CSF.

PART I

Top-Down proteomics of human saliva highlights anti-inflammatory, antioxidant, and antimicrobial defence responses in Alzheimer disease

INTRODUCTION AND AIM OF THE STUDY

AD is an incurable disease which is estimated will affect 130 million people worldwide by 2050 among the elderly population[5]. The diagnosis of diseases affecting the CNS, including forms of dementia like AD, represents still a challenge. Indeed, usually the tests performed for the diagnosis of neurological conditions involves CSF analyses, which require a lumbar puncture that results in pain, discomfort and side effects for patients, highlighting the necessity to look for alternative and less invasive testing methods which can also lead to disease identification at its early stages[60].

In this context, saliva is a biological fluid rich in proteins and peptides that gained attention in the last decade for its promising role in the search of potential disease-related biomarkers, including neurodegenerative disorders[40]. In fact, it was suggested that the alteration of the autonomic nervous system, which occurs in AD, could affect the activity of the major salivary glands and, thus, saliva production and composition[60]. According to literature, Gleerup et al. found that Aβ42 and tau protein could be candidates as salivary biomarkers of AD[40]. Furthermore, several proteins detectable in human saliva by proteomic approaches[55] have been investigated for their implication in AD in other body districts. For instance, S100A proteins which are strongly involved in inflammation process were found increased in the sera and brain lysates from AD patients[70,71], as well as α-defensins, which are antimicrobial peptides, found increased both in the sera and CSF of AD patients[72]. Another peptide involved in antimicrobial and anti-inflammatory activities is Thymosin β4 (Tβ4). Tβ4 was found highly expressed in reactive microglia of AD patients, where it was proposed to suppress the pro-inflammatory signaling to play a protective role on the CNS[73]. Additionally, the potential role of cystatins in AD pathobiology has been suggested because of the discovery of their co-localization with Aβ[74]. Cystatin C was the first one found co-localized with accumulated Aβ in the brain, and

biochemical studies have shown that the binding of cystatin C to soluble Aβ prevents Aβ oligomerization, fibril formation, and amyloid deposition[74]. On the other hand, cystatin B has been found overexpressed both in conditions of oxidative stress and neurodegeneration[75] and its aggregates have also been reported in association with senile plaques in AD and Parkinson disease as well as in patients suffering from senile dementia, suggesting that cystatin B could be considered as amyloid constituent[76]. Furthermore, *in vitro* studies have demonstrated that Aβ fibril growth is totally inhibited by tetramers of wild-type cystatin B, while there is no inhibition by monomers, dimers or higher oligomers, suggesting this protein may play a role in preventing Aβ fibril formation[77].

These examples concern proteins and peptides which are involved in many biological functions also at the level of the nervous system, so it is not surprising that they may be implicated in molecular processes associated with AD pathogenesis. Considering the difficulties in the early diagnosis of AD and the presence in saliva of different proteins/peptides already found associated with the pathology, we decided to investigate the salivary protein profile of AD patients with respect to healthy controls (HCs). Specifically, the research group I collaborated with applies a top-down proteomic approach for the disease biomarker discovery, so in the present study, qualitative/quantitative variations of salivary proteoforms detectable in the acid-soluble fraction of saliva have been analyzed by RP-HPLC-ESI-IT-MS and MS/MS analyses. The experiments and research activity of the present section was performed at the University of Cagliari (Cagliari, Italy) under the supervision of Professor Tiziana Cabras. The results of the present study were published in the journal Frontiers in Neurosciences as follows: "*Top-Down Proteomics of Human Saliva Highlights Anti-inflammatory, Antioxidant, and Antimicrobial Defense Responses in Alzheimer Disease*, Contini Cristina, Alessandra Olianas, Simone Serrao, Carla Deriu, Federica Iavarone, Mozhgan Boroumand, Alessandra Bizzarro,

Alessandra Lauria, Gavino Faa, Massimo Castagnola, Irene Messana, Barbara Manconi, Carlo Masullo, Tiziana Cabras. Front. Neurosci. 2021 doi: 10.3389/fnins.2021.668852".

MATERIALS AND METHODS

Reagents and Instruments

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, United States). HPLC-low-resolution ESI-MS (HPLC-(LR)-ESI-MS), analyses were performed with a Surveyor HPLC system connected to an LCQ Advantage mass spectrometer (Thermo Fisher Scientific, San Jose, CA, United States). The chromatographic column was a Vydac C8 reverse-phase column (Grace, Hesperia, CA, United States) (2.1 x 150 mm, particle diameter 5 mm). HPLC/high-resolution ESI-MS and MS/MS (HPLC-(HR)-ESI-MS and MS/MS) experiments were carried out using an UltiMate 3000 micro HPLC apparatus (Dionex, Sunnyvale, CA, United States) equipped with an FLM-3000 flow manager module and coupled to an LTQ-Orbitrap Elite or LTQ Orbitrap XL apparatus (Thermo Fisher). The column was a Zorbax 300SB-C8 (1.0 x 150 mm; 3.5 mm particle diameter). All the chemicals and reagents for immunodetection were purchased from Bio-Rad (Hercules, California, United States); the primary mouse monoclonal antibodies (Abs) for α-defensins, cystatin A, cystatin B, S100A8, S100A9, cystatin SN, and Tβ4 were purchased from Santa Cruz Biotechnology (Dallas, TX, United States). The secondary rabbit anti-mouse Ab was from Invitrogen (Waltham, MA, United States). Standard Tβ4 was provided by Bachem (Bubendorf, Switzerland).

Study subjects and clinical data

Thirty-five subjects affected by AD (23 females and 12 males; mean age and standard deviation (SD): 80 ± 6) were recruited by the Neurology Department of the "Fondazione Policlinico Universitario A. Gemelli," Catholic University of Rome. The clinical diagnosis of AD has been

carried out according to standardized criteria[1]. The control group included thirty-four healthy volunteers (18 females and 16 males; mean age and SD: 78 ± 5). The informed consent process agreed with the latest stipulations established by the Declaration of Helsinki. The study was approved by the formal ethical committee of the Catholic University of Rome. AD patients and controls were carefully selected as non-smokers. They were not affected by any major oral disease (periodontitis, caries, or dry mouth). Thirteen patients were classified in the moderate stage, and the remaining patients were in the mild stage. Demographic and clinical features of the included patients are reported in Table 1.1, as well as the pharmacological therapy. Nineteen patients underwent ChEI therapy (donepezil 5 mg or rivastigmine 9.5 mg daily, group G1). Eight patients were treated with ChEI in association with memantine, an antagonist of the NMDA receptor (group G2). Eight patients were treated with memantine 20 mg daily (group G3).

Sample collection and treatment

All the samples of non-stimulated whole saliva were collected between 9:00 and 12:00 a.m. Donors, in fasting conditions, 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. Variable volumes of whole saliva were collected depending on donor capacity and disposal, in any case, a minimum of 0.3 mL was guaranteed every time. For each donor, 0.1 mL was designed to the analysis of the acidic-soluble fraction of the salivary proteins (Part I and II of the thesis), while 0.2 mL was utilized for interactome study (Part III of the thesis).

For the acidic protein fractionation, salivary samples were immediately diluted in a 1:1 v/v ratio with a 0.2% solution of 2,2,2-trifluoroacetic acid (TFA) containing 50μM of leu-enkephalin as internal standard. Then, each sample was centrifuged at 20000 g for 15 minutes

at 4°C. Finally, the supernatant was separated from the precipitate and immediately analyzed by RP-HPLC-ESI-MS or stored at -80°C until the analysis for up to two weeks. Aliquots of 50 µL of each sample were used for the determination of the total protein concentration (TPC) by a bicinchoninic acid (BCA) assay kit (Sigma-Aldrich/Merck, Darmstadt, Germany) in duplicate, following the provided instructions. The TPC determined for each salivary samples is indicated together with demographic data in Table 1.1.

RP-HPLC-low resolution ESI-MS analysis

Thirty microliters of acidic extracts were injected in HPLC-(LR)-ESI-MS, applying procedures and conditions optimized in our previous studies[78,79]. The total ion current (TIC) chromatographic profiles were analyzed to selectively search and quantify the peptides/proteins reported in Table S1.1 (Supplementary Section), which shows UniProtKB codes, elution times, and experimental and theoretical average mass values (Mav) of the proteins/peptides included in the study. Table S1.1 also reports the multiply charged ions used for the eXtracted Ion Current (XIC) search, which were selected by excluding values common to other closely eluting proteins. A window of ± 0.5 Da was used to extract ion current peaks. Experimental Mav were obtained by deconvolution of averaged ESI-MS spectra automatically performed by using MagTran 1.0 software[80]. Mav and elution times of proteins/peptides were compared with those determined on salivary samples, under the same experimental conditions, in our previous studies[55]. Experimental Mav were also compared with the theoretical ones available at the UniProtKB human data bank (https://www.expasy.org/). To confirm the identification of proteins and peptides a (HR)-MSMS analysis was performed.

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

Twenty-two salivary samples (4 and 18 from the AD and HC groups, respectively) were subjected to HPLC-(HR)-ESI-MS/MS (LTQ-Orbitrap Elite or LTQ-Orbitrap XL) to perform a top-down characterization confirming the identity of the peptides and proteins investigated in this study and reported in Table S1.1 (Supplementary Section). The instrument operated in "Intact Protein Mode," with the delta HCD (higher-energy collisional dissociation) vacuum pressure reduced to 0.1. The chromatographic separation was carried out using eluent A (0.1% (v/v) aqueous formic acid (FA)) and eluent B (0.1% (v/v) FA in ACN/water 80/20). The gradient was 0–2 min 5% B, 2–10 min from 5 to 25% B (linear), 10–25 min from 25 to 34% B, 25–45 min from 34 to 70% B, and 45–55 min from 70 to 90% B at a flow rate of 50 ml/min. The injection volume was 19 µl. Full MS experiments were performed in positive ion mode with mass ranging from 400 to 2,000 m/z at a resolution of 120,000 (at 400m/z). Capillary temperature was 275 °C, source voltage 4.0 kV, and S-lens radiofrequency level 69%. In the data-dependent acquisition mode, the five most abundant ions were acquired and fragmented by using collision-induced dissociation (CID) and HCD with 35% normalized collision energy for 10 ms, isolation width of 5 m/z, and activation q of 0.25. HPLC-ESI-MS and MS/MS data were generated by Xcalibur 2.2 SP 1.48 (Thermo Fisher Scientific, CA, United States) using default parameters of the Xtract program for the deconvolution. MS/MS data were analyzed by both manual inspection of the MS/MS spectra recorded along the chromatogram and the Proteome Discoverer (PD) 2.2 software elaboration based on the SEQUEST HT cluster search engine (University of Washington, licensed to Thermo Electron Corporation, San Jose, CA, United States) against the UniProtKB human data bank (188,453 entries, release 2019_03). For peptide matching, high-confidence filter settings were applied: the peptide score threshold was 2.3, and the limits were Xcorr scores >1.2 for singly charged ions and 1.9 and 2.3 for doubly and triply charged ions, respectively. The false discovery rate (FDR) was set to 0.01 (strict)

and 0.05 (relaxed), and the precursor and fragment mass tolerance were 10 ppm and 0.5 Da, respectively. N-terminal pyroglutamination of Glu or Gln residues, phosphorylation on Ser and Thr residues, N-terminal acetylation, oxidation of Met and Trp residues, glutathionylation, nitrosylation, and sulfonic acid of Cys residues were selected as dynamic modifications. Because of the difficulties of the automated software in detecting with high confidence every protein and their fragments, the structural information derives in part from manual inspections of the MS/MS spectra, obtained by both CID and HCD fragmentations, against the theoretical ones generated by MS-Product software (http://prospector.ucsf.edu/prospector/mshome.htm). All the MS/MS spectra were manually verified by using all the fragmentation spectra with an acceptable number of fragment ions. The mass spectrometry proteomics data have been deposited into the ProteomeXchange Consortium (http://www.ebi.ac.uk/pride) via the PRIDE[81] partner repository with the dataset identifier PXD021538.

Quantification and statistical analysis

Quantification of peptides/proteins was performed by using the XIC peak areas measured by HPLC-(LR)-ESI-MS analysis, with the following peak parameters: baseline window 15, area noise factor 50, peak noise factor 50, peak height 15%, and tailing factor 1.5. The estimated percentage error of the XIC analysis 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, as described in a previous study[78]. Then, the corrected XIC peak area values of each peptide/protein were normalized on the TPC of each sample. GraphPad Prism software (version 5.0) was used to calculate means and SDs of the normalized protein XIC peak areas and to perform statistical analysis. Data distributions were tested for normality by the D'Agostino-Pearson test. A comparison between the groups of patients and controls has been performed with Mann–Whitney and Welch-corrected t-tests,

depending on the distribution (skewed or normal) and the variance (unequal or homogeneous). Statistical analysis has been considered significant when the *p*-value was <0.05. A nonparametric ANOVA with the Kruskal–Wallis test and Dunn's post-test was applied to compare the groups of patients treated with different pharmacological therapies.

Preparative RP-HPLC separation for enriched salivary cystatin B

To obtain a cystatin B-enriched sample by a preparative RP-HPLC separation, a total of 6 ml of acid-soluble fraction of human saliva from a healthy donor was injected in a HPLC system UHPLC UltiMate 3,000 (Dionex, Sunnyvale, CA, United States). The chromatographic separation was carried out on an RP Vydac C8 column (Grace, Hesperia, CA, United States), 250 x 10 mm, 5 mm diameter particles, with the following eluents: A, 0.056% aqueous TFA; and B, 0.05% TFA in ACN/water 80:20 (v/v). The applied gradient was from 0 to 60% B in 40 min, from 60 to 100% in 5 min with a flow rate of 2.8 ml/min. RP-HPLC fractions collected during the separation were analyzed by HPLC-ESI-MS, and that containing the cystatin B proteoforms were lyophilized, the powder was suspended in 250 μl of 0.1% TFA, and an aliquot was subjected to a BCA assay (Sigma-Aldrich/Merck, Darmstadt, Germany) to determine the TPC to be used as standard for the immunodetection of cystatin B in the salivary pool from AD and HC groups.

Immune-blotting analyses

For immunoblotting analysis, 25 μg of proteins from 28 AD patients (16 females and 12 males, mean age and SD: 80 ± 6) and from 28 HC (15 females and 13 males, mean age and SD: 78 ± 4) was separately mixed in pools to reach a final TPC of 10 μg/μl for both pools. Dot-blotting was applied for the immunodetection of α-defensins, cystatins A and B, S100A8, S100A9, and Tβ4. 2 μl of each pool diluted 1:10 was blotted in triplicate in a nitrocellulose membrane to

detect α-defensins, S100A8, and S100A9. For detection of cystatins A and B and Tβ4, 2 μl of the 10 μg/μl concentrated pools was used. After 1 hour, the membranes were incubated under stirring for 45 min in the blocking solution composed by 5% blotting-grade blocker (Bio-Rad) in TBS-T (Tris 0.02 M, NaCl 0.15 M 0.05% Tween-20, pH 7.6) and 1 hour with the primary Ab (1:1,000 in TBS-T for α-defensins, S100A8, S100A9, and cystatins A and B and 1:200 for Tβ4). After washing with TBS-T, the membranes were incubated for 1 hour with the secondary rabbit anti-mouse Ab horseradish peroxidase (HRP)-conjugated [IgG (H + L), 1:5,000 in TBS-T for α-defensins, S100A8, S100A9, and cystatins A and B, and IgG2b, 1:50,000 for Tβ4]. Hence, the membranes were treated with the detection solution Clarity Western ECL Substrate (Bio-Rad) according to the manufacturer's instructions. The signals were acquired in high-sensitivity mode with the ChemiDoc MP Imaging System (Bio-Rad) and the images analyzed with the Image Lab software (4.0.1 version). The intensities of the signals, appropriately corrected by the software background subtraction, were normalized with respect to that of cystatin SN, leading to quantitatively unvaried results between patients and controls. Cystatin SN, indeed, is a protein found not variated both according to XIC areas, both according to dot blot analysis (Table S1.2 for XIC areas and Figure S1.1 for dot blot. Supplementary Section). To obtain a dot-blotting in the presence of monoclonal Abs against cystatin SN, the same experimental conditions reported above were applied, except for the secondary HRP-conjugated Ab, which was a rabbit anti-mouse IgG2b, diluted 1:50,000 in TBST. Tβ4 signal normalization was performed with respect to the signal of 0.26 nmol of the standard peptide. Cystatin B signal normalization was performed with respect to the signal of 0.2 μg of the cystatin B-enriched sample.

Table 1.1. Demographic data of HC involved in the study. Demographic, clinical features and pharmacological treatment of AD patients included in the study. For each subject it is indicated the TPC of the acid soluble fraction of saliva determined by BCA assay.

Controls	Sex and age	TPC μg/ml	Patients	Sex and age	TPC μg/ml	Diagnosis	Pharmacological treatment
#1	M, 70	1310.0	#1	M, 82	1860.0	AD	Rivastigmine + Memantine
#2	M, 85	1443.0	#2	F, 80	182.6	AD	Rivastigmine + Memantine
#3	F, 84	1011.0	#3	M, 85	1142.0	AD	Rivastigmine
#4	M, 82	1226.0	#5	F, 63	1116.0	Pre-senile AD	Rivastigmine + Memantine
#5	F, 81	1580.0	#6	M, 78	1853.0	Dementia	Donezepil
#6	F, 81	810.0	#7	M, 85	2316.0	AD	Rivastigmine + Memantine
#7	F, 79	1772.0	#8	F, 81	277.1	AD	Rivastigmine
#8	M, 74	1050.0	#9	F, 78	564.7	AD	Memenatine
#9	M, 71	1366.0	#10	M, 85	580.3	AD	Rivastigmine + Memantine
#10	M, 78	1115.0	#11	F, 80	781.9	AD	Rivastigmine + Memantine
#11	M, 76	1382.0	#13	F, 79	1136.0	AD	Rivastigmine
#12	F, 77	1828.0	#14	F, 82	2051.0	AD	Donezepil
#13	M, 74	891.7	#16	F, 83	827.0	AD	Donezepil + Memantine
#14	M, 87	1376.0	#17	F, 63	442.7	AD	Donezepil
#15	M, 73	638.6	#18	F, 80	756.0	AD	Rivastigmine
#16	F, 81	1082.0	#19	F, 80	867.8	AD	Rivastigmine
#17	F, 82	1108.0	#20	M, 87	302.4	AD	Memantine
#18	F, 72	1047.0	#21	M, 81	472.7	AD	Rivastigmine
#19	F, 86	1115.0	#22	M, 87	484.0	AD	Donezepil
#20	F, 73	2121.0	#23	F, 75	529.1	AD	Rivastigmine
#22	F, 78	792.5	#24	F, 75	163.1	AD	Rivastigmine

Controls	Sex and age	TPC μg/ml	Patients	Sex and age	TPC μg/ml	Diagnosis	Pharmacological treatment
#23	F, 79	1241.0	#25	F, 83	499.7	AD	Rivastigmine
#24	F, 78	648.9	#26	F, 84	791.5	AD	Rivastigmine
#25	F, 75	1636.0	#27	F, 81	873.5	AD	Rivastigmine
#26	M, 75	897.8	#28	F, 84	313.7	AD	Memantine
#27	F, 89	1147.0	#29	F, 92	363.2	AD	Memenatine
#28	M, 78	625.8	#30	M, 86	721.2	AD	Memenatine
#29	M, 73	655.5	#31	M, 77	382.2	AD	Donezepil
#30	F, 76	694.7	#32	F, 88	331.6	AD	Memantine
#31	F, 81	455.6	#33	F, 81	312.3	AD	Rivastigmine
#32	M, 81	411.8	#34	F, 77	863.9	AD	Rivastigmine + Memantine
#33	M, 84	212.0	#35	M, 87	1307.0	AD	Memantine
#34	F, 72	307.3	#36	M, 84	576.4	AD	Memantine
#35	M, 80	722.2	#37	F, 77	1400.0	AD	Rivastigmine
			#38	F, 78	801.8	AD	Rivastigmine

RESULTS

The results obtained in the present study concern the fraction of salivary peptides and proteins soluble in acidic solution and directly analyzable by RP-HPLC-ESI-MS by a top-down approach. The investigated components belong to the following protein families: aPRPs, statherin, histatins, S-type cystatins, cystatins A, B, C, and D, α-defensins, Tβ4, SLPI, S100A proteins and all the variants and PTMs that we have characterized in human saliva by our approach (Table S1.1, Supplementary Section). Overall, 56 proteoforms were investigated in each salivary sample, including modified proteoforms generated by phosphorylation, proteolysis, N-terminal acetylation, methionine or tryptophan oxidation, and cysteine oxidation (formation of disulfide bridges, glutathionylation, cysteinylation, and nitrosylation) as reported in Table S1.1 of Supplementary Section. Protein/peptide levels, measured by MS analysis with a standardized XIC procedure, were compared between AD patients and controls. Considering that TPC was lower in AD patients (807 μg/ml ± 544) than in controls (1,054 μg/ml ± 445) with a *p*-value of 0.01, all the XIC peak areas were normalized with respect to the TPC. Figure 1.1 shows the typical TIC chromatographic profile of a healthy subject involved in the study and obtained by RP-HPLC-(LR)-ESI-MS analysis. The investigated protein families at their different elution time are reported in the chromatogram and it is highlighted, as an example, the XIC procedure used to measure cystatin B S-Glutathionylated (cystatin B SSG). UniProtKB codes, elution times, multiply charged ions (m/z) used for the XIC search and those used for HR-MS/MS characterization, experimental and theoretical Mav and monoisotopic ($[M+H]^+$) mass of all the proteins/peptides included in the study are listed in Table S1.1 of Supplementary Section.

XIC peak area values were used for the label-free quantitation (LFQ) of the proteins and peptides revealed in this study in both the groups, controls and patients. Proteins/peptides showing significantly different levels between the AD and HC groups are reported in Table 1.2, while the results of the statistical analysis of proteins/peptides with similar abundance in the two groups are reported in Table S1.2 (Supplementary Section).

Higher levels in AD group than in HC were observed for statherin and three of its naturally occurring fragments (desF$_{43}$, des1–9, and des1–13), for the P-C peptide, and for Histatin (Hst)-1 proteoforms, both phosphorylated and non-phosphorylated (Table 1.2). No changes in the abundance of other secretory proteins/peptides reported in Table S1.1 (Supplementary Section) were observed. AD patients showed also significantly higher levels of α-defensinss 1–4, Tβ4, cystatin A, and the dimeric and glutathionylated proteoforms of cystatin B (Table 1.2). Moreover, the total S100A8 was more abundant in AD than in HC. The unmodified proteoform of S100A8 was observed in only 9 AD patients and in 2 HC subjects. In 5 AD patients, S100A8 was observed just as a hyperoxidized form (carrying Met$_{1/78}$ and Trp$_{54}$ oxidation and Cys$_{42}$ oxidized to sulfonic acid). The nitrosylated proteoform at Cys$_{42}$ (S100A8-SNO) was observed only in 12 AD patients, and in 7 of these patients was the unique proteoform detected. All the proteoforms of S100A9(S) were more abundant in AD patients, and thus Table 1.2 reports the sum of the XIC peak areas of all the S100A9(S) proteoforms (short and short phosphorylated at Thr$_{108}$ and their correspondent oxidized derivatives at Met$_{89}$ or Met$_{78}$ or Met$_{76}$ or Met$_{58}$, (Table S1, Supplementary Section). Finally, the long glutathionylated S100A9 was significantly more abundant in the AD group than in the HC group. Immunodetection approaches confirmed some of the MS results. Dot-blotting analysis, shown in Figure 1.2 A, evidenced a higher abundance of total α-defensins in AD samples than in HC salivary samples (p = 0.02), in accordance with the MS data obtained considering the sum of the XIC peak area values of all the α-defensins in each sample (p = 0.0005, Figure 1.2 B). Similar results were

obtained for S100A8 and S100A9, and also in this case, dot-blotting was used to immunodetect total S100A8 (Figure 1.3 A) and total S100A9 (Figure 1.3 C). The signal intensity of total S100A8 was significantly higher in the AD salivary pool (p = 0.03) than in the HC one, in accordance with MS data obtained on total S100A8 (Figure 1.3 B). Analogously, total S100A9 (short and long proteoforms) exhibited a signal intensity higher in the AD salivary pool than in the HC one (p = 0.007, Figure 1.3 C). The result of the dot-blotting was in accordance with the MS results obtained on the total S100A9 (p < 0.0001, Figure 1.3 D). Similarly, the immunodetection of Tβ4 confirmed the highest levels of the peptide in AD patients (p = 0.02) obtained by the MS approach (Figures 1.4 A,B). Finally, dot-blotting experiments showed more abundance of total cystatin B (dimeric and monomeric proteoforms) in AD patients than in HC (p = 0.03, Figure 1.5 A), confirming the MS results (Figure 1.5 B). Immunodetection of cystatin A did not provide reliable results, probably because the cystatin A levels were under the sensitivity limits of our method.

The comparison among the three AD patient groups treated with different pharmacological therapies, named G1(ChEI therapy, donepezil 5 mg or rivastigmine 9.5 mg daily), G2 (ChEI in association with memantine), and G3 (antagonist of NMDA receptor, memantine 20 mg daily), was limited by the very low number of subjects of the G2 (n=8) and G3 (n=8) groups and thus provided preliminary results that were reported in the Table 1.3. These data highlighted similar TPC values in the three patient groups (Figure S1.2, Supplementary Section). The ANOVA showed that patients in the G1 and G2 groups, rather than G3, could contribute mainly to the highest levels of the peptides and proteins reported in Table 1.2, with the exception of statherin and its proteoforms found abundant also in the G3 group and P-C peptide, for which G3 was the group with the highest values.

Fig. 1.1: Typical TIC chromatographic profile obtained by RP-(LR)-HPLC-ESI-MS on the salivary acid soluble fraction of a healthy subject (A). XIC peak area of cystatin B SSG is obtained by specific multiply charged ions (m/z) generated by the protein under exam (B). Experimental Mav is obtained by deconvolution of averaged ESI-MS spectra automatically performed by using MagTran 1.0 software (C).

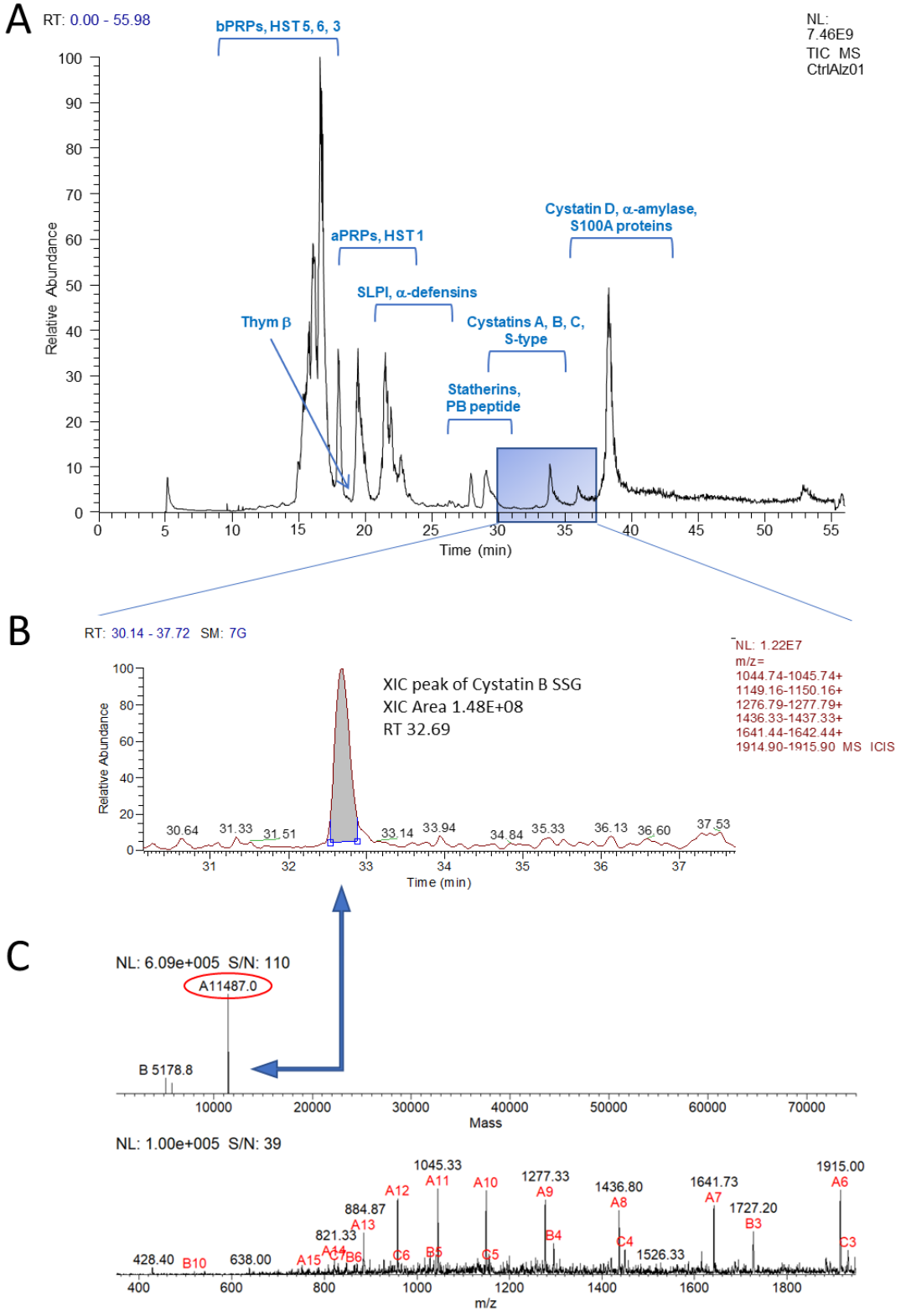


Table 1.2: XIC peak areas values (mean ± SD) normalized on TPC and frequencies of salivary proteins/peptides resulted significantly different between the AD group and HC group. (•), not statistically significant *p*-values; NA, not detected.

Protein/peptide	**XIC Peak Areas x 10^5(mean ± SD) and Frequency**		**p-values**
	AD	**HC**	**AD vs HC**
P-C peptide	9.8 ± 7.7 (33/35)	5.9 ± 4.4 (34/34)	p = 0.02 AD↑
Statherin 2P	9.7 ± 9.6 (35/35)	1.1 ± 0.9 (33/34)	p < 0.0001 AD↑
Statherin desF_{43}	1.5 ± 1.3 (35/35)	0.4 ± 0.7 (33/34)	p < 0.0001 AD↑
Statherin des1-9	0.6 ± 0.6 (29/35)	0.08 ± 0.1 (24/34)	p < 0.0001 AD↑
Statherin des1-13	0.2 ± 0.3 (24/35)	0.08 ± 0.1 (16/34)	p = 0.01 AD↑
Hst-1 1P	1.6 ± 1.5 (33/35)	0.8 ± 0.8 (28/34)	p = 0.01 AD↑
Hst-1 0P	0.3 ± 0.4 (24/35)	0.1 ± 0.2 (17/34)	p = 0.02 AD↑
Cyst A	1.9 ± 1.6 (32/35)	1.0 ± 0.9 (29/34)	p = 0.007 AD ↑
Cystatin B-SSG	0.9 ± 1.2 (29/35)	0.3 ± 0.3 (27/34)	p = 0.006 AD ↑
Cystatin B-SSC	0.2 ± 0.3 (20/35)	0.08 ± 0.1 (18/34)	•
Cystatin B-SS dimer	0.6 ± 0.5 (31/35)	0.2 ± 0.3 (27/34)	p = 0.0006 AD ↑
Cystatin B tot	1.7 ± 1.9 (33/35)	0.6 ± 0.6 (27/34)	p = 0.002 AD ↑
α-defensin 1	1.8 ± 2.0 (32/35)	0.5 ± 0.6 (24/34)	p = 0.0001 AD ↑
α-defensin 2	1.2 ± 1.3 (30/35)	0.4 ± 0.4 (27/34)	p = 0.0009 AD ↑
α-defensin 3	0.6 ± 0.5 (25/35)	0.2 ± 0.3 (17/34)	p = 0.0003 AD ↑
α-defensin 4	0.2 ± 0.3 (13/35)	0.05 ± 0.1 (11/34)	p = 0.01 AD ↑
α-defensin tot	3.9 ± 4.0 (32/35)	1.2 ± 1.5 (24/34)	p = 0.0005 AD↑
Tβ4	0.7 ± 0.8 (25/35)	0.2 ± 0.4 (16/34)	p = 0.01 AD ↑
S100A8	0.6 ± 1.3 (9/35)	0.04 ± 0.2 (2/34)	p = 0.02 AD ↑
S100A8 hyperoxidized	0.2 ± 0.4 (5/35)	0.04 ± 0.2 (3/34)	•
S100A8-SNO	0.8 ± 1.8 (12/35)	NA	NA
S100A8 tot	1.6 ± 2.7 (22/35)	0.08 ± 0.2(5/34)	p <0.0001 AD ↑
S100A9(S) (all proteoforms)	3.3 ± 3.2 (32/35)	0.5 ± 0.8 (16/34)	p <0.0001 AD ↑
S100A9(L)-SSG	1.5 ± 2.4 (22/35)	0.3 ± 0.4 (16/34)	p = 0.004 AD ↑
S100A9 tot	4.8 ± 5.4 (33/35)	0.9 ± 1.2 (19/34)	p <0.0001 AD ↑

Table 1.3: XIC peak areas values (mean ± SD) normalized on TPC, frequencies and p-values obtained by statistical analysis when comparing the three patients' groups treated with different therapies by non-parametric ANOVA with the Krustal-Wallis test and Dunn's post-test. NA, not detected.

Protein/ peptide	XIC Peak Areas x 10^5(mean ± SD) and Frequency				ANOVA p value
	G1 (nr 19)	G2 (nr8)	G3 (nr8)	HC (nr34)	
α-defensin 1	2.2 ± 2.3 (19)	2.3 ± 1.8 (8)	0.6 ± 0.8 (5)	0.5 ± 0.6 (24)	p < 0.0001 (G1vs G3*, G2 vs G3*, G1 vs HC***, G2 vs HC**)
α-defensin 2	1.5 ± 1.5 (19)	1.5 ± 1.2 (7)	0.4 ± 0.6 (5)	0.4 ± 0.4 (27)	p = 0.0003 (G1vs G3*, G1 vs HC**, G2 vs HC*)
α-defensin 3	0.7 ± 0.5 (16)	0.8 ± 0.5 (8)	0.2 ± 0.4 (2)	0.2 ± 0.3 (17)	p < 0.0001 (G1vs G3*, G2 vs G3*, G1 vs HC***, G2 vs HC**)
α-defensin 4	0.2 ± 0.3 (8)	0.4 ± 0.2 (8)	0.04 ± 0.08 (2)	0.05 ± 0.1 (11)	p = 0.0004 (G1vs G2*, G2 vs G3**, G2 vs HC***)
Tβ4	0.5 ± 0.6 (14)	1.1 ± 0.9 (8)	0.2 ± 0.4 (3)	0.2 ± 0.4 (16)	p = 0.002 (G2vs G3**, G2 vs HC**)
Statherin 2P	8.8 ± 7.7 (19)	5.9 ± 4.4 (6)	11.0 ± 7.1 (8)	1.1 ± 0.9 (33)	p < 0.0001 (G1 vs HC***, G2vs HC*, G3 vs HC***)
Statherin Des1-9	0.5 ± 0.6 (16)	0.6 ± 0.9 (6)	0.8 ± 0.6 (7)	0.08 ± 0.1 (24)	p = 0.0002 (G1 vs HC**, G3 vs HC**)
P-C peptide	9.3 ± 5.1 (19)	8.6 ± 12.6 (6)	12.3 ± 6.9 (8)	5.6 ± 4.4 (34)	p = 0.01 (G3 vs HC*)
Cystatin B tot	2.0 ± 2.2 (18)	1.9 ± 1.6 (8)	0.9 ± 1.2 (7)	9.3 ± 5.1 (27)	p = 0.006 (G1vs HC*, G2 vs HC*)
Cystatin A	1.8 ± 1.6 (17)	2.9 ± 1.9 (8)	1.2 ± 0.9 (7)	1.0 ± 0.9 (29)	p = 0.02 (G2 vs HC*)
S100A8 tot	1.1 ± 1.5 (12)	0.9 ± 1.1 (6)	0.4 ± 0.5 (4)	0.08 ± 0.2(5)	p < 0.0001 (G1 vs HC***, G2 vs HC**)
S100A8-SNO	1.1 ± 2.3 (9)	0.7 ± 1.0 (3)	NA	NA	p < 0.0001 (G1 vs G3*, G1 vs HC***)
S100A9(S) tot	1.1 ± 1.5 (18)	0.9 ± 1.1 (7)	0.4 ± 0.5 (7)	0. .5 ± 0.8 (16)	p < 0.0001 (G1 vs HC***, G2 vs HC*)

Fig. 1.2. Dot-blotting immunodetection of total α-defensins in salivary pools of AD and HC samples (A) and distribution plot of the XIC peak area values measured by HPLC-(LR)-ESI-MS of total α-defensins (B) in each salivary sample from the AD and HC groups (XIC peak areas of the identified defensins (Table S1.1, see Supplementary Section) were considered in the sum; *, $p < 0.05$; ***, $p < 0.001$).

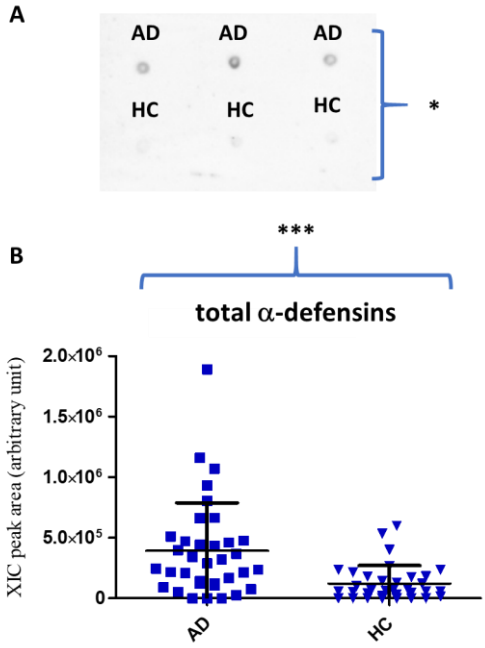


Fig. 1.3. Dot-blotting immunodetection of total S100A8 (A) and total S100A9 (C) in salivary pools of AD and HC samples and distribution plot of the XIC peak area values of total S100A8 (B) and total S100A9 (D) measured by HPLC-(LR)-ESI-MS in each salivary sample from the AD and HC groups (XIC peak areas of all the identified proteoforms of S100A8, such as S100A9, were considered in the sum (Table S1.1); *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

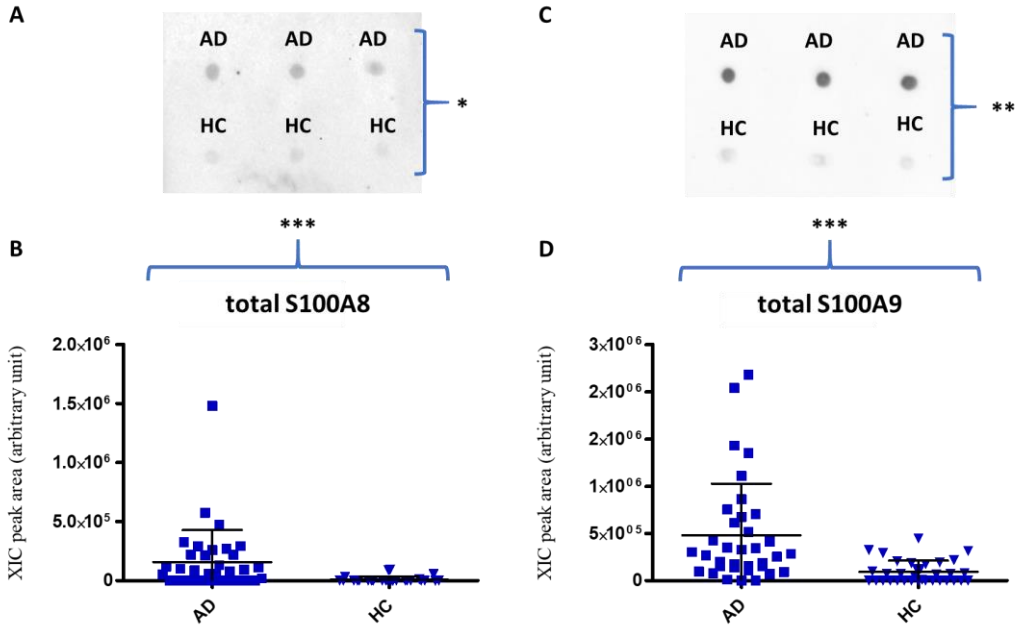


Fig. 1.4. Dot-blotting immunodetection of Tβ4 in salivary pools of AD and HC samples (A) and distribution plot of the XIC peak area values of Tβ4 measured by HPLC-(LR)-ESI-MS (B) in each salivary sample from the AD and HC groups (*, $p < 0.05$; **, $p < 0.01$). Normalization of dots was performed on the signal of the standard Tβ4 at 0.26 nmol (C).

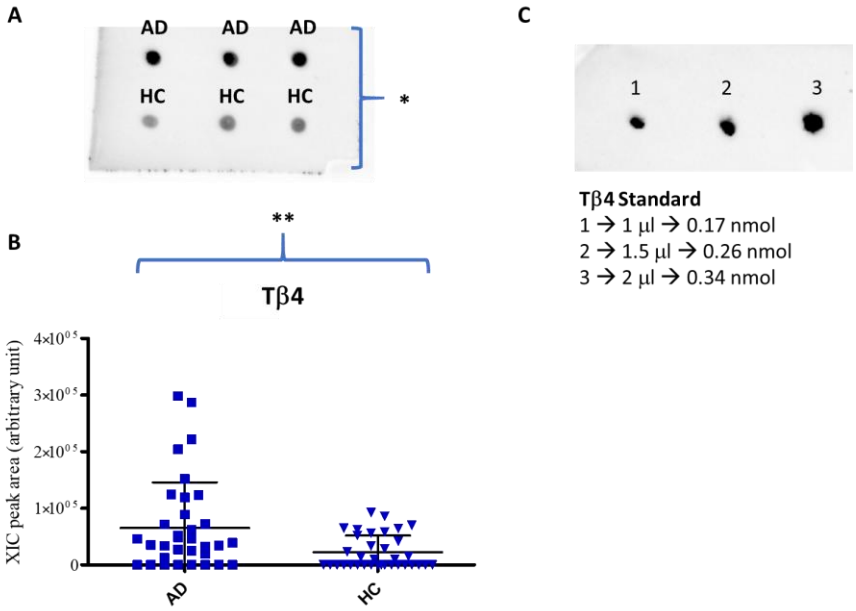
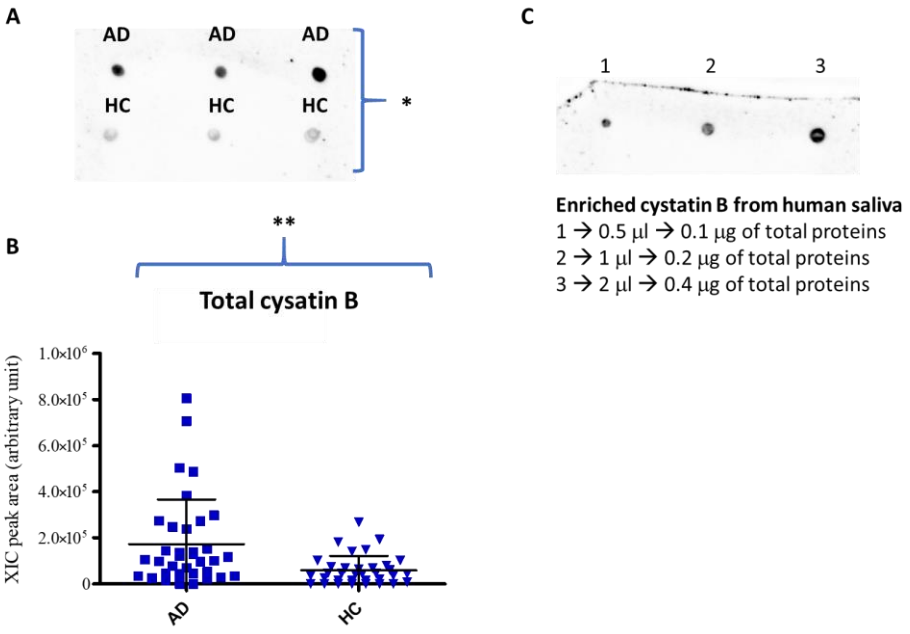


Fig. 1.5. Dot-blotting immunodetection of total cystatin B in salivary pools of AD and HC samples (A) and distribution plot of the XIC peak area values of total cystatin B measured by HPLC-(LR)-ESI-MS (B) in each salivary sample from the AD and HC groups (XIC peak areas of the three identified proteoforms of cystatin B (see Table S1.1, Supplementary Section) were considered in the sum (*, $p < 0.05$; **, $p < 0.01$). Normalization of dots was performed on the signal at 0.2 µg of enriched cystatin B (C).



DISCUSSION

Saliva, as a mirror of oral and systemic health, provides valuable information because it contains not only proteins specifically secreted by the salivary glands[54] but also proteins derived from the gingival crevicular fluid[82,83] from oral microflora[84] and plasmatic proteins transported from blood to saliva by both intracellular and extracellular pathways. The present work evidenced increased levels of some proteins and peptides both secreted and not secreted by the salivary glands from AD patients compared to HCs.

Among the proteins not secreted by the salivary glands, this study highlighted a higher level of proteins with a multifaceted nature, namely, S100A8 and S100A9, α-defensins 1–4, Tβ4, and cystatins A and B. They are involved in many biological functions also at the level of the nervous system, and thus, it is not surprising that they may be implicated in molecular processes associated with AD pathogenesis. S100A8 and S100A9 mainly derive from neutrophils and macrophages, which participate in inflammatory process. During inflammation, S100A8 and S100A9 are actively released and exert a critical role in modulating the inflammatory response by stimulating leukocyte recruitment and inducing cytokine secretion. S100A8 and S100A9 may play a dual role in inflammation, and their pro-inflammatory activity can switch to anti-inflammatory probably depending on the local microenvironment[85]. The results obtained did not allow to establish if the increased levels of S100A9 and S100A8 observed in saliva of AD patients were linked to a pro- or anti-inflammatory role; however, they agree with studies performed by other research groups. Indeed, as a consequence to neuroinflammation, the overexpression of S100A proteins had already been observed in AD where they seem to be involved in several processes related to APP processing, Aβ levels, tau protein PTMs, formation of protein inclusions, and multiple signaling pathways[86]. S100A8 was found to be upregulated in the sera of AD patients[70] and in the hippocampus of mice models of AD[87]. Similarly, S100A9 was found to be strongly increased in brain lysates of both AD patients and

AD mice[71,88] and also in activated microglia and neurons with tau NFTs[89]. Several studies evidenced a strong positive correlation between S100A9 levels and AD pathology: knockdown of S100A9 improved cognition on model mice of AD and reduced global levels of Aβ and APP C-terminal fragments due to decreased activity of the beta-site amyloid precursor protein cleaving enzyme 1 (BACE-1)[71]. Moreover, a correlation between S100A9 and calcium dysregulation in AD has been also evidenced since knockdown of S100A9 significantly diminished the increase of Ca^{2+} levels determined by APP C-terminal fragments or by Aβ[90]. This study on saliva from AD patients highlighted, with respect to HC, higher levels not only of S100A8 and S100A9 but also of their oxidized proteoforms, namely, the hyperoxidized proteoform of S100A8, S100A8-SNO, and glutathionylated long S100A9. Oxidative modifications of proteins are common in neurological disorders since the extracellular milieu is strongly oxidizing as a result of generation of reactive oxygen species (ROS) and reactive nitric oxide species (RNS)[91,92]. Various environmental stimuli are responsible for ROS/RNS production by either inducing neuroinflammation or stimulating/preventing several signaling pathways, which leads to increased or reduced enzyme activity that boosts the addition of oxidative products in neuronal cells[16]. Indeed, high concentrations of Fe^{3+} in NFTs and Aβ aggregates increased the levels of H_2O_2 and advanced glycation end products[91]. Furthermore, it is known that Aβ chemotaxis of microglia and amyloid fibril phagocytosis represent inflammatory stimuli and play a critical role in ROS production[93]. Also, the NADPH oxidase system has been identified as an important source of intracellular ROS, thus playing a key role in the generation of oxidative stress in neurodegenerative diseases[94]. Therefore, it is worth underlining that S100A8 and S100A9 themselves, increasing intracellular NADPH oxidase activity, give a contribution to ROS generation[95]. Oxidation may represent a switch, whereby the modified proteins display anti-inflammatory functions, as demonstrated for S100A8-SNO[96]; S100A9 and S100A8 are able also to act as ROS/RNS scavengers; in particular,

S100A8 is sensitive to oxidative cross-linking and massive oxidation and shows capacity to reduce oxidative damage[97]. Changes in the inflammatory microenvironment, PTMs of S100A8 and S100A9, binding with transition metals (Zn^{2+} and Ca^{2+}), and S100A8/S100A9 heterodimer formation most likely play an important role in the functional switching of these pleiotropic proteins[85]. In addition to exposure to oxidative stress, the microbiota induced neuronal inflammation is a gradually emerging concept promoted by the discovery that brain infections, involving external risk factors such as bacteria or viruses, can trigger Aβ deposition and AD development[17]. In fact, chronic brain infections, caused by various pathogens, have been reported for AD patients[98,99]. In particular, it has been hypothesized that oral and gut microbiota, or their released endotoxins, may alter the permeability of the blood–brain barrier (BBB), facilitate the cerebral colonization by opportunistic pathogens, and induce microglia activation, resulting in increased levels of pro-inflammatory cytokines, which lead to neuronal loss and neurodegeneration[93]. The inflammatory response thereafter may indirectly lead to the upregulation of Aβ production[100].

It has been supposed that neuropathological alterations might be associated with abnormal expression and/or regulation of antimicrobial peptides, including defensins[101], and it is worth noting that Aβ itself exerted antimicrobial activity against bacteria and viruses[102]. Such antimicrobial peptides open an intriguing prospect for the detection and follow-up of such cerebral infection since they, as part of the innate immune system, can efficiently penetrate the BBB and target microbes. Besides their main role as antibacterial and antiviral peptides, defensins also exert numerous immunological effects[103]. This study evidenced higher levels of α-defensins 1–4 in the saliva of AD patients than in HCs. In accordance with these results, Watt and colleagues demonstrated that the levels of α-defensins 1–3 were elevated in both CSF and sera of AD patients, indicating that an inflammatory condition is present not only at brain level but also in other body districts[72]. In the context of AD, it is also fascinating the

antimicrobial and anti-inflammatory activities of Tβ4[104] that was found to be more abundant in the saliva of AD patients than in HCs. Tβ4 is the most abundant Thymosin-β in human tissues, including CNS, where it exerts multiple biological functions, such as downregulation of inflammatory chemokines and cytokines, promotion of cell migration, blood vessel formation, cell survival, stem cell maturation, inhibition of microbial growth, and antiapoptotic factor on gingival fibroblasts[105], as well as neuroprotective and neurodegenerative effects[73,106]. The increased level of Tβ4 in saliva from AD patients is interesting and probably reflects an increase in the brain, in view of the elevated levels of Tβ4 found in reactive microglia of AD patients, where it suppresses the pro-inflammatory signaling[73]. Interestingly, the saliva of AD patients was also characterized by increased levels of cystatin B, which is widely expressed in numerous tissues, including the brain, and of cystatin A, expressed mainly in epidermal cells, lymphoid tissues, and oral squamous epithelia[107]. It has been suggested that cystatins could play a role in AD; indeed, cystatins A and B have been reported to co-localize with amyloid plaques of various origins[76,108]. It is interesting to underline that cystatins B and A, like cystatin C, are considered potential Aβ-binding proteins in vitro[77], able to break down amyloid aggregates in cells, and for this reason, they are also called "amateur chaperones." Indeed, it was demonstrated that cystatin B can bind Aβ and inhibit its fibrillization. Binding is dependent on the oligomeric state[77] of cystatin B, with the tetramer showing the highest affinity. Unfortunately, it was not possible to determine the oligomeric state of cystatin B in saliva. This opens a new perspective for further studies on the salivary proteome of AD patients. Cystatins A and B may also play an important role as regulation factors of inflammation through the inhibition of cathepsins[109]. In particular, cystatin B, the main natural inhibitor of cathepsin B, may exert a neuronal protective key role in AD. In fact, it should be highlighted that the chronic systemic exposure to lipopolysaccharide produced by periodontal bacteria may result in an overexpression of cathepsin B, which has been recently demonstrated to play a critical role in

initiating neuroinflammation and neural dysfunction[110]. Cathepsin B is a beta secretase enzyme, and similar to the BACE-1, it cleaves APP at the beta-cleavage site, creating amyloid fragments[111]. Besides the principal role of cystatin B as an inhibitor of cathepsin B, recent data show additional interesting roles in maintenance of cell homeostasis, reduction of oxidative stress[112], and prevention of apoptosis[113].

Saliva of AD patients compared to the HC group also showed higher levels of some peptides secreted by the salivary glands that are involved in the metal homeostasis and defense of the oral cavity, namely, Hst-1, both phosphorylated and not phosphorylated, and statherin. Hst-1 is a peptide with multifaceted actions, including bacterial and host protease inhibitory properties, wound healing processes, and stimulation of cell migration[114]. Moreover, Hst-1, similar to cystatin B, is able to inhibit bacterial lipopolysaccharide activity and consequently inflammatory cytokines and production of other pro-inflammatory factors[115]. Indeed, Hst-1 together with other salivary histatins, defensins and cathelicidins, represent the main family of antimicrobial peptides in mammals[116]. Histatins carry two metal-binding motifs in the N-terminal domain, by which they exert their antimicrobial action: the Cu(II)/Ni(II) binding (ATCUN) motif and the Zn(II)-binding motif HEXXH[117]. Interestingly, Huang *et al.* evidenced the presence of a similar Zn(II)-binding motif (EXHH) in Aβ 1–40 amyloid peptides[118]. Antibacterial activities against oral pathogens have been demonstrated also for statherin and its C-terminal fragments[119], which, in particular, retain the specific binding sites for Porphyromonas gingivalis, the key-stone pathogen in periodontitis[120]. The increased level of these antimicrobial peptides in saliva of AD patients can be linked to an oral dysbiosis, which is often observed in such patients. It is worth noting that recent epidemiological studies have identified a strong association between the periodontal pathogen *P. gingivalis* and neurodegeneration. *P. gingivalis* has been detected in the brain autopsy and CSF of individuals diagnosed with AD[121]. Another recent study demonstrated the presence of *P. gingivalis* derived

lipopolysaccharides in brain samples from AD patients[122] and the presence of gingipains, a class of *P. gingivalis* proteases, in neurons, tau tangles, and Aβ of individuals with AD[123]. Oral dysbiosis in elderly people with AD was frequently observed as well as the propensity to periodontal problems; AD individuals may also suffer from xerostomia and oral lesions, such as stomatitis and candidiasis[124]. These conditions have been associated with a decrease of submandibular salivary flow in patients with AD without medications[125], which is considered an effect of the disease itself or associated with the anticholinergic therapy often applied[124]. Nevertheless, the results obtained in the present study did not appear linked to pathological conditions occurring in the oral cavity, since the enrolled patients were selected on the basis of absence of manifest oral diseases. Moreover, these results could not be considered a consequence of glandular flow variations due to the therapy; indeed, no statistically significant differences in the TPC were measured among the three groups of patients with diverse drug treatments (G1, G2, and G3). Additionally, among all the peptides/proteins of glandular origins, only some peptides, such as statherin, Hst-1, and P-C peptide, changed their abundance in the saliva of AD patients with respect to the controls. This evidence suggested a connection between the attendance of the AD and the overexpression and/or oversecretion of these specific peptides. Finally, preliminary results from the comparison among G1, G2, and G3 groups appeared to display that almost all proteoforms found at high levels in the AD saliva were more concentrated in G1 and G2, treated with ChE inhibitors, than in G3 group; this last group, treated with an NMDA receptor antagonist, showed levels similar to those of the HC group. Only statherin proteoforms and P-C peptides were found at the highest levels in the G3 group. A further investigation focused on implications of the pharmacological therapy is desirable with a greater number of patients. The data collected in the present study did not allow us to demonstrate that infections and/or inflammation occurred in AD patients at the moment of the sample collection or to affirm that these results were a consequence of previous injury events.

Certainly, peptides and proteins involved in the innate immune-protection, both specific of the oral cavity and also expressed in other body districts, were more abundant in salivary samples of AD patients, particularly proteins acting as ROS/RNS scavengers and with a neuroprotective role, such as S100A8, S100A9, and cystatin B; proteins with antimicrobial activity, such as α-defensins, cystatins A and B, Hst-1, statherin, and Tβ4; and peptides involved in the homeostasis of the oral cavity. Almost all these peptides/proteins are able to act as anti-inflammatory factors, and among these, Tβ4 has an important regulatory role at the microglia level. Moreover, several experimental and clinical data confirm a key role of oral and gut dysbiosis in neurodegeneration. The merging of oral-/gut derived inflammatory response together with aging and poor diet in the elderly may contribute to the pathogenesis of AD. The results obtained in this study support the view that, in the AD patients under study, several protective mechanisms against pathogen-targeting agents generating ROS and RNS, infections, and inflammatory conditions have been established.

PART I SUPPLEMENTARY SECTION

Table S1.1: UniProt-KB code, experimental and theoretical Mav ± SD, Mav and monoisotopic ([M+H]$^+$), and elution times of proteins and peptides analyzed. Table reported also the m/z values and charge of the multiply-charged ions selected for XIC search in HPLC-LR-ESI-MS, and those ones used for HR-MS/MS characterization.

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) [M+H]$^+$ ± SD	m/z (charge) for MS/MS	PTMs
Acid Proline-Rich Proteins						
PRP-1 2P (P02810)	22.2	15515 ± 2 (15514-15515)	1293.9 (+12) 1194.4 (+13) 1035.3 (+15) 970.7 (+16) 913.6(+17)	15506.1± 0.1 (15506.24-15507.22)	1293.19(+12) 1034.75(+15) 970.14(+16) 862.46(+18)	N-Term(Gln->pyro-Glu), S_8(Phospho), S_{22}(Phospho)
PRP-1 1P	22.9	15435 ± 2 (15434-15435)	1287.2 (+12) 1188.3 (+13) 1030.0 (+15) 965.7 (+16) 908.9 (+17)	15426.3 ± 0.1 (15426.27-15427.26)	1286.52 (+12) 1187.64 (+13) 1029.42 (+15)	N-Term(Gln->pyro-Glu), S_8 or S_{22}(Phospho)
PRP-1 0P	23.2	15355 ± 2 (15354-15355)	1280.5 (+12) 1182.1 (+13) 1024.6 (+15) 960.7 (+16) 904.2 (+17)	15346.3± 0.1 (15346.31-15347.29)	903.72(+17) 853.57(+18)	N-Term(Gln->pyro-Glu)
PRP-1 3P	21.6	15595 ± 2 (15594-15595)	1418.7 (+11) 1300.5 (+12) 1200.6 (+13) 1040.6 (+15) 975.7 (+16)	15586.3± 0.1 (15586.21-15587.19)	1732.80(+9) 1114.30(+14) 866.90 (+18) 1040.08(+25)	N-Term(Gln->pyro-Glu) S_8, S_{17}, S_{22} (Phospho)

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) $[M+H]^+$ ± SD	m/z (charge) for MS/MS	PTMs
PRP-3 2P (P02810)	22.8	11161 ± 1 (11161-11162)	1595.5 (+7) 1396.2 (+8) 1015.7 (+11) 931.1 (+12) 859.6 (+13)	11155.10± 0.07 (11156.08-11157.06)	1015.19 (+11) 930.67 (+12) 797.86 (+14) 744.74 (+15)	N-Term(Gln->pyro-Glu) S_8, S_{22} (Phospho) Fragment 1-106 of PRP-1
PRP-3 1P	23.4	11081 ± 1 (11081-11082)	1584.1 (+7) 1386.2 (+8) 1008.4 (+11) 924.5(+12) 853.4 (+13)	11076.00± 0.07 (11076.11-11077.09)	1385.51 (+8) 1008.61 (+11) 924.01 (+12) 853.01 (+13)	N-Term(Gln->pyro-Glu) S_8 or S_{22} (Phospho)
PRP-3 0P	23.8	11001 ± 1 (11001-11002)	1376.2 (+8) 1101.2 (+10) 917.8 (+12) 786.8 (+14)	10996.01± 0.07 (10996.14-10997.13)	1000.56 (+11) 786.37 (+14)	N-Term(Gln->pyro-Glu)
PRP-3 2P desR_{106}	22.8	11004 ± 1 (11005-11006)	1573.2 (+7) 1223.8 (+9) 1001.5 (+11) 847.6 (+13)	10999.85± 0.07 (10999.97-11000.96)	1001.00 (+11) 917.66 (+12)	N-Term(Gln->pyro-Glu) S_8, S_{22} (Phospho), R_{106} removal
P-C peptide (P02810)	15.0	4370.9 ± 0.4 (4370.8)	1457.9 (+3) 1093.7 (+4)	4369.19 ± 0.02 (4369.18)	1093.05(+4)	Fragment 107-150 of PRP-1
Statherin						
Statherin (P02808)	29.2	5380.0 ± 0.5 (5379.7)	1794.2 (+3) 1345.9 (+4) 1076.9 (+5)	5377.46 ± 0.03 (5377.45)	1345.12 (+4)	S_2(Phospho); S_3(Phospho)

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) $[M+H]^+$ ± SD	m/z (charge) for MS/MS	PTMs
Statherin 1P	28.9	5299.9 ± 0.5 (5299.7)	1767.6 (+3) 1325.9 (+4) 1060.9 (+5)	5297.50 ± 0.03 (5297.48)	1325.13 (+4)	S_3(Phospho)
Fr. des-F_{43}	27.8	5232.4 ± 0.5 (5232.5)	1745.1 (+3) 1309.1 (+4) 1047.5 (+5)	5230.38 ± 0.03 (5230.38)	1308.35 (+4)	C-Term F_{43}removal
Fr. desT_{42}-F_{43}	27.9	5131.2 ± 0.5 (5131.4)	1711.4 (+3) 1283.8 (+4) 1027.2 (+5)	5129.33 ± 0.03 (5129.33)	1283.09 (+4)	C-Term T_{42}-F_{43}removal
Fr. desD_1	28.7	5264.7 ± 0.5 (5264.6)	1755.9 (+3) 1317.2 (+4) 1053.9 (+5)	5262.41 ± 0.03 (5262.42)	1316.36 (+4)	N-Term D_1removal
Fr. des1-9	28.5	4127.9 ± 0.4 (4127.6)	1376.9 (+3) 1032.9 (+4)	4125.99 ± 0.02 (4125.99)	1376.00 (+3)	N-Term 1-9 residue removal
Fr. des1-10	28.0	3971.3 ± 0.4 (3971.4)	1986.7 (+2) 1324.8 (+3)	3969.90 ± 0.01 (3969.89)	1323.97 (+3)	N-Term 1-10 residue removal
Fr. des1-13	27.5	3645.2 ± 0.4 (3645.0)	1823.6 (+2) 1216.1 (+3)	3643.68 ± 0.01 (3643.68)	1215.56 (+3) 911.92 (+4)	N-Term 1-12 residue removal
P-B peptide						
P-B peptide (P02814)	30.0	5792.9 ± 0.5 (5792.7)	1932.0 (+3) 1449.2 (+4) 1159.6 (+5)	5790.06 ± 0.03 (5790.04)	965.85(+6)	N-Term(Gln->pyro-Glu)

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) $[M+H]^+$ ± SD	m/z (charge) for MS/MS	PTMs
Fr. des1-4	30.0	5371.0 ± 0.5 (5371.3)	1791.4 (+3) 1343.8 (+4) 1075.3(+5)	5368.84 ± 0.03 (5368.82)	1343.21 (+4) 1074.77 (+5)	N-Term 1-4 residue removal
Fr. des1-5	30.3	5215.0 ± 0.5 (5215.1)	1739.4 (+3) 1304.8 (+4) 1044.0 (+5)	5212.75 ± 0.03 (5212.73)	1303.94 (+4)	N-Term 1-5 residue removal
Fr. des1-7	30.1	5060.1 ± 0.5 (5060.9)	1688.0 (+3) 1266.2 (+4) 1013.2 (+5)	5058.67 ± 0.03 (5058.65)	1265.42 (+4)	N-Term 1-7 residue removal
Fr. des1-12	27.5	4549.0 ± 0.5 (4549.3)	1517.5 (+3) 1138.3 (+4)	4547.10 ± 0.02 (4547.41)	1137.85 (+4) 910.48 (+5)	N-Term 1-12 residue removal
Histatins						
Hst-1 (P15515)	21.9	4928.2 ± 0.5 (4928.2)	1644.1 (+3) 1233.5 (+4)	4926.21 ± 0.02 (4926.20)	704.61 (+7)	S_2(Phospho)
Hst-1 0P	22.0	4848.2 ± 0.5 (4848.2)	1617.4 (+3) 1213.5 (+4)	4846.24 ± 0.02 (4846.23)	693.18 (+7)	
Hst-3 (P15516)	17.7	4062.2 ± 0.4 (4062.4)	1355.1 (+3) 1016.6 (+4)	4060.98 ± 0.02 (4060.98)	813.00 (+5) 677.67 (+6)	
Hst-3 1/25	14.3	3192.4 ± 0.3 (3192.5)	1065.1 (+3) 799.1 (+4)	3191.62 ± 0.01 (3191.62)	532.78 (+6) 456.81 (+7)	Fragment 1-25 of Hst-3

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) $[M+H]^+$ ± SD	m/z (charge) for MS/MS	PTMs
Hst-3 1/24	14.6	3036.5 ± 0.3 (3036.3)	1013.2 (+3) 760.1 (+4)	3035.53 ± 0.01 (3035.52)	607.91 (+5) 506.76 (+6)	Fragment 1-24 of Hst-3
Cystatins						
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)	11000.65± 0.07 (11000.67)	1375.96 (+8) 847.13 (+13) 786.69 (+14)	
A Nα-Ac.	33.0	11047.43 ± 2 (11048.5)	1005.41 (+11) 1105.85 (+10) 1228.61 (+9) 1382.06 (+8) 1579.36 (+7) 1842.42 (+6)	11042.55± 0.07 (11042.68)	1381.21 (+8) 1227.85 (+9) 789.69 (+14)	N-Term-α-Acetylation
B-SSG (P04080)	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)	11480.69± 0.07 (11480.68)	1149.58 (+10) 1045.25 (+11)	C_3Glutathionylation
B-SSC	32.9	11299.8 ± 2 (11300.7)	1884.5 (+6) 1615.4 (+7) 1413.6 (+8) 1256.7 (+9)	11294.54 ± 0.07 (11294.61)	-	C_3Cysteinylation

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) $[M+H]^+$ ± SD	m/z (charge) for MS/MS	PTMs
			1131.1 (+10) 1028.6 (+11)			
B S-S dimer	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)	nd		1 interchain disulfidebridge
C (P01034)	35.1	13342 ± 2 (13343.1)	1483.57 (+9) 1335.32 (+10) 1214.02 (+11) 1112.93 (+12) 1027.40 (+13)	13335.32 ± 0.08 (13335.58)	1112.21 (+12) 834.41 (+16)	2 intrachain disulfidebridges
D-R_{26} 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)	13509.43± 0.08 (13509.65)	1501.96 (+9) 1351.86 (+10) 965.90 (+14)	N-Term(Gln->pyro-Glu) after 1 residue removal, 2 intrachain disulfidebridges

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) $[M+H]^+$ ± SD	m/z (charge) for MS/MS	PTMs
Cystatins S-type						
S1 (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)	14256.66± 0.09 (14256.77)	1296.98 (+11) 1097.60 (+13) 1019.27 (+14)	S_3(Phospo) on cystatin S, 2 intrachain disulfide bridges
S1ox	35.3	14281 ± 2 (14280.7)	1786.40 (+8) 1589.70 (+9) 1429.30 (+10) 1299.50 (+11) 1191.30 (+12) 1099.70 (+13)	14272.66 ± 0.09(14272.77)	1428.18 (+10) 1298.43 (+11) 1190.31 (+12) 892.99 (+16)	S_3(Phospo), W_{23} oxidation, 2 intrachain disulfide-bridges
S2	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)	14336.58 ± 0.09 (14336.74)	1434.57 (+10) 1195.65 (+12) 896.98 (+16)	S_1, S_3 (di-Phospho) on cystatin S 2 intrachain disulfide-bridges
SN (P01037)	34.6	14312 ± 2 (14313)	1790.0 (+8) 1591.2 (+9) 1432.2 (+10) 1302.1 (+11) 1193.7 (+12)	13304.03 ± 0.09 (13304.09)	951.22 (+14) 740.06 (+18)	2 intrachain disulfide-bridges

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) $[M+H]^+$ ± SD	m/z (charge) for MS/MS	PTMs
			1101.9 (+13) 1023.3 (+14)			
SNox	34.6	14328 ± 2 (14328)	1792.30 (+8) 1593.20 (+9) 1434.00 (+10) 1303.30 (+11) 1195.20 (+12) 1103.30 (+13)	14320.10± 0.09 (14320.09)	1302.74 (+11) 1194.26 (+12) 1102.47 (+13) 955.61 (+15)	W_{23}oxidation
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)	14338.02± 0.09 (14338.01)	1103.85 (+13) 1025.07 (+14) 956.80 (+15) 897.06 (+16) 844.35 (+17) 797.50 (+18)	1 intrachain disulfide-bridge
Antileukoproteinase						
SLPI (P03973)	26.2	11702.2 ± 1 (11706)	1952.64 (+6) 1673.84 (+7) 1464.73 (+8) 1302.10 (+9)	11703.29 ± 0.07 (11702.36)	1301.26 (+9) 1171.24 (+10)	8 intrachain disulfide bridges
α-Defensins						
α-defensin 1 (P59665)	23.5	3442.5 ± 2 (3442.1)	1772.03 (+2) 1148.36 (+3) 861.52 (+4)	3440.45 ± 0.01 (3440.52)	861.13 (+4) 689.10 (+5)	2 intrachain disulfide-bridges

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) $[M+H]^+$ ± SD	m/z (charge) for MS/MS	PTMs
α-defensin 2 (P59665/6)	23.5	3370.4 ± 1 (3370.9)	1686.49 (+2) 1124.66 (+3) 843.75 (+4)	3370.41 ± 0.01 (3370.44)	1124.48 (+3) 843.61 (+4) 675.09 (+5)	2 intrachain disulfide-bridges
α-defensin 3 (P59666)	23.5	3485 ± 2 (3486.1)	1744.03 (+2) 1163.03 (+3) 872.52(+4)	3484.53 ± 0.01 (3484.51)	872.13 (+4)	2 intrachain disulfide-bridges
α-defensin 4 (P12838)	27.2	33708 ± 1 (3709.4)	1855.71 (+2) 1237.48 (+3) 928.36(+4)	3707.68 ± 0.01 (3707.77)	927.94 (+4)	2 intrachain disulfide-bridges
Thymosins β4						
Tβ4 (P62328)	18.5	4964.0± 1 (4963.5)	1655.51 (+3) 1241.88 (+4) 993.71 (+5)	4961.50 ± 0.02 (4961.48)	877.76 (+6)	N-Term-α-Acetylation
S100A proteins						
S100A12 (P80511)	40.0	10444 ± 2 (10443.9)	1306.5 (+8) 1161.4 (+9) 1045.4 (+10) 950.4 (+11)	10438.74±0.07 (10438.49)	1044.75 (+10) 803.88 (+13) 696.83 (+15)	M_1 removal
S100A7 (P31151)	37.0	11367 ± 2 (11367.8)	1422.0 (+8) 1264.1 (+9) 1137.8 (+10) 1034.4 (+11)	11360.43 ± 0.07(11361.52)	1263.28 (+9) 1033.77 (+11) 947.71 (+12) 874.89 (+13)	M_1 removal, N-Term-α-Acetylation, D_{27} variant

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) $[M+H]^+$ ± SD	m/z (charge) for MS/MS	PTMs
S100A8 (P05109)	40.4	10833 ± 2 (10834.5)	1355.3 (+8) 1204.8 (+9) 1084.5 (+10) 985.9 (+11)	10826.60 ± 0.07 (10828.657)	985.42 (+11) 903.39 (+12) 833.97 (+13)	
S100A8 hyperoxidized	39.3	10915 ± 2 (10914.6)	1365.3 (+8) 1213.7 (+9) 1092.5 (+10) 993.2 (+11)	10908.40 ± 0.07 (10908.631)	840.05 (+13)	C_{42}-SO_3H and W_{54} dioxidation or C_{42}-SO_3H and W_{54} oxidation and $M_{1/78}$ oxidation
S100A8 SNO	40.8	10863 ± 2 (10863.5)	1358.9 (+8) 1208.1 (+9) 1087.3 (+10) 988.6 (+11)	10858.67 ± 0.07(10857.647)	836.13 (+13) 776.47 (+14) 724.78 (+15)	C_{42} nitrosylation
S100A9(S) (P06702)	42.2	12690 ± 2 (12689.2)	1410.9 (+9) 1269.9 (+10) 1154.6 (+11) 1058.4 (+12) 977.1 (+13)	12682.21 ± 0.08 (12682.293)	1410.03 (+9) 1269.13 (+10) 1057.77 (+12) 976.48 (+13) 846.42 (+15) 705.52 (+18)	N-Term-α-Acetylation after 1-5 residue removal
S100A9(S) 1P	42.2	12770 ± 2 (12769.2)	1419.8 (+9) 1277.9 (+10) 1161.8 (+11) 1065.1 (+12) 983.3 (+13)	12762.05 ± 0.08 (12762.26)	912.96 (+14) 852.29 (+15) 798.96 (+16)	N-Term-α-Acetylation after 1-5 residue removal, T_{108} (Phospho)
S100A9(S) M-ox	41.3	12706 ± 2 (12705.2)	1412.7 (+9) 1271.5 (+10) 1156.0 (+11)	12698.39 ± 0.08 (12698.288)	907.95 (+14) 847.49 (+15)	N-Term-α-Acetylation after 1-5 residue removal,M_{89} or $_{78}$ or $_{76}$ or $_{58}$ oxidation

Proteins/peptides	El. time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	Exper. (theor) $[M+H]^+$ ± SD	m/z (charge) for MS/MS	PTMs
			1059.8 (+12) 978.3 (+13)			
S100A9(S) M-ox 1P	41.3	12786 ± 2 (12785.2)	1421.9 (+9) 1279.5 (+10) 1163.3 (+11) 1066.4 (+12) 984.5 (+13)	12778.21 ± 0.08 (12778.255)	1420.70 (+9) 983.87 (+13) 913.66 (+14)	N-Term-α-Acetylation after 1-5 residue removal, T_{108} (Phospho), M_{89} or $_{78}$ or $_{76}$ or $_{58}$ oxidation
S100A9(L) SSG	41.5	13459 ± 2 (13458.1)	1346.8 (+10) 1224.5 (+11) 1122.5 (+12) 1036.3 (+13) 962.3 (+14)	13450.51 ± 0.08 (13450.55)	1035.58 (+13) 961.68 (+14) 841.60 (+16) 792.15(+17) 748.20 (+18)	M_1 removal, N-Term-α-Acetylation, C_2glutathionylation

Fig. S1.1: Immune-detection analysis of Cystatin SN by dot-blotting of a triplicate of a salivary pools of AD samples and HC samples (A), statistical analysis of the spot intensity of the AD and HC spots (B). Cystatin SN signal was used to normalize the dot blot intensity signals of total α-defensins, total S100A8 and total S100A9.

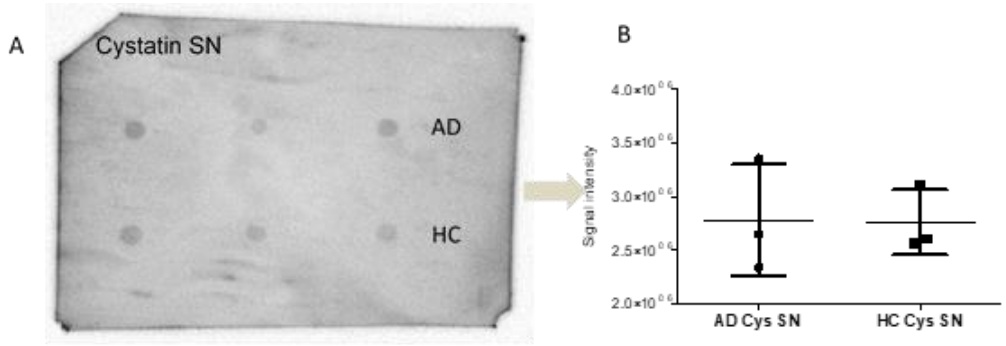


Fig. S1.2: TPC measured in the subgroups of patients with different pharmacological treatment.

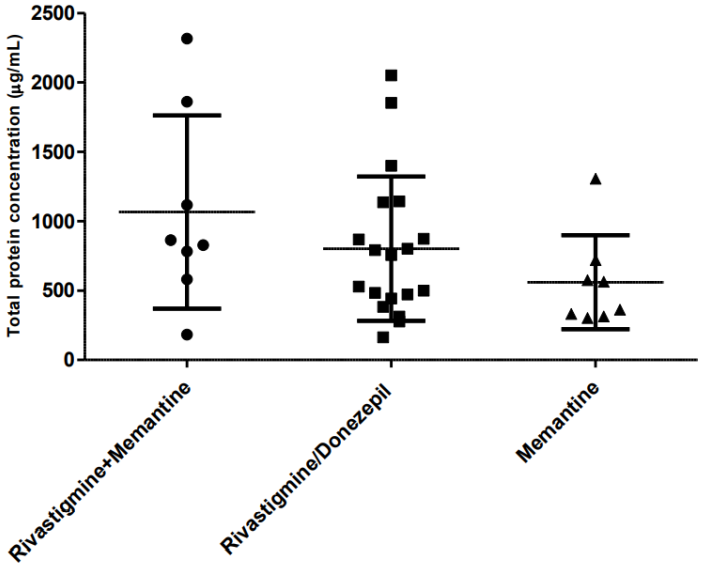


Table S1.2: XIC peak areas values (mean ± SD) normalized on TPC and frequencies of proteins/peptides from glandular origin resulted with similar levels between AD group and HC group. (•), not statistically significant *p*-values.

Protein/peptide	XIC Peak Areas x 10^5(mean ± SD) and Frequency		p-values
	AD	HC	AD vs HC
PRP-1 2P	37.6 ± 27.2 (34/35)	28.8 ± 18.7 (34/34)	•
PRP-1 1P	5.0 ± 4.3 (34/35)	3.8 ± 3.5 (31/34)	•
PRP-1 0P	0.2 ± 0.6 (8/35)	0.3 ± 0.3 (19/34)	•
PRP-1 3P	0.09 ± 0.2 (8/35)	0.3 ± 0.4 (16/34)	•
PRP-3 2P	13.0 ± 10.4 (35/35)	9.4 ± 6.7 (34/34)	•
PRP-3 1P	2.1 ± 1.9 (35/35)	1.5 ± 1.1 (32/34)	•
PRP-3 0P	0.06 ± 0.1 (10/35)	0.04 ± 0.07 (9/34)	•
PRP-3 2P, desR_{106}	1.8 ± 1.8 (31/35)	2.4 ± 2.2 (28/34)	•
Stath. 1P	0.3 ± 0.4 (29/35)	0.3 ± 0.2 (24/34)	•
Stath. desT_{42}-F_{43}	0.4± 0.4 (26/35)	0.2 ± 0.2 (30/34)	•
Stath. desD_1	0.4 ± 0.5 (32/35)	0.3 ± 0.3 (30/34)	•
Stath. des1-10	0.3 ± 0.3 (26/35)	0.2 ± 0.2 (27/34)	•
P-B peptide	7.8 ± 6.6 (35/35)	6.4 ± 4.4 (34/34)	•
P-B des1-4	0.6± 0.8 (2/35)	0.7 ± 0.6 (32/34)	•
P-B des1-5	0.8 ± 0.8 (34/35)	1.0 ± 1.1 (34/34)	•
P-B des1-7	1.3 ± 2.3 (34/35)	0.8 ± 0.5 (34/34)	•
P-B des1-12	0.7 ± 0.7 (34/35)	0.8 ± 1.3 (31/34)	•
SLPI	0.1 ± 0.2 (13/35)	0.06 ± 0.09 (13/34)	•
Hst-3	0.6 ± 1.0 (18/35)	0.4 ± 0.7 (18/34)	•
Hst-3 1/25	0.2 ± 0.3(9/35)	0.2 ± 0.3 (17/34)	•
Hst-3 1/24	1.1 ± 1.5 (20/35)	0.6 ± 0.9 (10/34)	•
Cyst A Nα-Ac	0.2 ± 0.2 (24/35)	0.2 ± 0.5 (23/34)	•
Cyst. C	0.1 ± 0.4 (6/35)	0.08 ± 0.2 (5/34)	•
Cyst. D-R_{26} des1-5	0.3 ± 0.4 (13/35)	0.3 ± 0.4 (16/34)	•
Cyst S1 tot	6.4 ± 8.8 (31/35)	4.1 ± 3.7 (30/34)	•
Cyst. S2	1.6 ± 2.6 (20/35)	1.2 ± 1.2 (26/34)	•
Cyst. SN tot	12.3 ± 21.3 (32/35)	6.7 ± 6.6 (29/34)	•
Cyst. SA	1.4 ± 3.4 (8/35)	0.5 ± 1.2 (7/34)	•
S100A7	0.3 ± 0.7 (13/35)	0.2 ± 0.3 (11/34)	•

PART II

Comparative salivary proteomics among adult, elderly and Alzheimer individuals

INTRODUCTION AND AIM OF THE STUDY

Aging is a complex biological process characterized by a steady decline in different physiological functions that can lead to physical and cognitive impairment in humans[126]. Aging, indeed, is a risk factor for many pathologies, including neurodegeneration, cancer, osteoarthritis and others, so the availability of true biomarkers of biological age has been pointed as highly relevant to improve the clinical health and to succeed in the early identification of patients at high age-related risk[127]. In humans, APOE gene, which is strictly implicated in AD, appeared to be the most representative contributor to lifespan[128]. However, genetic studies showed large limits in the understanding of molecular mechanisms that influence aging and longevity due to individual variability[128], while the potential of proteomic and metabolomic studies has been suggested[127,128]. Indeed, the search for metabolites and proteins involved in lifespan has recently gained attention, revealing how proteomic studies can be highly relevant in the study of aging clock[129]. Johnson and colleagues performed a systematic review on 36 studies involving a total of 3301 subjects aged 18-76 years old[130], mainly focused on proteomic investigation of blood, plasma, liver, bone marrow, skin, urine. The results highlighted the increased expression of 23 proteins, *e.g.* vascular endothelial growth factor, pleiotrophin, fibrinogen alpha, in older subjects[130]. Proteomic studies also highlighted the increase in the content of proteins involved in iron transport, homeostasis, immune response and apoptosis in the plasma from neonatal age to adulthood[129], as well as the potential correlation between proteins post-translational modifications (PTMs) and age or age-related diseases[131]. Human saliva composition varies between individuals depending on a multitude of factors, including sex, health status, circadian rhythms, habits, nutritional factors, and age[50]. Age-related changes in salivary proteome were highlighted by several studies performed on populations from 180 days after birth to adulthood[132–136]. Anyways, to date, few proteomic studies concerning the effect of age on salivary protein profile have been

performed on populations older than 60 years and they are summarize in Table 2.1. Johnson and colleagues evidenced, by cationic gel-electrophoresis procedure, a significant decrease in histatins levels[137]; mucin (MUC) 1 and 2 using radiolabeled SDS-page[138] and real-time PCR[139], lactoferrin and transferrin by ELISA[140], and peroxidase activity by spectrophotometry[141] were also found decreased in elderly. Conversely, increased levels of lysozyme, amylase, albumin and sIgA were observed by sialometrical analysis[142] and sIgA enzyme immune assay[140]. Manconi and colleagues evidenced, with a top-down proteomic approach, high levels of cystatins A and B and small proline-rich protein 3 in old edentulous subjects with respect to younger subjects[143]. In a recent review, telomere length, DNA methylation, MUC1 expression and protein carbonylation levels, have been indicated as the main biological hallmarks measurable in saliva and related to aging process[144], but further investigation about the changes in salivary protein composition in advanced age are essential[144,145]. Indeed, defining the salivary protein profile typical of elderly subjects could be useful in the investigation of biomarkers of age-related pathologies, as AD, especially considering the globally increasing size of old population and the high risk of age-related diseases. Thus, there is a critical need of deeply investigate the salivary proteome in the elderly to better understand if and how it can reflect the biological age of subjects and aging-related risks. In this study, we performed a statistical investigation, based on exact Mann-Whitney t-test, Random Forest (RF), Multidimensional Scaling (MDS) and Hierarchical Cluster Analysis (HCA), which compared the salivary protein profiles, detectable by HPLC-ESI-IT-MS, of adult (under 70 years old) and elderly (over 70 years old) subjects. This kind of analysis was chosen to obtain a classification of the subjects into the two groups depending on their salivary proteome and thus highlighting that the changes at proteomic level in oral cavity are related to the biological age of the subjects. Moreover, the comparison was extended to the protein profile of pathological subjects affected by AD, in order to confirm

the results obtained in the previous study[146] with a different statistical approach, to validate the goodness of the classification method and to determine their feasibility for the health/disease status-related classification of the subjects.

Table 2.1: Significant changes ($p < 0.05$) in human saliva composition reported in elderly (subjects over 60 years old) with respect to young population.

Protein	Increase ↑ or decrease ↓	Reference
Histatins	↓	137
MUC1, MUC2	↓	138,139
Lactoferrin, transferrin	↓	140
Peroxidase activity	↓	141
Lysozyme	↑	142
Amylase	↑	142
Albumin	↑	142
IgA	↑	140,142
Cystatin A and B, small proline-rich protein 3	↑	143

The research activity of the present section was performed at the University of Cagliari (Cagliari, Italy) under the supervision of Prof. Tiziana Cabras and in collaboration with Prof. Giacomo Diaz.

MATERIALS AND METHODS

Study subjects

In this study we analysed, by a different statistical approach, quantitative proteomic data obtained on salivary samples in our two previous studies[79,146]. We used proteomic data of thirty-five adult healthy controls (aHC) (18 females, 17 males, 46 ± 12, mean age ± SD) selected among those enrolled in the study of Serrao and colleagues[79] in order to have a group matched for sex and number with the group of elderly healthy controls (eHC), (18 females and 16 males, 78 ± 5 mean age ± SD). For the eHC group as for the AD group (23 females and 12 males; 80 ± 6 mean age ± SD), we utilized proteomic data of eHC and AD subjects included in the study of Contini and colleagues[146]. Table 2.2 reports demographic features and TPC values determined by BCA assay of all the subjects involved. The informed consent process for sample collections agreed with the latest stipulations established by the Declaration of Helsinki. The study approval was obtained by the formal ethical committees of the Catholic University of Rome and of the University of Cagliari, being the samples from healthy controls collected in Cagliari and the samples from AD subjects at the Neurology Department of the "Fondazione Policlinico Universitario A. Gemelli", Catholic University of Rome. The diagnosis of AD has been made according to standardized criteria[1]. None of the subjects included were affected by any major oral disease (periodontitis, caries, or dry mouth), moreover, they had not history of radiotherapy or chemotherapy and were carefully selected as no-smokers. The elderly subjects enrolled as controls suffered from common age-related illness, such as hypertension, and were treated with standard drugs. However, none of them used antidepressants or anticholinergic drugs.

Sample treatment and HPLC-ESI-MS and MS/MS analysis

The procedure of sample collection and treatment, and the experimental conditions of the HPLC-ESI-IT-MS and MS/MS analysis have been described in our previous studies[79,146]. All the samples were collected and analysed in the same period of time and, thus, with the same chromatography conditions and MS parameters. The experimental procedure was standardized for the analysis of the acid soluble protein fraction of human saliva, obtained after treatment of the whole saliva with 0.2% trifluoracetic acid in a 1:1 (v/v) ratio. All the peptides/proteins included in this investigation were previously identified in the acid soluble fraction of saliva by HPLC-HR-ESI-MS/MS analysis (LTQ-Orbitrap Elite or LTQ-Orbitrap XL)[79,146–148]. LR-MS analyses were utilized for the LFQ of identified peptides and proteins with XIC approach. In the present study we considered the 61 peptides and proteins listed in the supplementary material as Table S2.1, which reports UniProt-KB codes, elution times, experimental and theoretical Mav, multiply-charged ions used for the XIC, and the current PTMs.

Quantification of proteins/peptides

Quantification of peptides/proteins was performed using the XIC peak areas measured by LR-MS analysis[79,146]. The quantification of some proteoforms, not included or included in only one of the two previous studies was performed in this study with the following peak parameters: baseline window 15, area noise factor 50, peak noise factor 50, peak height 15%, and tailing factor 1.5. The estimated percentage error of the XIC analysis 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 50 μM used as internal standard. Finally, XIC peak area values of each peptide/protein were normalized on the total protein concentration (TPC) measured as described in Contini *et al.*[146] and expressed in μg/μl.

Statistical analysis

Statistical analysis considered both single proteoforms and the sum of proteoforms of the same protein and thus we refer to them as components. In total, the number of components examined for the purposes of this study is 76. The difference of TPC within the three groups was tested by Kruskal-Wallis test for non-parametric multiple comparison followed by Dunn's Post Test using GraphPad Prism software (version 5.0). XIC areas of all proteins/peptides showed a considerable deviation from normality using Kolmogorov-Smirnov and a number of goodness-of-fit tests (Shapiro-Wilk, Anderson-Darling, Lilliefors, with p-values < 0.0001 in almost all tests, data not shown). This suggested the adoption of non-parametric exact Mann-Whitney tests (between groups) and Kendall correlations (within groups). Significant p-values of simultaneous multiple tests were verified by the Benjamini-Hochberg procedure[149] to keep a cumulative false discovery ratio (FDR) among all test less than 5%. MDS was applied to Kendall correlations to obtain a dimensionally reduced diagram of co-expressed proteins. The classification of subjects was obtained using RF analysis. Algorithm parameters, such as the number of trees to grow and the number of features randomly sampled for each split, were preliminarily tuned to minimize the classification error. RF was applied to three data set combinations: (1) aHC and eHC, characterized by differences in age; (2) eHC and AD, characterized by differences in normal or pathological conditions; (3) aHC, eHC and AD, characterized by differences both in age and normal or pathological conditions. Classification accuracy was calculated as the proportion of correct assessments (both true positive and true negative) to the total number of assessments. The use of a subset proteins/peptides selected according to the Boruta method[150], compared to the use of all proteins/peptides analyzed, resulted in a consistent increase of classification accuracy. The relative importance of each protein/peptide for classification was expressed by the mean decrease of the Gini index (MDG). The Gini index is a measure of impurity. For a single

decision tree, the Gini index ranges from 0 (no impurities, 100% correct classification) to 1 (total impurity, elements are randomly distributed across classes). MDG averages the decrease in impurity for each tree of the whole 'forest', produced by each protein. Dimensionally reduced diagrams of RF classifications were obtained by MDS and HCA using the RF proximity values (the normalized frequency of trees that contain the two samples in the same end node). For HCA we used the Ward's agglomerative method and 1-proximity as distance between each pair of samples. MDS was computed using the singular value decomposition method which ensures a matrix factorization numerically accurate even in the presence of a high degree of multicollinearity (i.e., multiple correlation). Analyses were made using R (RCoreTeam. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. http://www.R-project.org/).

Table 2.2. Demographic data of adult healthy controls (aHC), elderly healthy controls (eHC) and AD patients involved in the study. For each subject it is indicated the TPC of the acid soluble fraction of saliva determined by BCA assay.

aHC	Sex and age	TPC μg/μl	eHC	Sex and age	TPC μg/μl	AD	Sex and age	TPC μg/μl
#1	F, 44	0.27	#1	M, 70	1.31	#1	M, 82	1.86
#2	M, 53	0.37	#2	M, 85	1.44	#2	F, 80	0.18
#3	F, 60	1.6	#3	F, 84	1.01	#3	M, 85	1.14
#4	F, 43	0.61	#4	M, 82	1.23	#5	F, 63	1.12
#5	F, 38	0.63	#5	F, 81	1.58	#6	M, 78	1.85
#6	F, 39	1.21	#6	F, 81	0.81	#7	M, 85	2.32
#7	F, 23	0.67	#7	F, 79	1.77	#8	F, 81	0.28
#8	F, 49	0.25	#8	M, 74	1.05	#9	F, 78	0.56
#9	M, 55	0.6	#9	M, 71	1.37	#10	M, 85	0.58
#10	M, 54	0.61	#10	M, 78	1.12	#11	F, 80	0.78
#11	M, 36	0.62	#11	M, 76	1.38	#13	F, 79	1.14
#12	M, 24	0.37	#12	F, 77	1.83	#14	F, 82	2.05
#13	M, 27	1.03	#13	M, 74	0.89	#16	F, 83	0.83
#14	M, 53	0.94	#14	M, 87	1.38	#17	F, 63	0.44
#15	M, 53	1.1	#15	M, 73	0.64	#18	F, 80	0.76
#16	M, 58	0.52	#16	F, 81	1.08	#19	F, 80	0.87
#17	F, 43	1.14	#17	F, 82	1.11	#20	M, 87	0.3
#18	M, 45	1.01	#18	F, 72	1.05	#21	M, 81	0.47
#19	F, 64	0.96	#19	F, 86	1.12	#22	M, 87	0.48
#20	M, 38	0.51	#20	F, 73	2.12	#23	F, 75	0.53
#21	F, 52	1.36	#22	F, 78	0.79	#24	F, 75	0.16
#22	M, 36	0.71	#23	F, 79	1.24	#25	F, 83	0.5
#23	F, 57	0.57	#24	F, 78	0.65	#26	F, 84	0.79
#24	F, 60	0.68	#25	F, 75	1.64	#27	F, 81	0.87
#25	F, 59	1.56	#26	M, 75	0.9	#28	F, 84	0.31
#26	F, 40	0.26	#27	F, 89	1.15	#29	F, 92	0.36
#27	F, 27	0.26	#28	M, 78	0.63	#30	M, 86	0.72
#28	M, 33	0.86	#29	M, 73	0.66	#31	M, 77	0.38
#29	F, 54	1.02	#30	F, 76	0.69	#32	F, 88	0.33
#30	F, 67	0.75	#31	F, 81	0.46	#33	F, 81	0.31
#31	M, 46	0.27	#32	M, 81	0.41	#34	F, 77	0.86
#32	M, 55	0.6	#33	M, 84	0.21	#35	M, 87	1.31
#33	M, 56	0.37	#34	F, 72	0.31	#36	M, 84	0.58
#34	F, 30	0.44	#35	M, 80	0.72	#37	F, 77	1.4
#35	M, 62	0.69				#38	F, 78	0.8

RESULTS

The results obtained in the present study concerned the fraction of salivary peptides and proteins soluble in acidic solution and directly analyzable by RP-HPLC-ESI-MS using a top-down approach. The main salivary protein families were investigated: aPRPs, statherin, Hst-1, 3, 5 and 6, P-B peptide, cystatins A, B, C, D and salivary (S-type), α-defensins 1-4, Tβ4, SLPI, S100A proteins (Supplementary Table S2.1). Overall, 61 proteins/peptides were included in the study, among which modified proteoforms generated by phosphorylation, proteolysis, N-terminal acetylation, methionine or tryptophan oxidation, and cysteine oxidation (formation of disulfide bridges, glutathionylation, cysteinylation, and nitrosylation) (Table S2.1). For the purpose of this study, we also included cystatin S2 mono-oxidized, S100A12, S100A9 long glutathionylated ((L)-SSG) and its phosphorylated and oxidized proteoforms (Table S2.1), which were not originally included in our previous studies[79,146]. The XIC area values measured for each peptide/protein in every sample was normalized on TPC, indeed, the TPC resulted significantly increased in eHC both with respect to aHC (p-value < 0.01) and with respect to AD patients (p-value < 0.05). (Figure 2.1). The XIC search was performed along the TIC chromatographic profiles. Typical TICs obtained by RP-HPLC-ESI-MS analysis from one aHC, one eHC and one AD subject involved in this study are shown in Figure 2.2 (panels A, B and C respectively). XIC area values of the investigated proteoforms of the three groups were analyzed using different statistical methods: (1) multiple Mann-Whitney tests to identify differentially expressed proteins/peptides included between groups, (2) multiple Kendall correlations to identify co-expressed proteins/peptides included within groups, and (3) RF analysis to provide a classification of single subjects into different groups.

Fig.2.1: TPC distribution of acid soluble protein fraction of saliva from aHC, eHC and AD patients. On the bottom it is reported the mean ± SD for each group, also shown by horizontal bars.

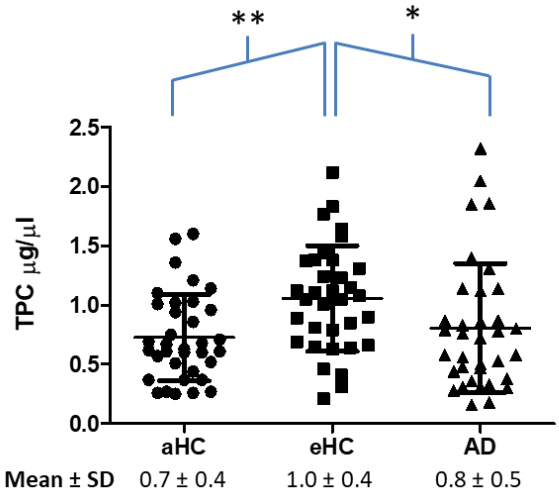
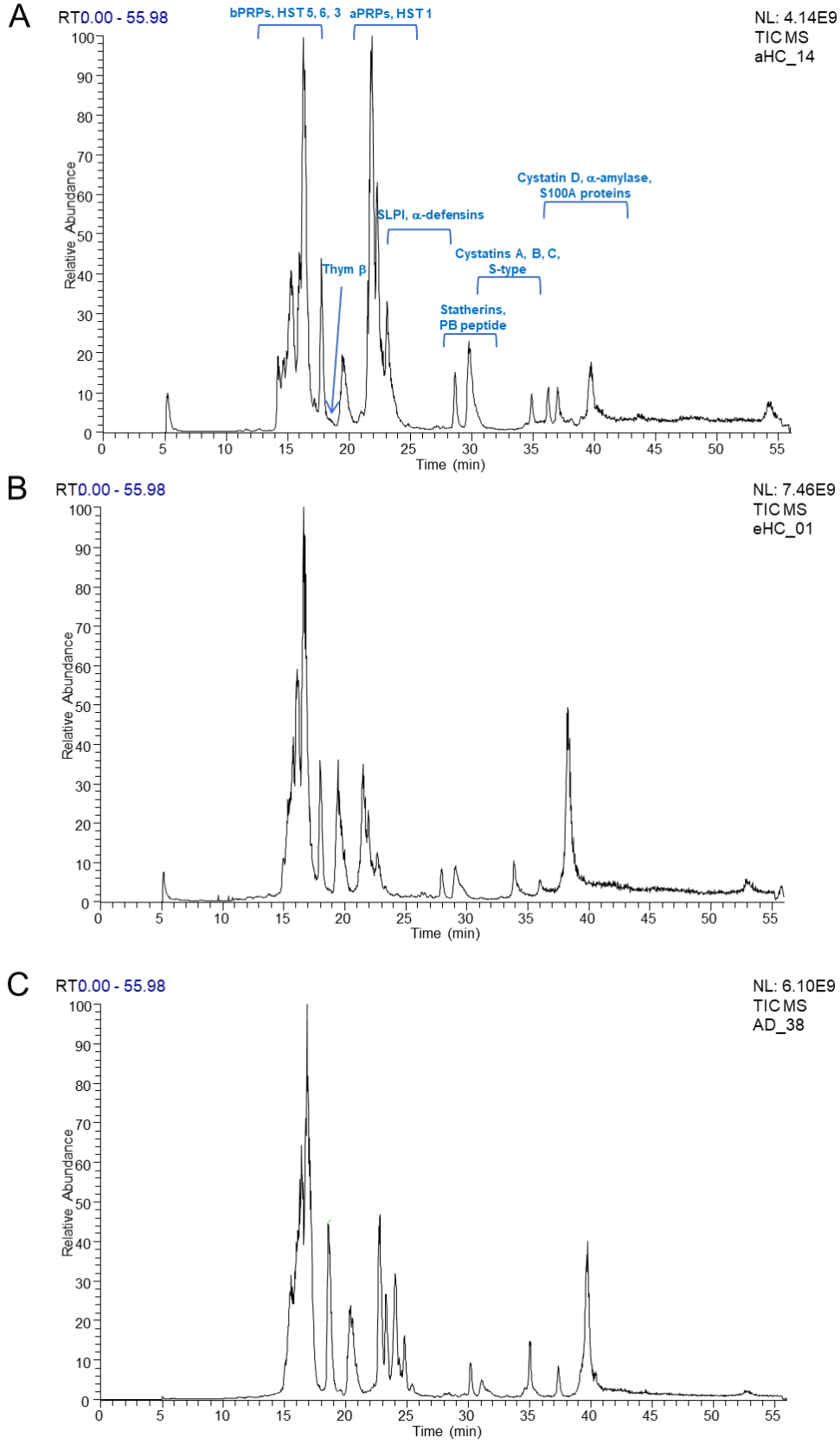


Fig.2.2: TIC profile of acidic-soluble fraction of saliva from aHC (panel A), eHC (panel B) and AD patient (panel C) obtained by RP-HPLC-ESI-LR-MS analysis.



1. Differentially expressed proteins/peptides

The XIC peak areas, the frequencies, and Mann-Whitney comparisons of all the components measured in aHC, eHC and AD patients, are shown in Table 2.3. In the case of aPRPs, statherin, P-B peptide, Histatins, cystatin A, B, S1, S2, SN, S100A8, S100A9 and α-defensins the sum of the XIC peak areas of all their proteoforms (Table S2.1) was also considered and reported in Table 2.3, leading to a total of 76 components to be submitted to statistical analysis. Results of Mann-Whitney tests are also graphically shown in supplementary Figure S2.1. The comparison between aHC and eHC (Table 2.3 and Figure S2.1, panel A) showed significant changes in 69 of the 76 components tested (91%). All these changes, with no exception, consisted in a decrease of protein levels in eHC. The highest significant down-regulation concerned the P-B peptide and particularly its fragment des1-7, the sum of PRP-3 and PRP-1 proteoforms and the oxidized proteoforms of S-type cystatins (indicated as "ox"). On the other hand, components which did not vary significantly between the two groups of controls were S100A7, the form of PRP3 missing for Arginine in position 106 ($desR_{106}$), cystatin B cysteinylated (SSC), the non-phosphorylated (0P) proteoforms of PRP-1 and Hst-1, Hst-3, and SLPI. The comparison between eHC and AD (Table 2.3 and Figure S2.1, panel C) showed significant changes in 37 of the 76 components tested (49%), all showing higher protein levels in AD. These were statherin and its proteoforms des1-9, des1-13 and, as consequence, the sum of all proteoforms; Hst-1, both phosphorylated and non-phosphorylated; P-C peptide, cystatin A, the several proteoforms of S100A8, S100A9, and cystatin B, cystatin SA, the oxidized form of cystatin S2 and SN, α defensins and Tβ4. The comparison between AD patients and aHC (Table 2.3 and Figure S2.1, panel B) showed significant lower abundance of cystatin A, S-type cystatins, Hst-5 and Hst-6, PRP-1, PRP-3, with the exception of its non-phosphorylated form, P-C peptide, statherin and its truncated

forms des$T_{42}F_{43}$ and desD_1, and P-B peptide. Only the sum of the oxidized forms of S100A8 showed an opposite trend, being increased in AD group with respect to aHC.

Table 2.3. XIC peak areas (median and interquartile range) normalized on TPC, and frequencies (F) of the salivary proteins/peptides in aHC, AD patients and eHC. On the right, results of multiple Mann-Whitney tests for group comparisons, with an FDR <5%. The range of color tones from yellow to red denotes the magnitude of significant p-values. Color tones are continuous and more accurate over significance thresholds.

	aHC				AD				eHC				aHC vs eHC		aHC vs AD		eHC vs AD	
	XIC Peak Area			F	XIC Peak Area			F	XIC Peak Area			F						
Components	25th perc	median	75th perc		25th perc	median	75th perc		25th perc	median	75th perc		p-value	change	p-value	change	p-value	change
S100A12	1.2E05	1.9E05	3.9E05	8/35	1.0E05	1.7E05	3.1E05	3/35	8.1E04	9.7E04	1.6E05	5/34	<0.001	↓eHC	ns		ns	
S100A8	1.4E05	2.3E05	7.8E07	11/35	1.1E05	2.0E05	1.3E08	9/35	7.4E04	9.4E04	1.5E05	2/34	<0.00001	↓eHC	ns		<0.01	↑AD
S100A8-Hyperox	1.0E05	1.6E05	2.5E05	2/35	1.1E05	1.8E05	3.3E05	5/35	7.7E04	9.5E04	1.5E05	3/34	<0.01	↓eHC	ns		<0.01	↑AD
S100A8-SNO	1.1E05	1.6E05	2.7E05	2/35	1.3E05	2.8E05	6.0E07	12/35	7.4E04	9.4E04	1.4E05	0/34	<0.0001	↓eHC	ns		<0.00001	↑AD
Sum_S100A8-ox	2.2E05	3.3E05	5.4E05	5/35	3.5E05	1.1E06	1.1E08	18/35	1.5E05	1.9E05	3.0E05	3/34	<0.001	↓eHC	<0.01	↑AD	<0.00001	↑AD
Sum_S100A8	4.1E05	6.8E05	9.4E07	14/35	5.8E05	1.1E08	2.208	24/35	2.3E05	2.9E05	4.6E05	5/34	<0.0001	↓eHC	ns		<0.00001	↑AD
S100A7	1.4E05	2.0E05	1.2E07	10/35	1.1E05	1.9E05	9.2E06	9/35	9.0E04	1.5E05	3.5E07	11/34	ns		ns		ns	
S100A9S	1.7E05	4.0E05	4.4E08	17/35	2.0E05	9.3E07	2.2E08	20/35	8.8E04	1.5E05	5.9E07	15/34	<0.01	↓eHC	ns		<0.001	↑AD
S100A9S-1P	1.4E05	1.9E05	5.2E07	9/35	1.2E05	3.0E05	9.3E07	13/35	7.7E04	9.7E04	1.6E05	4/34	<0.001	↓eHC	ns		<0.01	↑AD
S100A9S-ox	1.7E05	4.7E07	2.1E08	19/35	1.2E05	3.0E05	1.1E08	14/35	8.1E04	1.1E05	1.7E07	10/34	<0.0001	↓eHC	ns		<0.05	↑AD
S100A9S-1Pox	1.0E05	1.7E05	3.2E05	5/35	1.1E05	2.0E05	3.5E05	7/35	7.7E04	9.5E04	1.5E05	1/34	<0.001	↓eHC	ns		<0.01	↑AD
Sum_S100A9S_and_S-1P	3.3E05	8.0E05	5.6E08	18/35	4.4E05	1.0E08	3.2E08	23/35	1.8E05	3.1E05	7.1E07	15/34	<0.01	↓eHC	ns		<0.001	↑AD
Sum_S100A9S-1P_and_S-1Pox	2.9E05	5.4E05	1.5E08	14/35	3.9E05	4.8E07	1.0E08	20/35	1.5E05	1.9E05	3.2E05	4/34	<0.0001	↓eHC	ns		<0.001	↑AD
Sum_S100A9S-ox_and_S-1Pox	3.3E05	7.1E07	2.6E08	19/35	2.5E05	6.1E05	1.5E08	15/35	1.6E05	2.1E05	2.9E07	10/34	<0.0001	↓eHC	ns		<0.05	↑AD
Sum_S100A9S	7.1E07	3.0E08	7.6E08	32/35	1.1E08	2.0E08	3.5E08	33/35	3.5E05	6.2E05	9.6E07	16/34	<0.00001	↓eHC	ns		<0.00001	↑AD
Sum_S100A9L-SSG	5.8E05	1.5E06	2.4E08	20/35	9.0E05	4.9E07	1.8E08	26/35	3.5E05	6.2E05	5.0E07	16/34	<0.05	↓eHC	ns		<0.01	↑AD
Cystatin_A	1.7E08	3.3E08	5.2E08	33/35	8.8E07	1.5E08	2.6E08	32/35	4.0E07	6.5E07	1.6E08	29/35	<0.00001	↓eHC	<0.001	↓AD	<0.01	↑AD
Cyst_A-NAcetyl	2.4E07	5.9E07	1.1E08	29/35	1.8E06	2.2E07	3.6E07	26/35	1.7E05	1.1E07	2.4E07	23/35	<0.0001	↓eHC	<0.001	↓AD	ns	
Sum_cyst_A	2.1E08	3.9E08	5.6E08	34/35	9.4E07	1.7E08	3.0E08	32/35	4.7E07	7.9E07	1.9E08	30/34	<0.00001	↓eHC	<0.001	↓AD	<0.05	↑AD
Cystatin_B-SSG	4.3E07	7.2E07	1.3E08	29/35	2.1E07	4.1E07	1.2E08	30/35	1.1E07	1.9E07	3.1E07	27/34	<0.0001	↓eHC	ns		<0.01	↑AD

Components	aHC				AD				eHC				aHC vs eHC		aHC vs AD		eHC vs AD	
	XIC Peak Area			F	XIC Peak Area			F	XIC Peak Area			F	p-value	change	p-value	change	p-value	change
	25th perc	median	75th perc		25th perc	median	75th perc		25th perc	median	75th perc							
Cyst_B-SSC	1.5E05	1.5E07	4.1E07	18/35	2.6E05	1.5E07	2.8E07	21/35	1.1E05	4.0E06	1.3E07	18/34	ns		ns		<0.05	↑AD
Cyst_B-S-S_dimer	8.9E06	5.5E07	1.2E08	26/35	1.7E07	5.4E07	7.9E07	31/35	1.5E05	1.7E07	3.3E07	22/34	<0.01	↓eHC	ns		<0.001	↑AD
Sum_cyst_B	7.5E07	1.5E08	3.0E08	30/35	4.5E07	1.1E08	2.4E08	33/35	1.6E07	4.5E07	7.3E07	27/34	<0.0001	↓eHC	ns		<0.01	↑AD
Cystatin_C	1.6E05	3.8E05	6.8E07	16/35	1.2E05	1.8E05	3.2E05	6/35	8.1E04	9.7E04	2.4E05	6/34	<0.00001	↓eHC	<0.01	↓AD	ns	
Cystatin_D_R_{26}des1-5	1.7E05	4.0E07	1.7E08	20/35	1.3E05	2.8E05	5.3E07	13/35	9.0E04	3.6E05	5.8E07	16/34	<0.05	↓eHC	ns		ns	
Cystatin_S	1.6E05	4.0E05	1.4E08	17/35	1.2E05	1.9E05	3.2E05	5/35	8.1E04	1.2E05	2.1E07	9/34	<0.001	↓eHC	<0.01	↓AD	ns	
Cyst_S1	2.1E08	6.8E08	1.3E09	33/35	1.1E08	3.2E08	8.1E08	29/35	1.0E08	3.3E08	6.6E08	30/34	<0.05	↓eHC	ns		ns	
Cyst_S2	3.6E07	2.3E08	4.0E08	31/35	2.2E05	8.0E07	2.6E08	20/35	3.4E07	8.4E07	2.1E08	26/34	<0.01	↓eHC	<0.01	↓AD	ns	
Cyst_SN	5.7E08	1.6E09	2.4E09	34/35	1.6E08	4.8E08	1.6E09	30/35	2.6E08	5.4E08	8.8E08	29/34	<0.001	↓eHC	<0.05	↓AD	ns	
Cyst_SA	2.5E05	2.0E08	4.3E08	24/35	1.2E05	2.1E05	5.9E05	8/35	7.7E04	9.7E04	2.4E05	7/34	<0.00001	↓eHC	<0.01	↓AD	ns	
Cyst_S1-ox	3.3E05	1.9E08	4.9E08	24/35	1.1E05	2.0E05	3.5E07	10/35	7.7E04	1.1E05	2.0E05	7/34	<0.00001	↓eHC	<0.0001	↓AD	ns	
Cyst_S2-ox	1.9E05	2.3E07	1.4E08	18/35	1.0E05	1.7E05	2.9E05	3/35	7.4E04	9.4E04	1.5E05	1/34	<0.00001	↓eHC	<0.0001	↓AD	<0.05	↑AD
Cyst_SN-ox	8.0E05	2.3E08	5.7E08	25/35	1.2E05	2.6E05	5.2E07	12/35	7.7E04	1.1E05	1.4E07	9/34	<0.00001	↓eHC	<0.0001	↓AD	<0.05	↑AD
Sum_cyst_S1	4.4E08	8.0E08	1.8E09	33/35	1.5E08	3.4E08	8.1E08	31/35	1.2E08	3.3E08	6.6E08	30/34	<0.001	↓eHC	<0.01	↓AD	ns	
Sum_cyst_S2	1.4E08	2.5E08	6.8E08	32/35	4.4E05	9.6E07	2.6E08	21/35	3.5E07	8.4E07	2.1E08	26/34	<0.0001	↓eHC	<0.001	↓AD	ns	
Sum_cyst_SN	8.1E08	1.9E09	2.8E09	34/35	1.7E08	4.8E08	1.6E09	32/35	2.6E08	5.4E08	8.8E08	31/34	<0.00001	↓eHC	<0.01	↓AD	ns	
Hst-1-1P	3.7E05	2.4E08	5.7E08	24/35	3.2E07	1.3E08	2.3E08	33/35	1.5E07	6.5E07	1.1E08	29/34	<0.01	↓eHC	ns		<0.05	↑AD
Hst-1-0P	1.4E05	2.7E05	6.4E07	14/35	2.1E05	2.7E07	4.2E07	24/35	1.1E05	3.4E06	2.1E07	17/34	ns		ns		<0.05	↑AD
Sum_Hst-1	7.5E05	2.7E08	6.8E08	25/35	5.1E07	1.4E08	2.9E08	33/35	1.6E07	6.9E07	1.4E08	28/34	<0.01	↓eHC	ns		<0.05	↑AD
Hst-3_1-24_Hst-5	1.9E08	4.4E08	1.1E09	30/35	1.8E05	9.2E07	1.5E08	20/35	9.6E04	7.9E06	9.1E07	17/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
Hst-3_1-25_Hst-6	1.7E05	3.5E07	2.1E08	18/35	1.2E05	2.0E05	2.3E06	8/35	9.1E04	1.5E05	3.9E07	10/34	<0.001	↓eHC	<0.01	↓AD	ns	
Hst-3	1.5E05	3.8E05	2.1E08	17/35	1.5E05	6.3E05	5.8E07	17/35	1.1E05	1.6E07	5.6E07	18/34	ns		ns		ns	
Sum_Hst-3	2.8E08	5.9E08	1.5E09	33/35	6.1E05	1.4E08	2.7E08	25/35	4.0E05	3.8E07	2.1E08	22/34	<0.00001	↓eHC	<0.00001	↓AD	ns	

	aHC				AD				eHC				aHC vs eHC		aHC vs AD		eHC vs AD	
	XIC Peak Area			F	XIC Peak Area			F	XIC Peak Area			F						
Components	25th perc	median	75th perc		25th perc	median	75th perc		25th perc	median	75th perc		p-value	change	p-value	change	p-value	change
Tβ4	1.6E05	6.6E07	1.7E08	22/35	3.3E05	3.5E07	8.0E07	25/35	9.0E04	4.8E06	5.0E07	17/34	<0.001	↓eHC	ns		<0.01	↑AD
α-defensin1	5.1E07	1.8E08	3.2E08	32/35	6.2E07	1.1E08	2.3E08	32/35	2.7E06	3.4E07	7.5E07	25/34	<0.0001	↓eHC	ns		<0.0001	↑AD
α-defensin2	6.4E07	1.3E08	2.3E08	30/35	3.8E07	1.1E08	1.6E08	30/35	1.3E07	2.6E07	5.5E07	27/34	<0.0001	↓eHC	ns		<0.001	↑AD
α-defensin3	2.5E05	4.1E07	1.3E08	24/35	4.6E05	5.4E07	1.0E08	25/35	1.3E05	2.9E06	1.9E07	14/34	<0.001	↓eHC	ns		<0.0001	↑AD
α-defensin4	1.6E05	3.1E07	6.1E07	20/35	2.1E05	1.3E07	3.4E07	18/35	8.3E04	1.2E05	8.4E06	11/34	<0.0001	↓eHC	ns		<0.001	↑AD
Sum_ α -defensins	2.0E08	4.8E08	8.5E08	34/35	1.2E08	2.9E08	4.7E08	34/35	2.4E07	6.5E07	1.8E08	27/34	<0.00001	↓eHC	ns		<0.0001	↑AD
PRP-1-2P	6.7E09	1.0E10	2.0E10	35/35	1.4E09	3.2E09	5.6E09	34/35	1.1E09	2.5E09	4.6E09	33/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
PRP-1-1P	6.0E08	1.3E09	2.1E09	34/35	1.6E08	3.6E08	7.5E08	34/35	8.0E07	2.9E08	5.5E08	31/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
PRP-1-0P	1.4E05	1.1E07	1.1E08	17/35	1.2E05	1.9E05	4.4E05	7/35	9.4E04	7.4E06	3.8E07	17/34	ns		<0.05	↓AD	ns	
PRP-1-3P	6.4E07	2.2E08	3.4E08	33/35	1.2E05	2.1E05	5.9E05	8/35	1.2E05	4.0E05	4.0E07	15/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
Sum_PRP-1	8.2E09	1.2E10	2.3E10	35/35	1.8E09	3.7E09	6.1E09	35/35	1.3E09	2.9E09	5.2E09	33/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
PRP-3-2P	1.8E09	3.4E09	7.6E09	35/35	5.2E08	9.2E08	2.0E09	35/35	2.9E08	8.0E08	1.3E09	33/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
PRP-3-1P	2.5E08	5.2E08	8.9E08	35/35	1.1E08	1.5E08	2.9E08	35/35	3.7E07	1.5E08	1.9E08	31/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
PRP-3-0P	1.1E05	1.9E05	3.0E07	12/35	1.3E05	2.1E05	9.9E06	10/35	8.8E04	1.3E05	2.3E06	9/34	<0.05	↓eHC	ns		ns	
PRP-3-2P_desR_{106}	1.2E08	2.6E08	8.4E08	30/35	6.6E07	1.4E08	2.5E08	31/35	2.6E07	2.0E08	3.2E08	27/34	ns		<0.01	↓AD	ns	
Sum_PRP-3	2.5E09	4.7E09	9.6E09	35/35	8.2E08	1.4E09	2.4E09	35/35	5.8E08	1.1E09	1.8E09	33/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
P-C_peptide	1.1E09	2.3E09	4.3E09	35/35	4.7E08	8.4E08	1.3E09	34/35	2.0E08	4.9E08	8.2E08	34/35	<0.00001	↓eHC	<0.0001	↓AD	<0.01	↑AD
Statherin-2P	9.7E08	1.6E09	2.7E09	35/35	3.9E08	6.5E08	1.1E09	35/35	2.5E08	4.5E08	6.0E08	33/34	<0.00001	↓eHC	<0.001	↓AD	<0.05	↑AD
Stath-1P	3.3E05	3.0E07	4.5E07	25/35	4.4E06	1.3E07	3.2E07	27/35	2.8E05	8.5E06	1.5E07	25/35	<0.01	↓eHC	ns		ns	
Stath_des-F_{43}	9.6E07	1.9E08	4.2E08	34/35	6.5E07	1.1E08	2.1E08	34/35	3.6E07	7.9E07	1.7E08	33/34	<0.01	↓eHC	ns		ns	
Stath_des$T_{42}F_{43}$	2.7E07	5.5E07	1.1E08	34/35	1.3E07	2.5E07	5.3E07	33/35	9.5E06	1.8E07	4.0E07	30/34	<0.001	↓eHC	<0.01	↓AD	ns	
Stath_desD_1	4.8E07	9.7E07	1.5E08	35/35	2.9E05	1.9E07	6.1E07	24/35	9.7E06	2.9E07	7.0E07	32/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
Stath_des1-9	3.4E07	7.8E07	1.5E08	30/35	4.6E06	5.8E07	8.7E07	26/35	2.8E05	1.8E07	3.4E07	24/34	<0.0001	↓eHC	ns		<0.01	↑AD

	aHC				AD				eHC				aHC vs eHC		aHC vs AD		eHC vs AD	
	XIC Peak Area			F	XIC Peak Area			F	XIC Peak Area			F						
Components	25th perc	median	75th perc		25th perc	median	75th perc		25th perc	median	75th perc		p-value	change	p-value	change	p-value	change
Stath_des1-10	2.7E07	5.0E07	8.8E07	32/35	4.9E05	2.9E07	4.4E07	25/35	7.2E06	1.6E07	3.0E07	27/34	<0.00001	↓eHC	<0.01	↓AD	ns	
Stath_des1-13	1.5E07	3.5E07	5.9E07	33/35	1.3E05	1.8E07	2.7E07	23/35	9.3E04	3.5E05	1.4E07	16/34	<0.00001	↓eHC	<0.001	↓AD	<0.01	↑AD
Sum_statherin	1.5E09	2.6E09	3.7E09	35/35	5.6E08	1.0E09	1.7E09	35/35	4.3E08	7.1E08	9.6E08	34/34	<0.00001	↓eHC	<0.001	↓AD	<0.05	↑AD
P-B_peptide	1.4E09	3.0E09	3.9E09	35/35	3.7E08	5.5E08	9.8E08	35/35	3.0E08	5.0E08	9.8E08	33/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
P-B_des1-5	6.8E07	1.3E08	4.4E08	32/35	2.2E07	4.5E07	9.8E07	34/35	2.8E07	5.8E07	1.6E08	34/34	<0.01	↓eHC	<0.001	↓AD	ns	
P-B_des1-7	3.0E08	4.8E08	7.0E08	34/35	4.9E07	9.1E07	1.4E08	34/35	4.2E07	7.1E07	1.2E08	34/34	<0.00001	↓eHC	<0.00001	↓AD	ns	
P-B_des1-4	5.9E07	1.1E08	2.7E08	30/35	4.8E05	3.8E07	8.6E07	25/35	2.2E07	5.1E07	1.0E08	30/34	<0.01	↓eHC	<0.001	↓AD	ns	
P-B_des1-12	5.0E07	9.4E07	1.7E08	34/35	2.3E07	5.1E07	8.0E07	34/35	2.3E07	5.6E07	8.7E07	31/34	<0.05	↓eHC	<0.01	↓AD	ns	
Sum_P-B_peptide	2.2E09	4.2E09	5.2E09	35/35	5.7E08	9.4E08	1.4E09	35/35	4.8E08	8.3E08	1.2E09	35/35	<0.00001	↓eHC	<0.00001	↓AD	ns	
SLPI	1.1E05	1.6E05	2.7E05	5/35	1.4E05	3.2E05	1.6E07	14/35	8.3E04	1.9E05	8.7E06	14/34	ns		ns		ns	

2. Correlated proteins/peptides within groups

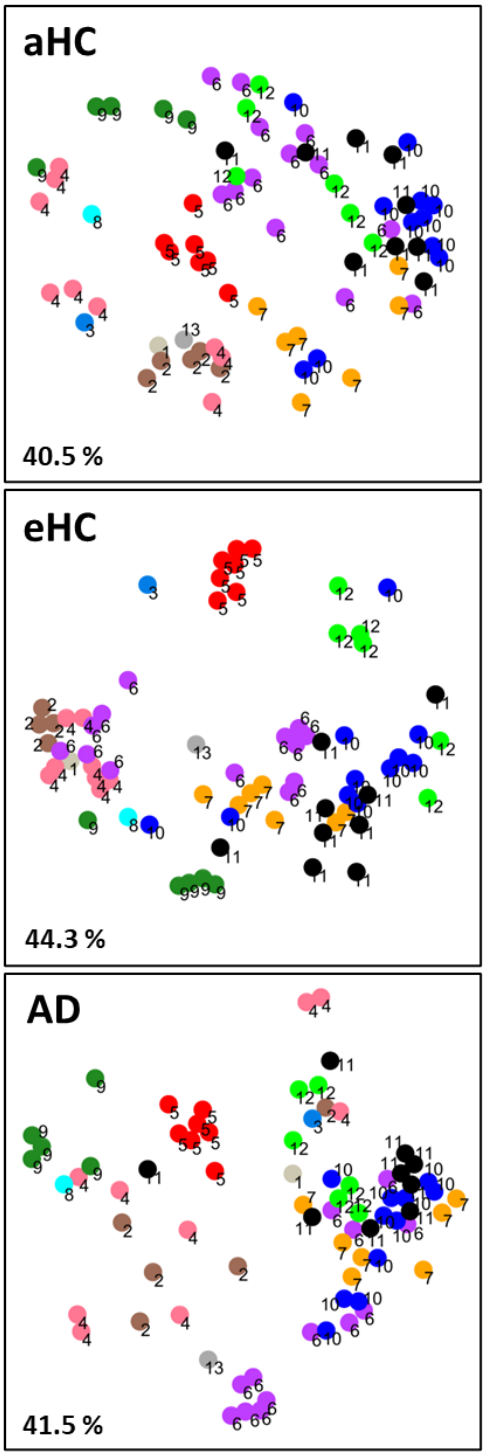
A diagram of correlated proteins/peptides within each group was obtained by MDS applied to Kendall correlations (Figure 2.3). In correlation-MDS diagrams, each component is surrounded by components with correlated expressions, so that clusters represent groups of correlated proteins/peptides. To facilitate the understanding of MDS diagrams, the 76 components were subdivided into 13 categories based on their structural/functional similitudes and secretory origin. The most compact cluster, in all groups, was that of cystatins A and B (category 5 in Figure 2.3). Less compact clusters were formed by histatins (category 7), α-defensins (category 9), aPRPs (category 10), statherin family (category 11) and P-B peptide family (category 12), without appreciable differences between groups. Conversely, differences were found in the degree of clustering of three categories: category 2, that included S100A8 and its proteoforms, and that was relatively more compact in aHC and eHC that in AD; category 4, that included S100A9 and its proteoforms, and that was relatively more compact in eHC that in aHC and AD; category 6, that included cystatins C, D and S, and that was relatively more compact in eHC and AD than in aHC. Other categories (1, 3, 8 and 13) were not compared as these were represented by single proteins/peptides (S100A12, S100A7D27, Tβ4 and SLPI, respectively). In addition to the relationships between components of single categories, a strong proximity between categories 2 and 4, including S100A8 and S100A9 peptides, respectively, was present in eHC.

3. Random forest classification

RF was applied to a subset of components selected by the Boruta method (Figure S2.2). The number of selected components varied in the three analyses. They were 47 for aHC-eHC, 22 for eHC-AD, and 38 for aHC-eHC-AD. Confusion matrices and sensitivity/specificity of classifications are shown in Figure 2.4 (panel A). Classification of aHC-eHC samples showed the highest accuracy (97%), followed by that of aHC-eHC-AD (82.7%) and eHC-AD (79.8%).

It should be noted that these findings were validated by 'out of bag' samples, a method that consists in creating separate sets of training and test samples, composed by 72% and 36% of the entire set of data, respectively. According to MDG scores, the most important components for the classification of aHC-eHC samples were the des1-7 fragment of P-B peptide and aPRPs, in particular PRP1 proteoforms. Differently from these, the most important components for the classification of eHC-AD samples were all S100A8 proteoforms, in particular the oxidized and nitrosylated forms, α-defensins, in particular α-def-3, and all S100A9 proteoforms. The classification of samples mixing the three groups together (aHC-eHC-AD) shared the components already identified in the previous two analyses. The 20 components with the highest MDG score of each analysis are shown in Table 2.4 (the limit of 20 components was chosen because of the different number of components preliminary selected by the Boruta method). MDG scores were in close agreement with the Boruta selection. However, both MDG and Boruta scores were not consistent with Mann-Whitney tests. For example, in the classification of eHC-AD mixed samples, PRP1-2P was the 8^{th} most 'important' component in the MDG ranking and the 10^{th} in the Boruta ranking, but the same protein did not appear to be differentially expressed by the Mann-Whitney test. This apparent contrast represents an essential difference between multivariate RF classification and univariate comparisons and will be discussed in detail in the next section. Diagrams of RF classifications were obtained by MDS, using the proximity between each pair of samples, that is the normalized frequency with which two samples occupy the same terminal node through one tree, averaged for all trees (Figure 2.4, panel B). RF classification of aHC-eHC-AD mixed samples was also shown by HCA (Figure 2.5), using the proximity values as distance and the Ward's agglomerative criterion. The HCA dendrogram shows three main clusters which represent the three groups of subjects with a distribution comparable to that shown by the confusion matrix.

Fig.2.3: MDS diagrams of Kendall correlations between components. To facilitate the understanding of MDS diagrams, the 76 components are grouped into different categories, numerically and color encoded, based on their structural/functional similitudes and secretory origin. The degree of clustering of points accounts for the degree of proteins/peptides correlation. Percent values indicate the percent of information contained in 2D MDS diagrams, relative to all the information contained in the whole multi-dimensional structure.



1 S100A12

2 S100A8
S100A8Hyperox
S100A8_SNO
Sum_S100A8_ox
Sum_S100A8

3 S100A7D27

4 S100A9_s
S100A9_s_p
S100A9_s_ox
S100A9_s_p_ox
Sum_S100A9_s_and_s_p
Sum_S100A9_s_p_and_p_ox
Sum_S100A9_s_ox_and_p_ox
Sum_S100A9_s
Sum_S100A9_l_g

5 Cystatin_A
Cystatin_A_Acetyl
Sum_Cystatin_A
Cystatin_B_s_glut
Cystatin_B_s_cyst
Cystatin_B_S_Sdimer
Sum_Cystatin_B

6 Cystatin_C
Cystatin_D_des_1_5
Cystatin_S
Cystatin_S1
Cystatin_S2
Cystatin_SN
Cystatin_SA
Cystatin_S1_ox
Cystatin_S2_ox
Cystatin_SN_ox
Sum_Cystatin_S1
Sum_Cystatin_S2
Sum_Cystatin_SN

7 Hst_1
Hst_1_0P
Sum_Hst_1
Hst_3_fr24
Hst_3_fr25
Hst_3
Sum_Hst_3

8 Thymosin_B4

9 a_defensin_1
a_defensin_2
a_defensin_3
a_defensin_4
Sum_a_defensins

10 PRP1_2P
PRP1_1P
PRP1_0P
PRP1_3P
Sum_PRP1
PRP3_2P
PRP3_1P
PRP3_0P
PRP_3_diphos_Des_Arg106
Sum_PRP3
P_C_peptide

11 Statherin_2P
Statherin_1P
Statherin_des_Phe43
Statherin_des_Thr42_Phe43
Statherin_des_Asp1
Statherin_des_1_9
Statherin_des_1_10
Statherin_des_1_13
Sum_Statherin

12 PB_peptide
PB_peptide_des_1_5
PB_peptide_des_1_7
PB_peptide_des_1_4
PB_peptide_des_1_12
Sum_PB_peptide

13 SLPI

Fig.2.4: RF applied to the three mixed data sets. In panel A the confusion matrices of RF classifications, validated by out-of-bag samples. Matrix rows represent the actual classes, while columns represent the predicted classes. Marginal columns show the frequency of false negatives, while marginal rows show the frequency of false positives. In panel B the MDS diagrams showing the relationships among subjects, using the proximity values calculated by RF. Each group is delimited by a dispersion ellipse with a confidence of 1.6 standard deviations. Percent values indicate the percent of information contained in 2D MDS diagrams, relative to all the information contained in the whole multi-dimensional structure.

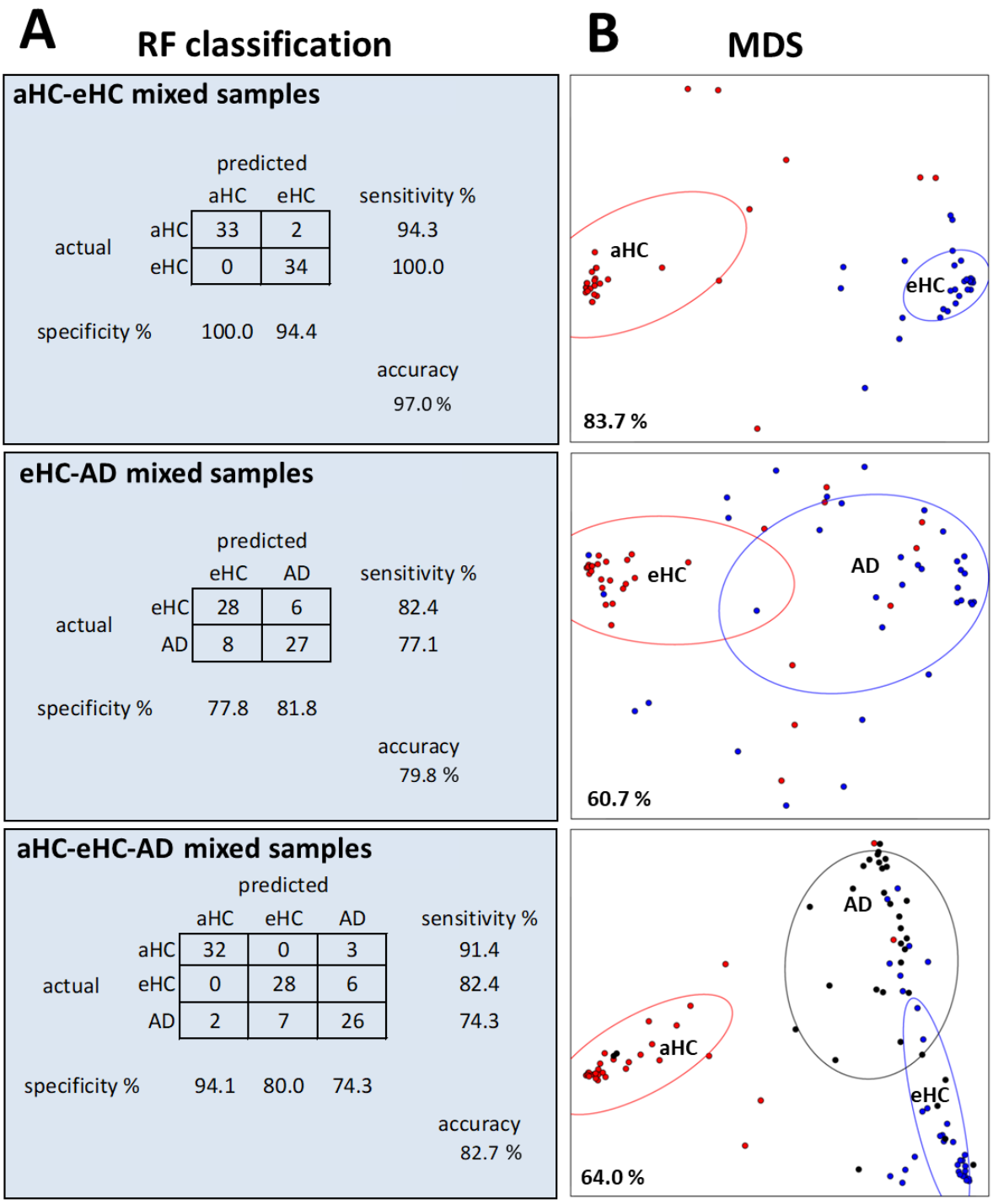
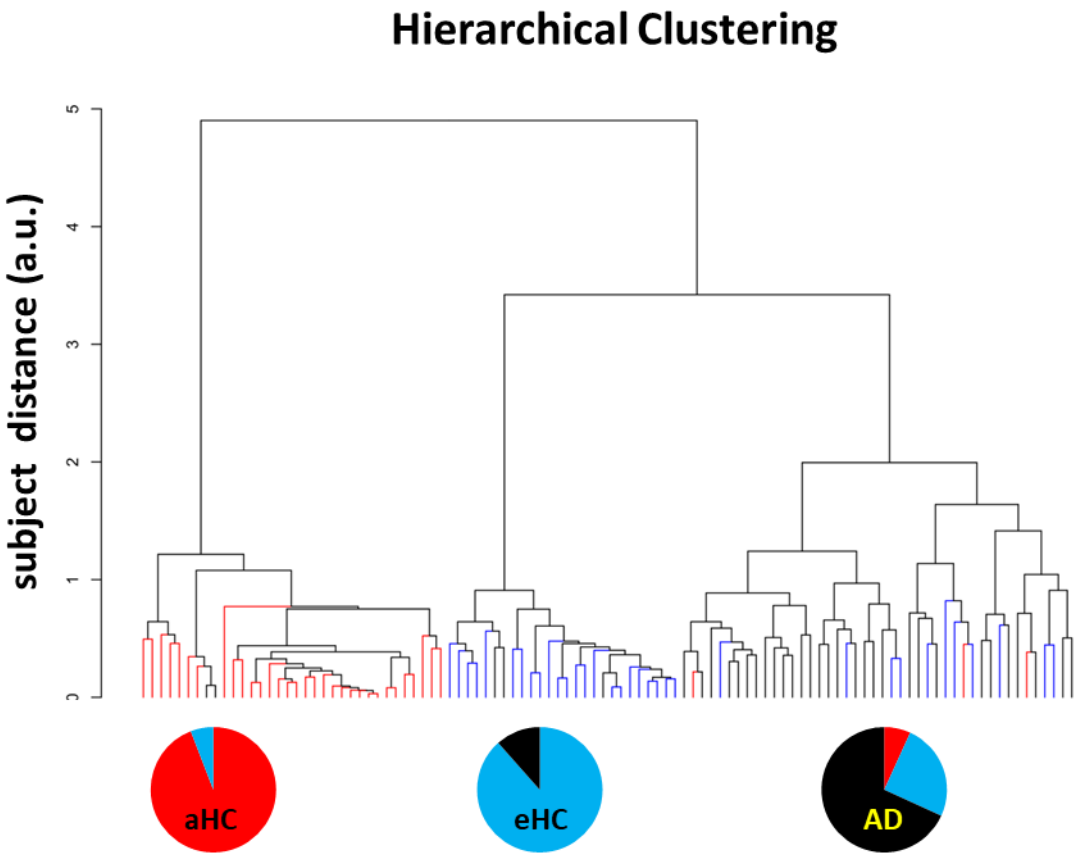


Table 2.4. Mean decrease of the Gini (MDG) scores of the 20 most important proteins/peptides (or their sum, generically indicated as components) identified by RF.

aHC-eHC		**eHC-AD**		**aHC-eHC-AD**	
Component	**MDG**	**Component**	**MDG**	**Component**	**MDG**
P-B_des1-7	11.8	Sum_S100A8	5.92	P-B_des1-7	14.6
PRP-1-2P	4.29	α-defensin3	4.14	Sum_S100A8	5.17
Sum_PRP-1	3.70	Sum_S100A8-ox	3.92	PRP-1-3P	4.93
Hst-3_1-24_(Hst-5)	1.19	Sum_S100A9(S)	2.81	Sum_PRP-1	3.56
Sum_P-B_peptide	1.10	S100A8-SNO	2.01	α-defensin3	3.37
Sum_S100A9(S)	1.00	Sum_S100A9(S)-1P_and_(S)-1Pox	1.80	Sum_S100A8-ox	3.08
P-B_peptide	0.94	Stath_des-F43	1.74	PRP-1-2P	3.03
PRP-3-2P	0.91	PRP-1-2P	1.50	Sum_S100A9(S)	2.94
Cyst_S2-ox	0.87	Stath_des1-13	1.17	S100A8-SNO	1.81
Cyst_SN-ox	0.72	Cystatin_A	1.14	Sum_S100A9(S)-1P_and_(S)-1Pox	1.80
PRP-1-3P	0.51	Stath_des1-9	1.08	Sum_α-defensins	1.64
Sum_cyst_SN	0.50	Sum_α-defensins	1.02	Hst-3_1-24_(Hst-5)	1.48
Sum_α-defensins	0.44	Cyst_B-S-S_dimer	0.99	α-defensin1	1.44
PRP-3-1P	0.43	α-defensin4	0.97	P-B_peptide	1.36
α-defensin4	0.42	α-defensin1	0.86	Sum_P-B_peptide	1.23
Cyst_S1-ox	0.40	Sum_S100A9(S)_and_(S)-1P	0.74	Sum_cyst_SN	1.23
Sum_PRP-3	0.35	α-defensin2	0.68	α-defensin2	1.06
Cystatin_A	0.31	S100A9(S)-1P	0.59	Cyst_B-S-S_dimer	1.06
P-C_peptide	0.30	Tβ4	0.54	α-defensin4	1.05
Sum_Hst-3	0.30	S100A8	0.37	Stath_des1-13	1.02

Fig.2.5: Hierarchical clustering of subjects using RF proximity and the Ward clustering criterion. Pie charts represent the relative frequency of aHC, eHC and AD in each of the three main clusters.



DISCUSSION

The results of this statistical study show different protein profiles in the acid soluble fraction of saliva in eHC, aHC and AD groups. In particular, eHC showed a general down-regulation of peptides and proteins with respect to both aHC subjects and AD patients, and the number and magnitude of significant changes found in the comparison between eHC and aHC was greater than that found in the comparison between eHC and AD. Indeed, almost all the analyzed components showed significant lower levels in elderly subjects than in adult ones. These results are conflicting with respect to the measured TPC, which showed the highest value in eHC group. Probably this incongruency is due to a higher concentration, in elderly subjects, of proteins that influenced the TPC but that have not been considered for the purposes of the present study, as N-glycoproteins[151], amylase and IgA[145]. Indeed, other researchers observed different ageing-related trends for diverse salivary proteins, highlighting the increasing in some and the decreasing in others, particularly mucins, peroxidase, lactoferrin, transferrin and reduced and oxidized glutathion[138,145]. It was suggested that an age-dependence influences the secretion of specific components, whereas, in others, the concentration effect is mainly driven by the decreased salivary output. Moreover, these findings, as well as our findings agreed with age-related histological and physiological changes in the salivary glands, indeed, the volume of fat and fibrovascular tissue increases in the parotid and submandibular glands in elderly individuals[152], as well as, the proportional volume of acinar cell secretion was reduced in elderly individuals[153], being considered as one of the major causes of dry mouth[154]. All these histological changes can result in overall salivary gland hypofunction[153].

When AD salivary protein profile was compared with that one of eHC subjects, we obtained a confirmation of the results obtained in our previous study[146], with an up-regulation of proteins and peptides either specific of the oral cavity or also expressed in other body regions.

Indeed, AD subjects exhibited high abundance of statherin and Hst-1 that are specifically involved in the wellness of the oral cavity[114,119], as well as proteins acting as ROS/RNS scavengers and with a neuroprotective role, such as S100A8, A9 and their glutathionylated and nitrosylated proteoforms[155], cystatin B and glutathionylated and dimeric derivatives[156]; proteins with antimicrobial activity were more abundant in AD group, such as α-defensins 1, 2, 3 and 4[157], cystatins A and B[107,158], Hst-1[116], statherin[119] and Tβ4[104]. Moreover, also the oxidized forms of S-type cystatins S2 and SN, which were not included in our previous study[146], resulted more abundant in AD than in eHC subjects, suggesting a possible role of these proteins as oral ROS scavengers like S100A8, S100A9 and cystatin B, and, in any case, the existence of oxidative state in oral cavity of AD patients.

The comparison between salivary protein profiles of AD and aHC subjects highlighted a down-regulation of a series of proteins and peptides in AD similar to that one observed in the comparison between elderly and adult subjects, they were Hst-5 and 6, P-B and its fragments, some fragments of statherin (desTF, desD1, des1-10), cystatin C, N-acetylated cystatin A, S-type cystatins with the exception of oxidized form of S2 and SN, the several proteoforms of aPRP family except P-C peptide and finally Hst-5 and 6, which agreed with the study of Johnson and colleagues[137]. These results suggested that the lower levels of these proteins/peptides in AD and eHC subjects with respect to aHC ones may be related to physiological aging variations[145,153].

P-C peptide, statherin and its fragment des1-13, oxidized cystatins S2 and SN, and unmodified cystatin A showed a different trend being less abundant in AD patients than aHC subjects, but at the same time more abundant with respect to eHC. Moreover, the total S100A8 oxidized proteoforms resulted up-regulated in AD subjects with respect to both control groups. These differences appeared to be specifically related to the health status, and the highest levels of oxidative forms of S100A8 in AD subjects suggested the existence of an oxidative stress

condition. Indeed, oxidative modifications of proteins are common in neurological disorders since the extracellular milieu is strongly oxidizing as a result of generation of ROS and RNS[91,92]. In this context, S100A8 can play an important role as ROS/RNS scavenger, being particularly sensitive to oxidative cross-linking and massive oxidation and it shows capacity to reduce oxidative damage[97].

Important differences among the salivary protein profiles in elderly, adult controls and AD subjects have been highlighted by the MDS correlation analysis, that showed a correlation of proteins and peptides with related structures, functions, or origin, but not all at the same level and not in all groups. The adult controls and AD patients exhibited a wide unclustered distribution of the analyzed components, even if proteins and peptides secreted from salivary glands (statherin, P-B peptide, aPRPs, histatins, cystatins S-type, C and D)[54] showed to cluster together with respect to those originated from leucocytes, epithelial cells, and plasmatic exudate (α-defensins, Tβ4, S100A proteins, cystatin A and B)[53,83,159]. The proteins/peptides of eHC subjects, instead, showed a higher level of correlation compared to both AD patients and adult controls. Correlation of specific proteins have been observed, cystatins A and B, particularly, were strongly clustered in all the three groups. This result was expected since they are proteins with same 3D structure, 80% sequence homology, and 52% identity[160]. They constitute the type I cystatin subfamily, also called "stefins", they are cathepsin inhibitors, and play an important role as regulation factors of inflammation[109] and in the innate immune response[107]. Cystatin A is an inhibitor of cathepsins B, L and H in epidermis, lymphoid tissue, and oral mucosa[161]. Cystatin B, expressed in a wide range of cells, is the main inhibitor of cathepsin B, and may exert additional functions in maintenance of cell homeostasis, reduction of oxidative stress[112], prevention of apoptosis[113], and neuronal protective role[109]. Interestingly, both cystatins A and B have been reported to co-localize in amyloid plaques of various origins[76,108], and they are considered able to bind Aβ and interrupt amyloid aggregation in

cells[77]. The constant co-expression of cystatin A and B, independently from age, suggested their protective role with respect to amyloid fibrilization, that probably is enhanced during AD occurrence, hypothesis in accordance with the increased abundance of cystatin A and B in the patients revealed by our previous study[146] and confirmed in the present.

In AD patients a correlation between α-defensins 1-4 and Tβ4 was observed, which was less strong in eHC and absent in aHC. They are peptides with antimicrobial and anti-inflammatory activities and immunological effects[104,162] and their correlated up-regulation in saliva of AD patients may reflects the same increase in the brain. Furthermore, this result is in strong accordance with the hypothesis that the microbiota-induced neuronal inflammation may trigger Aβ deposition and AD development[17]. As consequence, the neuropathological alterations might be associated with abnormal expression and/or regulation of antimicrobial and anti-inflammatory peptides such as defensins[163]. Tβ4 is a moonlighting peptide widely expressed in human tissues, where it exerts multiple biological functions[105], among them neuroprotective and neuro-regenerative effects[73,106]. Indeed, Tβ4 was found up-regulated in reactive microglia of patients with AD, where it suppresses the pro-inflammatory signaling[73].

S100A8 proteoforms were strongly clustered in the two control groups and correlated in both with S100A12 protein, and, only in elderly subjects, also with S100A9 proteoforms. The results suggested that the clustering of these S100A proteins may be related to two aspects: i) they are all involved in the modulation of the inflammatory processes[97,164,165]; ii) their participation in the inflammation appeared age-linked, since this correlation was not observed in AD patients, but was significantly downregulated in the elderly subjects.

S100A12 is a potent chemoattractant for monocytic cells[166], and it activates various cell types by binding the receptor for advanced glycosylation end products (RAGE), inducing expression of adhesion molecules and pro-inflammatory cytokines[167]. S100A9 can inhibit leukocyte recruitment, and its oxidation ended the chemo-repulsive effect on peripheral

neutrophils[168], moreover, S100A9 was considered as a molecular switch for oxidative control of inflammation regulated through its own methionine oxidation. As well as S100A9, also S100A8 may play a dual role in inflammation, and their pro-inflammatory activity can switch to anti-inflammatory probably depending on the local microenvironment, the oxidative modifications of the proteins and the binding with metal ions[85]. In particular, the S100A8 exerts anti-inflammatory activity when modified by nitrosylation on its cysteine residue[96]. S100A9 and S100A8 are both ROS/RNS scavenger and play protective roles against oxidative stress in several tissues[155]. Moreover, they can undergo to several oxidative modifications on Cys, Met, and Trp residues, which are all detected in our samples. The strong correlation between all the proteoforms, unmodified and oxidized, of S100A8 and S100A9 in elderly subjects suggested an enhanced request of protection from oxidative damages linked to the aging[164,169].

It is noteworthy that the differences in the correlation of S100A8 and S100A9 proteoforms are also important for the classification of the groups based on MDG scores. Indeed, among the good candidates for the classification of AD subjects with respect to the eHCs, it is important to cite all the analyzed proteoforms of S100A9 and S100A8, in particular the nitrosylated S100A8, cystatins A and B, Tβ4 and the α-defensins, mainly the α-defensin 3.

In general, the outcomes obtained by the differential expression and correlation analyses were in accordance with the RF classification of AD vs eHC subjects and of eHC vs aHC subjects, with some exceptions. It is relevant and in accordance with the Mann-Whitney test that the most discriminant components between elderly and adult subjects were peptides and proteins of salivary gland origin[54], such as the fragment of the P-B peptide, all the proteoforms of PRP1 and Hst-5.

For the classification, RF was preferred over other methods because of several reason, as (a) it is not conditioned by the data distribution; (b) it has a low risk of overfitting; (c) it does not

require supplementary samples for the validation of results, as each tree is built up omitting nearly one-third of the samples that are subsequently used to test the misclassification rate; (d) it provides a measure of the relative importance of each feature in the classification of samples (MDG); (e) it provides an estimation of the proximity (i.e., similarity) between each two samples that opens the possibility of applying MDS and HCA to obtain a visual representation of the classification. The 79.8% accuracy of the classification of eHC-AD mixed samples was relatively higher than that reported for clinical diagnostic methods[170], but lower than that obtained using ultrasensitive assays of biomarkers, namely Aβ and tau proteins, in plasma[171], and in salivary samples[172]. Nevertheless, present data may be of interest in view of the fact that they were obtained from a panel of proteins/peptides present in saliva, not associated with classical AD markers, that were recently related to the disease[146]. In addition, the use of MDS and HCA diagrams provides the possibility to correlate the position of misclassified subjects with their specific clinical profiles (i.e., grading of the disease, therapy, presence of comorbidities, etc.). This will be the subject of future investigations.

Apparently surprising is the fact that some proteins considered important for classification, according to Boruta and MDG scores, did not prove to vary significantly by Mann-Whitney tests (*e.g.,* Statherin desF43 and PRP-1-2P for eHC-AD samples), and *viceversa* proteins that varied significantly by Mann-Whitney tests were not considered important for classification (*e.g.*, Statherin des1-13 and Cystatin C for aHC-eHC samples). This apparent paradox is due to the modus operandi of decision trees, where the same variable can be split several times in a tree, at different threshold levels. As RF is an ensemble of decision trees, this fact has major implications about the meaning of the variable importance. Indeed, because of the use of multiple thresholds, it is possible that variables important for classification have nearly equivalent average values or average ranks in different groups, thus resulting not statistically significant by univariate comparisons. Conversely, variables not useful for classification may

turn out to be significantly different. Because of this behavior, despite the good performance shown by RF classification of aHC, eHC and AD subjects, the machinery of decision trees prevents the consideration of proteins important for classification as candidate markers of AD. For diagnostic purposes, the panel of differentially expressed proteins identified by Mann-Whitney tests appeared to provide more reliable indications.

CONCLUSIONS

The results obtained in the present section, confirmed those of our previous study[146] about the individuation of potential biomarkers candidates of AD in peptides and proteins involved principally in defense mechanisms of the innate immune system, in inflammation regulation and in protection against oxidative stress. Moreover, the present study added new interesting perspectives, firstly demonstrating that salivary protein profile strongly changes its composition, mainly at quantitative point of view, in relation to the aging, and that variation involved almost all the analyzed peptides and proteins and their proteoforms. It is to underline that age-related variation starts since pre-natal age, as demonstrated in previous studies[132–136] and continues with the growth until the old age[137,142,143,145,153,173], suggesting that saliva is a dynamic biofluid that adapt itself to the physiological transformations of the organism. These outcomes suggested the importance of the choice based on the age of the subjects enrolled for proteomic studies, particularly for biomarker disease investigations. From this point of view, the mainly fascinating results were produced by RF analysis, that demonstrated the feasibility of the salivary proteome to discriminate groups of subjects who are different in their age and health status, and thus patients with AD with respect to healthy controls. This outcome enforces the idea to use the salivary protein profile for diagnostic purpose and opens viable ways towards a technical verification of the possible biomarker candidates individuated in this study, particularly S100A8-SNO, S100A9 proteoforms and α-defensin 3.

PART II SUPPLEMENTARY SECTION

Table S2.1: UniProt-KB code, experimental and theoretical average mass values (Mav) ± standard deviations (SD), elution times of proteins and peptides analyzed, m/z values and charge of the multiply-charged ions selected for XIC search in HPLC-low resolution MS and their PTMs.

Proteins/peptides	El. Time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	PTMs
Acid Proline-Rich Proteins				
PRP-1 2P (P02810)	22.2	15515 ± 2 (15514-15515)	1293.9(+12), 1194.4(+13), 1035.3(+15), 970.7(+16), 913.6(+17)	N-Term(Gln->pyro-Glu), S_8(Phospho), S_{22}(Phospho)
PRP-1 1P	22.9	15435 ± 2 (15434-15435)	1287.2(+12), 1188.3(+13), 1030.0(+15), 965.7(+16), 908.9(+17)	N-Term(Gln->pyro-Glu), S_8 or S_{22}(Phospho)
PRP-1 nonphos.	23.2	15355 ± 2 (15354-15355)	1280.5(+12), 1182.1(+13), 1024.6(+15), 960.7(+16), 904.2(+17)	N-Term(Gln->pyro-Glu)
PRP-1 3P	21.6	15595 ± 2 (15594-15595)	1418.7(+11), 1300.5(+12), 1200.6(+13), 1040.6(+15), 975.7(+16)	N-Term(Gln->pyro-Glu) S_8, S_{17}, S_{22} (Phospho)
PRP-3 2P (P02810)	22.8	11161 ± 1 (11161-11162)	1595.5(+7), 1396.2(+8), 1015.7(+11), 931.1(+12), 859.6(+13)	N-Term(Gln->pyro-Glu) S_8, S_{22} (Phospho) Fragment 1-106 of PRP-1
PRP-3 1P	23.4	11081 ± 1 (11081-11082)	1584.1(+7), 1386.2(+8), 1008.4(+11), 924.5(+12), 853.4(+13)	N-Term(Gln->pyro-Glu) S_8 or S_{22} (Phospho)
PRP-3 nonphos.	23.8	11001 ± 1 (11001-11002)	1376.2(+8), 1101.2(+10), 917.8(+12) 786.8(+14)	N-Term(Gln->pyro-Glu)
PRP-3 2P desR_{106}	22.8	11004 ± 1 (11005-11006)	1573.2(+7), 1223.8(+9), 1001.5(+11), 847.6(+13)	N-Term(Gln->pyro-Glu) S_8, S_{22} (Phospho), R_{106} removal
P-C peptide (P02810)	15.0	4370.9 ± 0.4 (4370.8)	1457.9(+3), 1093.7(+4)	Fragment 107-150 of PRP-1
Statherin				
Statherin (P02808)	29.2	5380.0 ± 0.5 (5379.7)	1794.2(+3), 1345.9(+4), 1076.9(+5)	S_2(Phospho); S_3(Phospho)

Proteins/peptides	El. Time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	PTMs
Statherin 1P	28.9	5299.9 ± 0.5 (5299.7)	1767.6(+3), 1325.9(+4), 1060.9(+5)	S_3(Phospho)
Statherin SV-1 (des-F_{43})	27.8	5232.4 ± 0.5 (5232.5)	1745.1(+3), 1309.1(+4), 1047.5(+5)	C-Term. F_{43}removal
Statherin desT_{42}-F_{43}	27.9	5131.2 ± 0.5 (5131.4)	1711.4(+3), 1283.8(+4), 1027.2(+5)	C-Term. T_{42}-F_{43}removal
Statherin desD_1	28.7	5264.7 ± 0.5 (5264.6)	1755.9(+3), 1317.2(+4), 1053.9(+5)	N-Term. D_1removal
Statherin des1-9	28.5	4127.9 ± 0.4 (4127.6)	1376.9(+3), 1032.9(+4)	N-Term. 1-9 residue removal
Statherin des1-10	28.0	3971.3 ± 0.4 (3971.4)	1986.7(+2), 1324.8(+3)	N-Term. 1-10 residue removal
Statherin des1-13	27.5	3645.2 ± 0.4 (3645.0)	1823.6(+2), 1216.1(+3)	N-Term. 1-12 residue removal
P-B peptide				
P-B peptide (P02814)	30.0	5792.9 ± 0.5 (5792.7)	1932.0(+3), 1449.2(+4), 1159.6(+5)	N-Term(Gln->pyro-Glu)
P-B des1-4	30.0	5371.0 ± 0.5 (5371.3)	1791.4(+3), 1343.8(+4), 1075.3(+5)	N-Term. 1-4 residue removal
P-B des1-5	30.3	5215.0 ± 0.5 (5215.1)	1739.4(+3), 1304.8(+4), 1044.0(+5)	N-Term. 1-5 residue removal
P-B des1-7	30.1	5060.1 ± 0.5 (5060.9)	1688.0(+3), 1266.2(+4), 1013.2(+5)	N-Term. 1-7 residue removal
P-B des1-12	27.5	4549.0 ± 0.5 (4549.3)	1517.5(+3), 1138.3(+4)	N-Term. 1-12 residue removal
Histatins				
Hst-1 (P15515)	21.9	4928.2 ± 0.5 (4928.2)	1644.1(+3), 1233.5(+4)	S2(Phospho)
Hst-1 nonphos.	22.0	4848.2 ± 0.5 (4848.2)	1617.4(+3), 1213.5(+4)	
Hst-3 (P15516)	17.7	4062.2 ± 0.4 (4062.4)	1355.1(+3), 1016.6(+4)	
Hst-6	14.3	3192.4 ± 0.3 (3192.5)	1065.1(+3), 799.1(+4)	Fragment 1-25 of Hst-3
Hst-5	14.6	3036.5 ± 0.3 (3036.3)	1013.2(+3), 760.1(+4)	Fragment 1-24 of Hst-3

Proteins/peptides	El. Time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	PTMs
Cystatins				
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)	
A acetyl	33	11047.43 ± 2 (11048.5)	1005.41(+11), 1105.85(+10), 1228.61(+9), 1382.06(+8), 1579.36(+7), 1842.42(+6)	N-Term.-α-Acetylation
B-SSG(P04080)	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)	C_3Glutathionylation
B-SSC	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)	C_3Cysteinylation
B S-S dimer	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)	1 interchain disulfide bridge
C (P01034)	35.1	13342 ± 2 (13343.1)	1483.57(+9), 1335.32(+10), 1214.02(+11), 1112.93(+12), 1027.40(+13)	2 intrachain disulfide bridges
D-R_{26} 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)	N-Term(Gln->pyro-Glu) after 1-5 residue removal, 2 intrachain disulfide bridges
Cystatins S-type				
S	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)	2 intrachain disulfide bridges
S1 (P01036)	35.3	14266 ± 2 (14265)	1784.3(+8), 1586.1(+9), 1427.6(+10), 1297.9(+11),	S_3(Phospo) on cystatin S, 2 intrachain disulfide bridges

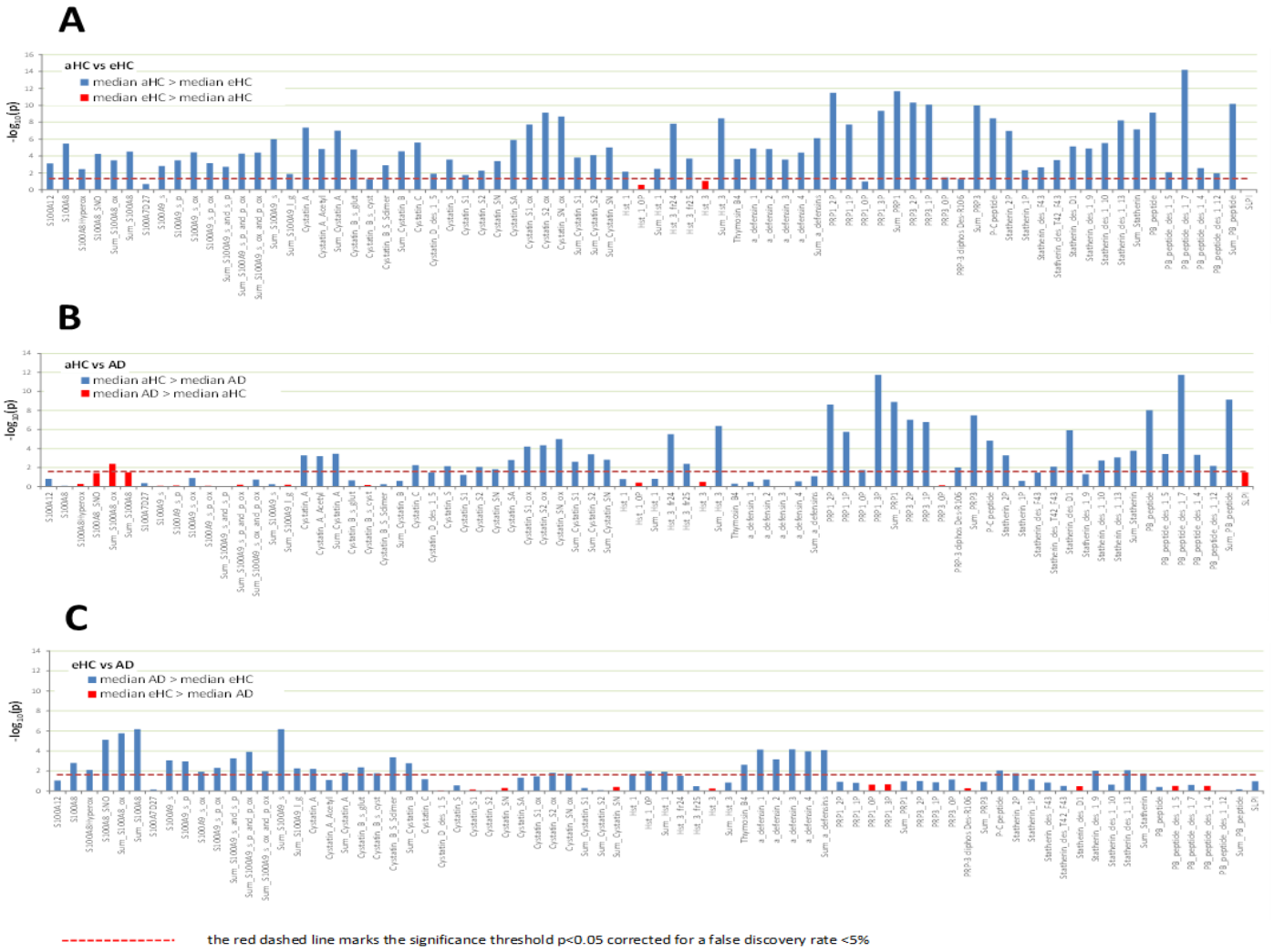
Proteins/peptides	El. Time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	PTMs
			1189.8(+12), 1098.4(+13), 1020.0(+14)	
S1 ox	35.3	14281 ± 2 (14280.7)	1786.40(+8), 1589.70 (+9), 1429.30 (+10), 1299.50 (+11), 1191.30 (+12), 1099.70 (+13)	S_3(Phospo), W_{23} oxidation, 2 intrachain disulfide bridges
S2	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)	S_1, S_3 (di-Phospho) on cystatin S, 2 intrachain disulfide bridges
S2 ox	35.3	14360 ± 2 (14361)	1596.64(+9), 1437.08(+10), 1306.52(+11), 1197.73(+12), 1105.68(+13)	S_1, S_3(di-Phospho) on cystatin S, 2 intrachain disulfide bridges, W_{23} oxidation
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)	2 intrachain disulfide bridges
SN ox	34.6	14328 ± 2 (14328)	1792.30(+8), 1593.20 (+9), 1434.00 (+10), 1303.30 (+11), 1195.20 (+12), 1103.30 (+13)	2 intrachain disulfide bridges, W_{23}oxidation
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)	1 intrachain disulfide bridge
Antileukoproteinase				
SLPI (P03973)	26.2	11702.2 ± 1 (11706)	1952.64(+6), 1673.84(+7), 1464.73(+8), 1302.10(+9)	8 intrachain disulfide bridges
α-Defensins				
α-defensin 1 (P59665)	23.5	3442.5 ± 2 (3442.1)	1772.03(+2), 1148.36(+3), 861.52(+4)	2 intrachain disulfide bridges
α-defensin 2 (P59665/6)	23.5	3370.4 ± 1 (3370.9)	1686.49(+2), 1124.66(+3), 843.75(+4)	2 intrachain disulfide bridges

Proteins/peptides	El. Time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	PTMs
α-defensin 3 (P59666)	23.5	3485 ± 2 (3486.1)	1744.03(+2), 1163.03(+3), 872.52(+4)	2 intrachain disulfide bridges
α-defensin 4 (P12838)	27.2	33708 ± 1 (3709.4)	1855.71(+2), 1237.48(+3), 928.36(+4)	2 intrachain disulfide bridges
Thymosins β4				
Tβ4 (P62328)	18.5	4964.0± 1 (4963.5)	1655.51(+3), 1241.88(+4), 993.71(+5)	N-Term.-α-Acetylation
S100A proteins				
S100A12 (P80511)	40.0	10444 ± 2 (10443.9)	1306.5(+8), 1161.4(+9), 1045.4(+10), 950.4(+11)	M_1 removal
S100A7 D_{27} (P31151)	37.0	11367 ± 2 (11367.8)	1422.0(+8), 1264.1(+9), 1137.8(+10), 1034.4(+11)	M_1 removal, N-Term.-α-Acetylation, D_{27} variant
S100A8(P05109)	40.4	10833 ± 2 (10834.5)	1355.3(+8), 1204.8(+9), 1084.5(+10), 985.9(+11)	
S100A8 hyperoxidized	39.3	10915 ± 2 (10914.6)	1365.3(+8), 1213.7(+9), 1092.5(+10), 993.2(+11)	C_{42}-SO_3H and W_{54}dioxidation or C_{42}-SO_3H and W_{54} oxidation and $M_{1/78}$oxidation
S100A8 SNO	40.8	10863 ± 2 (10863.5)	1358.9(+8), 1208.1(+9), 1087.3(+10), 988.6(+11)	C_{42}nitrosylation
S100A8 SO_3H Wox	40.4	10900 ± 2 (10899.6)	1363.46(+8), 1212.08(+9), 1090.97(+10), 991.88(+11), 909.26(+12), 839.39(+13)	C_{42} sulfonic acid and W_{54} oxidation
S100A9 (S) (P06702)	42.2	12690 ± 2 (12689.2)	1410.9(+9), 1269.9(+10), 1154.6(+11), 1058.4(+12), 977.1(+13)	N-Term.-α-Acetylation after 1-5 residue removal
S100A9 (S) 1P	42.2	12770 ± 2 (12769.2)	1419.8(+9), 1277.9(+10), 1161.8(+11), 1065.1(+12), 983.3(+13)	N-Term.-α-Acetylation after 1-5 residue removal, T_{108} (Phospho)
S100A9 (S) ox	41.3	12706 ± 2 (12705.2)	1412.7(+9), 1271.5(+10), 1156.0(+11), 1059.8(+12), 978.3(+13)	N-Term.-α-Acetylation after 1-5 residue removal, $M_{89\text{ or }78\text{ or }76\text{ or }58}$ oxidation

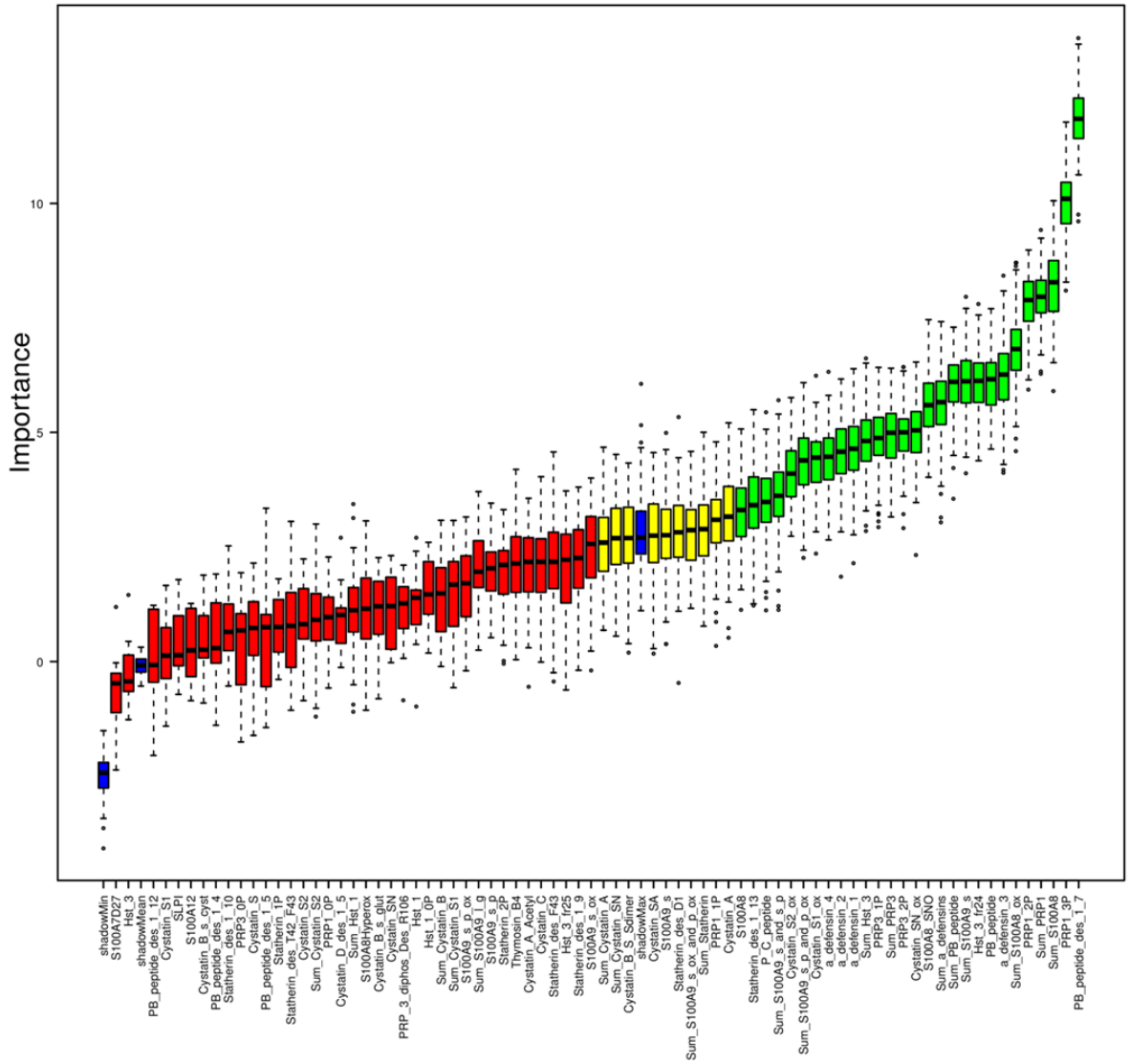
Proteins/peptides	El. Time (min ± 0.5)	Exper. (theor) Mav ± SD	m/z (charge) for XIC search	PTMs
S100A9 (S) 1P ox	41.3	12786 ± 2 (12785.2)	1421.9(+9), 1279.5(+10), 1163.3(+11), 1066.4(+12), 984.5(+13)	N-Term.-α-Acetylation after 1-5 residue removal, T_{108} (Phospho), $M_{89 \text{ or } 78 \text{ or } 76 \text{ or } 58}$ oxidation
S100A9 (L) SSG	41.5	13459 ± 2 (13458.1)	1346.8(+10), 1224.5(+11), 1122.5(+12), 1036.3(+13), 962.3(+14)	M_1 removal, N-Term.-α-Acetylation, C_2glutathionylation
S100A9 (L) SSG 1P	41.5	13538 ± 2 (13538.1)	1354.82(+10), 1231.75(+11), 1129.18(+12), 1042.40(+13), 968.02(+14)	M_1 removal, N-Term.-α-Acetylation, C_2glutathionylation, T_{108} (Phospho)
S100A9 (L) SSG ox	41.5	13476 ± 2 (13476.7)	1348.67(+10), 1226.16(+11), 1124.06(+12), 1037.67(+13), 963.63(+14)	M_1 removal, N-Term.-α-Acetylation, C_2glutathionylation, $M_{93 \text{ or } 82 \text{ or } 80 \text{ or } 62}$ oxidation
S100A9 (L) SSG 1P ox	41.5	13555 ± 2 (13555.1)	1507.13(+9), 1356.52(+10), 1233.29(+11), 1130.60(+12), 1043.71(+13)	M_1 removal, N-Term.-α-Acetylation, C_2glutathionylation, T_{108} (Phospho) $M_{93 \text{ or } 82 \text{ or } 80 \text{ or } 62}$ oxidation

SSG: glutathionylation on cysteine residue; SSC: cysteinylation on cysteine residue; S-S: formation of a disulfide bond; SNO: nitrosylation on cysteine residue; nonphos.: proteoform non phosphorylated; P: phosphorylation.

Supplementary Figure S2.1: Comparisons of proteins levels in aHC vs eHC (A), aHC vs AD (B), eHC vs AD subjects (C). The height of columns shows the negative $\log_{10}$ of Mann-Whitney tests p-values. The red-dashed line marks the significance threshold (1.3 on the $-\log_{10}$ scale, equivalent to a p-value of 0.05) adjusted for a cumulative false discovery rate <5% by the Benjamini-Hochberg procedure.



Supplementary Figure S2.2: Selection of components of aHC-eHC-AD mixed samples, by the Boruta method. Blue boxplots correspond to minimal and maximum scores of shadow variables obtained by random permutation of copies of the original variables. The max shadow score represents the threshold to select important features (green = confirmed; yellow = tentative). Red boxplots represent rejected features. Boxplots show the median, the interquartile range (IQR), the virtual minimum and maximum values, Q_1 - 1.5 IQR and Q_3 + 1.5 IQR, respectively (small bars) and outliers (dots).



PART III

Characterization of novel human cystatin B salivary interactome in Alzheimer's disease and healthy subjects

INTRODUCTION AND AIM OF THE STUDY

Cystatins are the largest group of endogenous cathepsin inhibitors expressed both intracellularly and extracellularly. They are single chain proteins divided into three inhibitory subfamilies: type 1 cystatins (cystatin A and B), type 2 cystatins (cystatin C, D, E, F, S, SN and SA) and type-3 cystatins (kininogens)[161]. Among them, cystatin B and C were found to be widely expressed in numerous tissues, including the brain, whereas the others have more restricted distribution[174].

Cystatin B is a small protein found involved in many diseases and biological processes, strongly suggesting additional multiple functions other than cathepsin inhibitor. Its ability to colocalize with amyloid plaques of AD has been proven and it has been proposed as amyloid constituent[76]. *In vitro* studies have shown that cystatin B oligomers can inhibit or facilitate Aβ fibril growth according to their size[77]. Actually, the existence of cystatin B polymeric structures in terms of different long homo-oligomers has also been demonstrated in cellular systems[175] as well as the ability of cystatin B to interact with other proteins[77,176] and to form a specific multiprotein complex in the cerebellum of rat[177].

Our previous work on the acid-soluble protein fraction of the saliva revealed the higher expression of cystatin B in AD patients[146], but the presence of high molecular weight oligomers or multiprotein complexes related to cystatin B has never been investigated in saliva. The first aim of the present study has been to demonstrate the existence of a multiprotein complex involving cystatin B in whole saliva samples and to characterize it, secondly, we further studied the cystatin B interactome to individuate any qualitative-quantitative difference between AD patients and healthy subjects age and sex matched. To these aims, we performed some preliminary tests mainly by western blot to look for high molecular weight positive signals related to cystatin B in saliva, and, then, in order to

characterize the presence of potential multiprotein complexes, we performed Co-IP assays followed by in gel tryptic digestion and RP-nanoHPLC-HR-ESI-MS/MS analysis.

The experiments and research activity of the present section begun at the University of Cagliari (Cagliari, Italy) where experimental design, preliminary tests and Co-IPs have been performed, under the supervision of Prof. Tiziana Cabras. The characterization of cystatin B interactome through RP-nanoHPLC-HR-ESI-MS/MS and data analysis has been performed at the Max Planck Institute of Psychiatry (Munich, Germany) under the supervision of Prof. Christoph Turck and Dr. Giuseppina Maccarrone.

MATERIALS AND METHODS

Reagents

Co-Immunoprecipitation (Co-IP) assay was performed with SureBeads™ Protein G Magnetic Beads purchased from Bio-Rad (Hercules, California, USA). The bait protein was captured with cystatin B mouse monoclonal antibody purchased from Invitrogen (Thermo Fisher Scientific, Waltham, MA, USA). Experimental negative control was obtained using normal mouse IgG Polyclonal Antibody from Merck (Darmstadt, German). Chemicals and materials used to perform SDS-PAGE and enzymatic in gel-digestion were purchased from Bio-Rad (Hercules, California, USA), SDS-PAGE analyses were run onto 4–15% Mini-PROTEAN® TGX™ Precast Protein Gels with 10-well of 30 µl capacity; trypsin (MS-approved) was purchased from SERVA (Heidelberg, Germany). Chemicals and reagents for Mass Spectrometry analyses were purchased from Sigma-Aldrich/Merck (Darmstadt, Germany). Chemicals and reagents for western blot were purchased from Bio-Rad (Hercules, California, USA).

Study subjects, Sample collection and treatment

For the purposes of this study, samples of unstimulated whole saliva were collected from the same subjects enrolled as described in the "Material and Methods" section of the Part I, with the exception of three AD patients (#12, #39 and #40) and five HC subjects (#21, #36, #37, #38 and #39) whose saliva has been treated in the same way but being enrolled later in the study, they were not included in the previous sections of the present thesis. Donors provided variable volumes of whole saliva according to their capacity and disposal. In some cases, the volume and TPC of available whole saliva was insufficient for the purpose of the present section, so it was not possible to include all the patients and healthy volunteers in the interactome study. For this reason, fourteen AD patients (#1, #2, #3, #5, #6, #7, #8, #9, #10,

#11, #21, #28, #29, #31) and five HC subjects (#1, #6, #29, #32, #33) who were part of the study subjects in the previous sections, have been excluded. Demographic features of the subjects involved in this section are reported in Table 3.1: twenty-four subjects affected by AD were included (17 females and 7 males; mean age and SD: 81 ± 5), comprising twenty-one subjects among those enrolled for the first study plus three subjects later enrolled (#12, #39 and #40); while the HC group was composed by thirty-four healthy volunteers including twenty-seven subjects among those enrolled for the first study plus five subjects later enrolled (#21, #36, #37, #38 and #39). Healthy volunteers have been split to form two different groups: one named "HC" composed by twenty-four subjects (16 females and 8 males; mean age and SD: 78 ± 4) chosen to be coherent with AD group for what concerns number, sex and age of the subjects; another one named "NEG" composed by twenty-six subjects (15 females and 11 males; mean age and SD: 77 ± 3), dedicated to serve as experimental negative control (NEG) in the Co-IP assay and thus composed by a higher number of subjects to be more balanced in gender contribution. Samples of whole saliva for interactome study were immediately diluted in a 1:1 v/v ratio with PBS buffer (270mM NaCl, 5mM KCl, 20mM $NaHPO_4$, 4mM KH_2PO_4) containing cOmplete™, EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich/Merck, Darmstadt, Germany), and gently centrifuged at 1000 g for 5 minutes at 4°C before being stored at -80° until the analysis.

Western blot analyses

All immunodetections by western blot have been performed as follows: SDS-PAGE was made with 4-15% T mini precast gels (Bio-rad, Hercules, California, USA) according to Laemmli protocol[178]. Electrophoretic separation was performed at 180 V and Bio-rad Precision Plus Protein™ WesternC™ Blotting Standards were used as molecular weight standards. Proteins were transferred to 0.2 μm PVDF membranes according to the instructions

provided with the Trans-blot Turbo system (Bio-Rad, Hercules, CA, USA). After the transfer, PVDF membranes were equilibrated for 1 hour with the blocking solution (5% Blotting-Grade Blocker, Bio-rad, in TBS containing 0.05% tween-20, TBS-T), and then, for 1 hour under stirring with cystatin B mouse monoclonal primary antibody diluted 1:1000 with TBS-T. After five x 5 min washing with TBS-T, membranes were incubated for 1 hour with the anti-mouse secondary Ab (HRP conjugated dil. 1:5000 in TBS-T). After five x 5 min washing with TBS-T, membranes were incubated with the detection solution (Clarity Max™ Western ECL Substrate, Bio-Rad) according to the manufacturer's instructions, and the detection was performed with the ChemiDoc MP Imaging System (Bio-Rad, Hercules, CA, USA) and analysed with Image Lab 4.0.1. When needed, statistics by GraphPad Prism software (version 5.0) was applied to the signal intensities obtained. Western blot immunodetections were performed in different set of samples:

1) The acid-soluble protein fractions of saliva prepared as described in the Part I. The residual volume of the acid-soluble fraction of saliva from 28 patients among those enrolled for Part I was pooled into a single pool (17 females and 11 males; mean age and SD: 80 ± 6, Supplementary Table S3.1). For comparison, the acid-soluble fraction of saliva from 28 HC subjects patients among those enrolled for Part I was also pooled (15 females and 13 males; mean age and SD: 78 ± 5, Supplementary Table S3.1). The total protein content for both pools was 60 μg and every subject contributed with the same protein quantity according to their TPC. Pools have been divided into aliquots of 20 μg/each to be analyzed in triplicate. As positive control, 1 μg of enriched cystatin B fraction (Material and Methods, Part I) has been analysed as well. All the samples have been mixed 1:1 (v/v) with 0.125 M Tris/HCl pH 6.8 containing 4% SDS, 20% glycerol, 0.02% bromophenol blue without reducing agent to keep intact any possible covalent interaction.

2) The whole saliva of adult and elderly HC subjects. Four aHC (two females aged 28 and 34; two males aged 30 and 39) and four eHC (two females aged 73 and 78; two males aged 78 and 79) were randomly selected. Whole salivary samples, already treated with PBS, were mixed 1:1 (v/v) with 0.125 M Tris/HCl pH 6.8 containing 4% SDS, 20% glycerol, 0.02% bromophenol blue without reducing agent to keep intact any possible covalent interaction. According to their TPC, a total of 35 μg *per* sample were utilized.
3) Whole salivary pools from 24 AD and 24 HC. Whole salivary samples treated with PBS were pooled into a single pool from 24 AD samples and a single pool from 24 HC samples (AD and HC, Table 3.1). The total protein content for both pools was 60 μg and every subject contributed with the same protein quantity according to their TPC. Pools have been divided into aliquots of 30 μg to be analyzed in duplicate. All the samples have been mixed 1:1 (v/v) with 0.125 M Tris/HCl pH 6.8 containing 4% SDS, 20% glycerol, 0.02% bromophenol blue without reducing agent to keep intact any possible covalent interaction.
4) Immunoprecipitated (IP) samples of AD and HC pools from whole salivary samples in PBS starting from 300 μg of total protein content. After Co-IP assay with cystatin B antibody, 4.5 μg of IP samples from 24 AD and 24 HC pools (AD and HC, Table 3.1) were mixed 1:1 (v/v) with 0.125 M Tris/HCl pH 6.8 containing 4% SDS, 20% glycerol, 0.02% bromophenol blue with or without 2% 2-mercaptoethanol to perform analysis in reducing (R-) and non-reducing (NR-) conditions. Co-IP assay has been performed as described in the corresponding following paragraph.

SDS-PAGE and in gel tryptic digestion of enriched cystatin B fraction and acid salivary pools from AD and HC subjects

To confirm the presence of cystatin B in all the gel slices showing a positive signal after western blot analysis, 100 μg/each of the two pools from acid-soluble protein fractions of saliva of 28 AD patients and 28 HC subjects (Supplementary Table S3.1) and 5 μg of enriched cystatin B fraction (Material and Methods, Part I) have been subjected to SDS-PAGE and in gel digestion. All the samples have been mixed 1:1 (v/v) with 0.125 M Tris/HCl pH 6.8 containing 4% SDS, 20% glycerol, 0.02% bromophenol blue without reducing agent to keep intact any possible covalent interaction and then boiled at 100°C for 5 min. After SDS-PAGE separation, gels were stained with Bio-Safe TM Coomassie G250 stain (Bio-Rad, Hercules, CA, USA) to be submitted to in-gel tryptic digestion and HR-MS/MS analysis. Only stained protein bands corresponding to cystatin B positive signals evidenced by western blot were manually excised from the gel and transferred to fresh tubes.

Gel pieces were vortex for 10 min two times in a solution of 25mM ammonium bicarbonate/ ACN (1:1, v/v) to be de-stained. Reduction and alkylation of cysteine residues were performed with 10mM dithiothreitol (30 minutes at 56°C, in the dark) followed by the addition of 55 mM iodoacetamide (30 min at RT, in the dark). Gel pieces were washed again two times with 25mM ammonium bicarbonate/ACN (1:1, v/v) and rehydrated in 25mM ammonium bicarbonate containing 100 ng of trypsin. The samples were incubated overnight at 37ºC to allow tryptic digestion. Trypsin reaction was blocked by adding 0,1% of FA. To extract peptides, samples were added 50 ml of 50% ACN, 1% of FA, sonicated for 5 min and vortexed for 20 min at room temperature. Extraction procedure was repeated twice and tryptic peptides were lyophilized again and stored at -20°C until the analysis.

RP-nanoHPLC-high resolution ESI-MS/MS analysis of enriched cystatin B fraction and acid salivary pools from AD and HC subjects

Samples from in gel digestion obtained on enriched cystatin B fraction and acidic AD and HC pools were resuspended in solvent A (0,1% FA) and analyses were performed with a Ultimate 3000 Nano System HPLC (Thermo Fisher Scientific, Sunnyvale, CA) coupled with a LTQ Orbitrap Elite (Thermo Fisher Scientific, San Jose, CA, USA). The Easy Spray reverse-phase nano column (150 mm x 50 µm inner diameter, Thermo Fisher Scientific) was a C18 with 2 µm beads and elution of peptides was achieved with solvent B (0.1% formic acid/80%ACN v/v) at flow rate of 0,3 µl/min with the following gradient: 0-3 min to 4%; 3-70 min to 50 % solvent B; 70-90 min to 80%; 90-92 min to 90%; 100-103 min to 98% solvent B. The mass spectrometer was operating at 1.7 kV in the data dependent acquisition mode with the capillary temperature set at 275 °C. Full MS experiments were performed in positive ion mode with mass ranging from 350 to 1600 m/z (resolution 120000). The 10 most intense ions were subjected to CID fragmentation setting 35% of normalized collision energy for 10 ms, isolation width of 2 m/z, and activation q of 0.25. Spectra were processed and analyzed by PD software (version 2.2, Thermo Fisher) using SEQUEST HT cluster search engine (University of Washington, licensed to Thermo Electron Corporation, San Jose, CA) against the UniProtKB human database (188,453 entries, release 2019_03). Database search parameters were as follows: carbamidomethylation of cysteine as fixed modification, oxidation of methionine and tryptophan, serine/threonine phosphorylation, C-terminal pyroglutamic residue, N-terminal acetylation and methionine loss as dynamic modification 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. Peptides were filtered for high confidence and a minimum length of 6 amino acids, while proteins were filtered for high FDR confidence excluding only keratins as skin proteins deriving from sample manipulation. Settings of FDR

were 0.01 (strict) and 0.05 (relaxed). The results obtained on the characterization of cystatin B in the different gel slides by high-resolution MS/MS experiments have been deposited to ProteomeXchange Consortium (http://ww.ebi.ac.uk/pride) via the PRIDE[81] partner repository with the dataset identifier PXD030561.

Co-IP assay

Three pools of whole salivary samples in PBS were prepared: one from the 24 AD samples included in this third part of the study (AD pool, 1200 μg of total protein content); one from 24 HC samples (HC pool, 1200 μg of total proteins); and one from 26 HC samples (NEG pool, 800 μg of total proteins) (Table 3.1). Every subject contributed with the same protein quantity according to their TPC. Pools have been divided into aliquots of 400 μg/each to be independently submitted to the experimental protocol: AD and HC in triplicate, NEG in duplicate. Co-IP assay was performed using 100 μl of Protein G Magnetic Beads (Bio-Rad) *per* experiment as suggested by the manufacturer instructions. Briefly, beads were first magnetize with a specific magnetic rack to discard their store-solution and then washed by resuspension in 1 mL of PBS buffer containing 0.1% Tween-20 (PBS-T). Supernatant was discarded after beads magnetization and the washing step was repeated 3 times. Then, beads were resuspended in 200 μl of PBS-T containing 4μg of cystatin B mouse monoclonal antibody (AD and HC in triplicate) or 4μg of normal mouse IgG polyclonal antibody (NEG in duplicate) and gently rotated for 10 min at room temperature. Only when the IP samples from whole salivary AD and HC pools were destined to western blot analysis 3 μg of cystatin B mouse monoclonal antibody were used in the Co-IP assay, because the initial total protein content of pools was 300 μg instead of 400 μg. To remove unbound antibody in excess, beads were magnetized and washed again three times with PBS-T. Salivary pools were added to the beads-antibody structure and gently rotated for 1 hour at room temperature. To remove

unspecific proteins binding, wash step with PBS-T was performed again three times. Finally, the protein complex pull-down was collected resuspending the beads in 50 μl of Laemmli buffer in absence of glycerol and bromophenol blue and incubating for 10 min at 70°C, in order to break the interaction between the G protein and the antibody and between the antibody and the proteins. In the end, we obtained three cystatin B IP from AD, three cystatin B IP from HC and two normal mouse IP from NEG.

Quantification, SDS-PAGE and tryptic digestion of IP samples

To quantify the pull-down protein complex, 10 μl of each IP was subjected to BCA assay performed in duplicate using a NanoDrop 2000 Spectrophotometer (Sigma-Aldrich/Merck, Darmstadt, Germany). When performing SDS-PAGE, samples were mixed and boiled for 5 min at 100°C with Laemmli buffer with or without 0.2% of β-mercaptoethanol to obtain results in R- and NR- conditions, respectively. IP obtained from AD and HC were divided in two aliquots: 6 μg of each replicate was destinated to SDS-PAGE in NR-condition and 6 μg to R-condition. SDS-PAGE of NEG IP was performed only in R-condition, in duplicate. After SDS-PAGE, gels were stained with Bio-Safe TM Coomassie G250 stain (Bio-Rad, Hercules, CA, USA) to be submitted to in-gel tryptic digestion and HR-MS/MS analysis. To perform in depth profile of the immunoprecipitated proteins in AD and HC pools, and in the NEG experiments, the SDS-PAGE gels obtained under NR- and/or R- conditions were submitted to an extensive excision along each lane. Stained protein bands were manually excised from the gel and transferred to fresh tubes. The zones of the lanes that did not appear stained were excised as well. De-stain procedure, trypsin digestion and peptides' extraction has been performed as previously described. Tryptic peptides were lyophilized under vacuum using a SpeedVac system (Sigma-Aldrich/Merck, Darmstadt, Germany). The dried peptides were resuspended in 2% FA and filtered with Corning® Costar® Spin-X® Plastic Centrifuge Tube

Filters, cellulose acetate membrane, pore size 0.22 μm (Sigma-Aldrich/Merck, Darmstadt, Germany) to filter out gel pieces residual. The samples were then subjected to desalting protocol prior to the MS/MS analysis by using Pierce C18 zip tip, 10 μl (Thermo Fischer, Thermo-Fisher Scientific San Jose, CA, USA) following the manufacturer instructions. The desalted peptides were lyophilized again and stored at -20°C until the LC-MS/MS analysis.

RP-nanoHPLC-high resolution ESI-MS/MS analysis

Tryptic peptides extracted from each gel slice after Co-IP assays and SDS-PAGE were resuspended in solvent A (0,1% FA) and analyses were performed with a Q-Exactive Plus Mass Spectrometer (Thermo Fisher, San Jose, CA, USA), coupled to the Nano Spray Flex source and connected to an Ultimate 3000R (Dionex, Sunnyvale, CA) nano HPLC system. Reverse phase chromatography was performed using a chromatographic nano column (150 mm x 75 μm inner diameter) in-house pulled (Puller P-1000, Sutter instrument, Germany), packed with ReproSil-Pur C18 beads 1.9 μm (Dr. Maisch GmbH, Germany). Elution of peptides was achieved with solvent B (0.1% FA/95% ACN v/v) at flow rate of 0,3 μl/min. When analyzing peptides from cystatin B Co-IPs the gradient used was as following: 0-15 min to 2%; 15-125 min to 30%; 125-145 min to 60%; 145-146 min to 98% solvent B. For elution of peptides from normal mouse Co-IPs the gradient used was as follows: 0-15 min to 2%; 15-75 min to 30%; 75-80 min to 40%; 80-90 min to 98% solvent B. The mass spectrometer was operating at 1.9 kV in data dependent acquisition mode with the capillary temperature set to 275 °C. A MS^2 method was set including a FT survey scan from 375 to 1400 m/z (resolution 70,000; AGC target 3E6). The 10 most intense ions were considered for HCD fragmentation (resolution 17,500; AGC target 1E5, (N)CE 27%, max. injection time 100 ms, dynamic exclusion 30s). Spectra were acquired with the support of Xcalibur software (Thermo Fisher).

All the gel pieces were digested with trypsin and analyzed individually by nano-HPLC-HR-MS/MS, thus, in total, MS analyses were performed on 20 samples from NEG IP in duplicate (2 x 10 gel pieces in R-condition), 78 samples from AD IP in triplicate (3 x 13 gel pieces for a total of 39 in NR- condition and 3 x 13 gel pieces for a total of 39 in R- condition, respectively) and 78 samples from HC IP in triplicate (3 x 13 gel pieces for a total of 39 in NR- condition and 3 x 13 gel pieces for a total of 39 in R- condition, respectively).

Spectra were processed and analyzed using Proteome Discoverer (PD, version 2.4, Thermo). PD analyses were conducted with SEQUEST HT cluster search engine (University of Washington, licensed to Thermo Electron Corporation, San Jose, CA) against the UniProtKB human database (188,453 entries, release 2019_03). Database search parameters were set as following: carbamidomethylation of cysteine as fixed modification, oxidation of methionine and tryptophan, serine/threonine phosphorylation, C-terminal pyroglutamic residue, N-terminal acetylation and methionine loss as dynamic modifications and the allowance for up to two missed tryptic cleavages. The peptide mass tolerance was set to10 ppm and 0.02 Da for peptide and fragment ion, respectively. Peptides were filtered for high confidence and a minimum length of 6 amino acids, while proteins were filtered for a minimum number of unique peptides of 2, excluding those considered contaminants deriving from both sample manipulation (i.e. skin proteins as keratins) and the sample itself (as hemoglobin or immunoglobulins). Settings of FDR were 0.01 (strict) and 0.05 (relaxed). Label Free Quantification (LFQ) has been performed for proteins deriving from SDS-PAGE in R-conditions, thus the spectra analysis was qualitative for NR-condition and both qualitative and quantitative for R-condition. MS raw files belonging to the same SDS-PAGE lane were loaded into the software as fractions and merged into an individual sample to avoid loss of information in LFQ of proteins detectable in adjacent gel positions. LFQ abundancies of proteins were normalized against the total peptide amount in the Precursor Ion Quantifier

node of PD software. Grouping and quantification was set as non-nested specifying a single categorical factor related to the condition (AD versus HC); the software calculated, for the comparison between the two groups, the median of ratios between abundancies of the peptides belonging to a specific protein. Spectra analysis from NEG samples was performed separately from AD and HC because of differences in the chromatographic gradient, but under the same PD conditions. All the results obtained by HR-MS/MS experiments have been deposited to ProteomeXchange Consortium (http://ww.ebi.ac.uk/pride) via the PRIDE[81] partner repository with the dataset identifier PXD030679.

Bioinformatic tools for statistics and data analysis

The minimum background proteins required by PD to perform a t-test and calculate *p*-values in the comparisons between groups was not reached, thus statistics was performed with Perseus (version 1.6.15.0, Max-Planck-Institute of Biochemistry) following the instructions provided within the Perseus use case for Label-free interaction data (available at http://www.coxdocs.org/doku.php?id=perseus:user:use_cases:start) [179].

Briefly, LFQ intensities calculated by PD software have been loaded into Perseus and data have been transformed into logarithm with the default formula "log2(x)" to improve the normality of data distribution. When a protein was not detected in one replicate or in one group, the LFQ value input in the matrix was "0". The proteins LFQ intensities measured in AD with respect to those measured in NEG and the protein~~s~~ LFQ intensities measured in HC with respect to those measured in NEG were compared separately. Replicates have been grouped assigning a categorical factor so that the tool could recognize LFQ intensities of proteins from replicates belonging to the same group. This step allows to remove from the matrix any protein that was randomly identified by PD in only one replicate, setting a minimum of 2 or 3 valid values in at least one group according to the number of replicates

per group. The imputation command was then used to replace missing values with random numbers drawn from a normal distribution. Specifically, we let Perseus assume normally distributed data, so that it could calculate width and center of the distribution and shrinks the distributions to a factor of "0.3" (width), shift it down by "1.8" (down shift) standard deviations and simulate random values that can fill up the missing values. When comparing AD or HC LFQ abundances with respect to NEG, we performed a two sample t-test (Student t-test) setting s0 = 1.5 as fold change and permutation-based FDR = 0.05 with 250 randomizations. Proteins were considered significant when the *p*-value was less than 0.05 and the fold change above to ± 1.5. When comparing LFQ abundances between AD and HC groups, we performed a two sample t-test (Student t-test) setting s0 = 1 as fold change and permutation-based FDR = 0.05 with 250 randomizations. Proteins were considered significant when the *p*-value was less than 0.05 and the fold change above to ± 1. Unspecific interactors were determined performing Perseus statistical analysis on the results obtained in Co-IP assays with cystatin B antibody with respect to those obtained in Co-IP assays with normal mouse IgG polyclonal antibody. Specifically, when proteins' intensities were higher in NEG assay or unchanged between NEG and AD or NEG and HC, proteins were filtered out as unspecific interactors. The remaining proteins candidates for being cystatin B interactors were further tested against the Contaminant Repository for Affinity Purification (CRAPome, version 2.0, available at https://reprint-apms.org/), which is a large database of standardized negative controls usable to exclude potential unspecific interactors[180]. The list of genes corresponding to the protein candidates were submitted to the repository selecting *Homo Sapiens* as organism and Tandem Epitope Tag AP-MS as experiment type. When a protein was measurable in more than 25% of the reported experiments (*i.e.* found more than 180 times for a total of 716 experiments) it was considered an unspecific interactor and thus excluded. The complete list of proteins identified, included those marked as unspecific interactors, is

reported in Supplementary table S3.2 where the multiply-charged ions (m/z) used by PD for high-resolution MS/MS characterization are indicated. To verify if there were any difference in cystatin B interactors between AD and HC groups, Perseus statistics was further applied by comparing candidate proteins' intensities. Finally, Perseus statistics was applied to the same proteins after calculating the ratio bait/prays in the two groups.

Interactions and functional enrichment analysis

The list of identified cystatin B interactors was submitted to STRING database (version 11.5, Szklarczyk *et al.*, 2021) to check for known interactions already characterized among the proteins involved. In the settings' panel, the active interaction sources selected were "experiments" and "databases" and minimum required interaction score set at 0.900 (highest confidence). The pathway analysis of the entire network of cystatin B and its interactors referring to Reactome database was also performed by STRING.

The functional enrichment analysis on all identified cystatin B interactors was performed using DAVID (Database for Annotation, Visualization and Integrated Discovery)[182]. Specifically, the list of cystatin B and its interactors were analyzed by the Functional Annotation Chart to test the statistical enrichment of Gene Ontology (GO)-terms (CC – Cellular Component, MF –Molecular Function, BP – Biological Process) associated to the identified proteins by *p* value through the EASE score (an alternative Fisher test) and applying Benjamini-Hochberg post-test as correction. A panoramic of all CC, MF and BP annotated GO-terms for cystatin B and all its interactors is shown in supplementary figure S3.1 obtained from g:profiler[183]. The analysis was performed using g:GOSt and visualized by Gantt-like charts (Fig. S3.1).

Table 3.1. Demographic data of 24 HC (named HC), 24 AD patients (named AD) and 26 HC (named NEG) included in the study and used to perform Co-IP. For each subject it is indicated the TPC of whole saliva determined by BCA assay.

HC	Sex and age	TPC μg/μl	AD	Sex and age	TPC μg/μl	NEG	Sex and age	TPC μg/μl
#3	F,84	1.14	#12	F,84	0.61	#2	M,85	1.8
#5	F,81	1.35	#13	F,79	0.71	#3	F,84	1.1
#9	M,71	1.48	#14	F,82	0.89	#4	M,82	1.6
#11	M,76	1.75	#16	F,83	0.56	#5	F,81	1.4
#12	F,77	1.22	#17	F,63	0.73	#7	F,79	1.6
#13	M,74	1.18	#18	F,80	0.87	#8	M,74	2.2
#14	M,87	2.86	#19	F,80	0.35	#9	M,71	1.5
#16	F,81	0.93	#20	M,87	0.49	#10	M,78	1.5
#17	F,82	1.48	#22	M,87	0.84	#11	M,76	1.7
#19	F,86	1.85	#23	F,75	0.37	#12	F,77	1.2
#20	F,73	3.20	#24	F,75	0.28	#13	M,74	1.2
#21	M,79	2.17	#25	F,83	0.16	#15	M,73	1.7
#22	F,78	1.73	#26	F,84	0.40	#16	F,81	0.9
#24	F,78	1.20	#27	F,81	0.34	#18	F,72	1.7
#25	F,75	2.12	#30	M,86	0.21	#20	F,73	3.2
#26	F,75	0.83	#32	F,88	0.30	#21	M,79	2.2
#30	F,76	0.36	#33	F,81	0.34	#22	F,78	1.7
#28	F,78	0.47	#34	F,77	0.42	#23	F,79	2.1
#31	F,81	0.38	#35	M,87	0.78	#24	F,78	1.2
#34	F,72	0.16	#36	M,84	0.99	#25	F,75	2.1
#35	M,80	1.18	#37	F,77	0.61	#26	F,75	0.8
#37	F,78	3.65	#38	F,78	1.19	#28	F,78	0.5
#38	F,78	2.18	#39	M,76	0.77	#29	M,73	0.3
#39	F,83	1.69	#40	M,85	1.61	#35	M,80	1.2
						#36	F,73	2.9
						#38	F,78	2.2

RESULTS

The results obtained in the present section concern the study of multiprotein complexes linked to cystatin B in whole saliva samples. The results of preliminary tests will be firstly illustrated and then the results obtained by Co-IP assays.

Preliminary tests

In a series of preliminary experiments, we observed the presence of positive signals to cystatin B antibody by western blot in molecular weight ranges higher than those expected, which are ca. 11 kDa for cystatin B monomeric form and ca. 22 kDa for cystatin B dimeric form. Cystatin B western blot immunodetection was performed after SDS-PAGE in NR-conditions of the acid-soluble protein fractions of saliva from AD and elderly HC group in pools. An enriched fraction of cystatin B purified from a young adult healthy control has been loaded in line 2, followed by AD pool in triplicate in lines 3, 4, 5 and HC pool in triplicate in lines 6, 7, 8 (Fig. 3.1, panel A). In the lines of AD pool several medium/high molecular weight signals were evident above and in correspondence of 250 kDa and below 75 kDa; as well as in the range of 25 kDa (cystatin B dimeric form) and between 10-15 kDa (cystatin B monomeric form). The same signals were present also in HC pool lines, even if less intense. In line 2, cystatin B was detected in its monomeric and dimeric form at 10-15 kDa and 25 kDa, respectively; two more signals were exclusively detected in this line below 50 kDa and below 37 kDa, while no positive signal for cystatin B was registered in the range of 250 or 75 kDa (Fig. 3.1, panel A).

Statistical analysis was performed on the signals of cystatin B in its monomeric form (Fig 3.2, panel A), dimeric form (Fig 3.2, panel B), and the sum of the two (Fig 3.2, panel C) with respect to those of the signals of the enriched cystatin B fraction used as standard at the same relative molecular weights. Statistics confirmed the results obtained in the Part I of the present

thesis, showing that AD patients have higher levels of cystatin B with respect to HC in its monomeric and dimeric form.

To confirm the presence of cystatin B in all the gel slices showing a positive signal after western blot analysis and to exclude the possibility of being observing unspecific signals of the antibody, we performed SDS-PAGE in NR-conditions followed by in gel digestion with trypsin and nanoHPLC-HR-ESI-MS/MS analysis of the same acid salivary pools and the enriched fraction of cystatin B. In panel B of Figure 3.1 it is shown which areas of the gel have been cut that correspond to the positive signals observed in western blot analysis (Fig. 3.1, panel A). Through the PD analyses, several proteins co-migrating together with cystatin B have been identified into the analyzed slices (data not shown), however, since their presence in the same slice is not indicative for a possible interaction, only results concerning cystatin B characterization are reported. The outcomes of PD analyses are listed in Table 3.2 for enriched cystatin B fraction. Table indicates Sequest HT score, percentage of sequence coverage, number of Peptide Spectrum Matches (PSMs) and unique peptides. Table 3.3 shows the same outcomes for AD and HC pools where relative LFQ abundances are also indicated. The presence of cystatin B has been confirmed in all the gel positions, except for slices 1 and 2 in HC pool, probably because the protein level was below our instrumental limit. To better understand the presence of the cystatin B signals between 50 and 30 kDa observed only in the standard of cystatin B, purified from saliva collection from a not elderly donor, another experiment was performed, where the whole saliva was directly used to obtain a wide-ranging overview of cystatin B immunodetection. Thus, in this case, immunodetection by western blot included only individual whole saliva samples from healthy subjects divided into 4 aHC and 4 eHC (Fig. 3.1, panel C). Cystatin B signals were easier detectable in the 4 eHC (from number 5 to number 8) than in the 4 aHC (from number 1 to number 4) both at low and at medium/high molecular weights (Fig. 3.1, panel C). The blue asterisks in the figure

highlight cystatin B signals in the blot for medium/high molecular weights which appeared in the same positions of those shown in panel A of the same figure below 75 kDa and above 250 kDa. On the contrary, the two signals detected below 50 and 37 kDa in the enriched fraction of cystatin B (Fig. 3.1, panel A, line 2) were not detected in the four aHC or in the four eHC (Fig. 3.1, panel C), thus, most likely, they represent signals proper of individual expression of cystatin B in the saliva of the young donor used to collect the enriched fraction.

The volume and TPC of whole saliva samples available were a limiting factor in AD group, so it was not possible to perform a cystatin B screening in individual samples. We decided instead to pool together the whole saliva from 24 AD patients and 24 HC subjects (AD and HC pools, Table 3.1) and to perform again cystatin B immunodetection by western blot (Fig. 3.1, panel D), in duplicate. In this case the intensity of signals between AD patients and HCs appeared to be comparable; blue asterisks evidence the immunodetection of the same cystatin B positive signals above 250 kDa and below 75 kDa, respectively, detected in the previous western blots (Fig. 3.1, panels A and C) with the addition of a faint positive signal around 150 kDa highlighted in the figure by gray circles (Fig. 3.1, panel D). Red asterisks in Fig. 3.1 refers to the immunodetection of monomeric and dimeric forms of cystatin B (panels A, C and D); in the western blot reported in panel D we were not able to detect cystatin B monomeric form.

The preliminary results we obtained through SDS-PAGE and western blot analyses evidenced the presence of cystatin B in medium/high molecular weight ranges suggesting the possibility that this protein could be involved in multi-protein complexes. With the aim of characterize the protein partners cystatin B was interacting with, we chose to perform Co-IP assay using cystatin B mouse monoclonal antibody to catch cystatin B as the bait protein together with all its partners. Specifically, according to the results obtained in the Part I, highlighting the role cystatin B may play in AD, and considering the high variability in the composition of saliva

with respect to subjects' age (Part II of the thesis), we investigated the cystatin B interactome in the whole saliva from AD patients and healthy subjects age and sex matched.

Figure 3.1: Western blot immunodetection of 1 µg of enriched cystatin B fraction and 20 µg/each of acid salivary pools from AD and HC subjects in triplicate (A). SDS-PAGE of 5 µg of enriched cystatin B fraction and 100 µg/each of acid salivary pools from AD and HC subjects (B) mirroring the western blot immunodetection (A). The dashed lines in both panels represent the excised slices corresponding to the positive signals of immunodetection. Cystatin B immunodetection by western blot of healthy subjects' individual whole saliva samples (35 µg TPC/each, panel C) and whole salivary pools from 24 HC subjects and 24 AD patients (HC and AD pools, Table 3.1) in duplicate (30 µg TPC/each, panel D). In both panels, blue asterisks highlight cystatin B positive signals detected in medium/high molecular weight ranges; red asterisks highlight positive signals corresponding to dimeric and monomeric cystatin B. Grey circle in panel D highlight a cystatin B positive signal detected at 150 kDa.

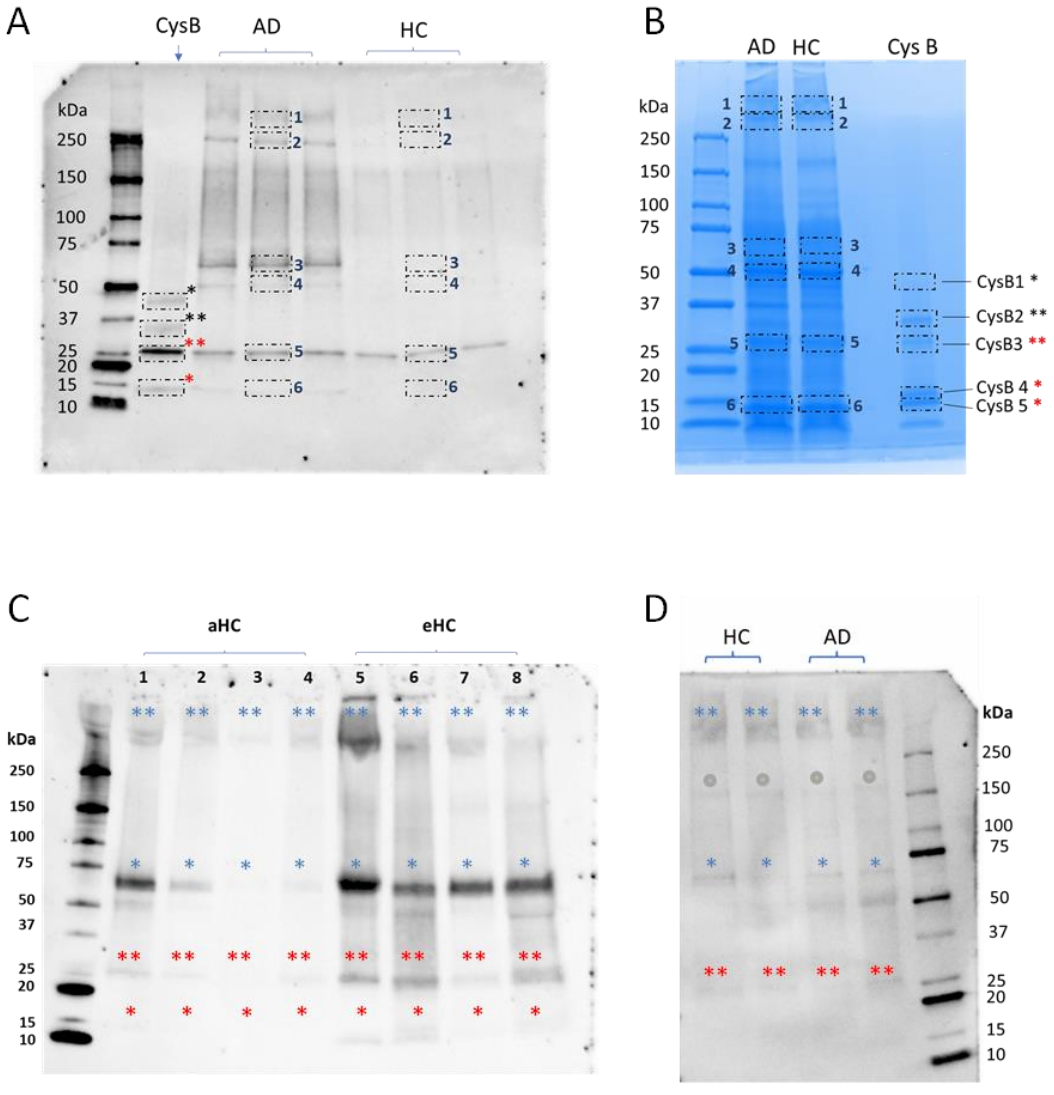


Fig. 3.2: Statistical analysis on the signals of cystatin B obtained by western blot (Fig. 3.1, panel A) in its monomeric (A) and dimeric (B) form in AD and HC groups. Signals have been normalized with respect to those of enriched cystatin B fraction signals at the same relative molecular wights. The sum of both dimeric and monomeric forms (C) has also been tested. *p* value has been considered significant when below 0.05 (*).

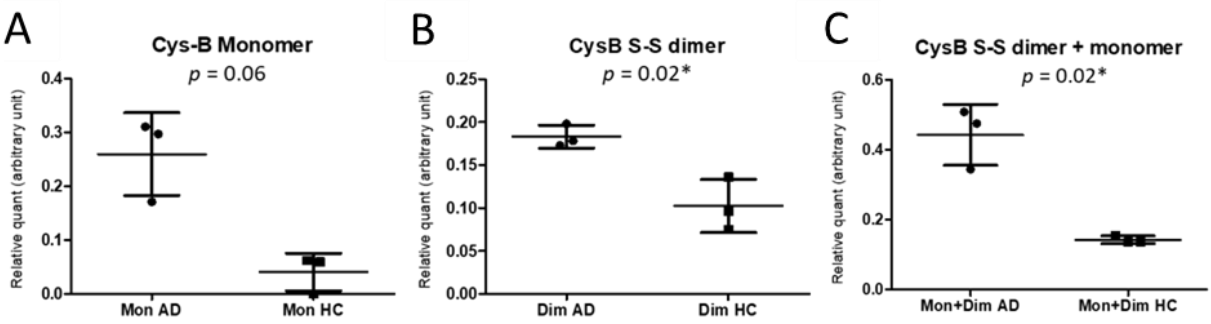


Table 3.2: Identification of cystatin B in gel slices from enriched cystatin B fraction.

Protein	Gel position	Coverage [%]	# PSMs	# Unique Peptides	Score Sequest HT
Cystatin B	CysB_1	12	8	1	24.9
Cystatin B	CysB_2	12	3	1	8.5
Cystatin B	CysB_3	12	1	1	2.6
Cystatin B	CysB_4	85	24	7	58.9
Cystatin B	CysB_5	58	7	3	16.1

Table 3.3: Identification of cystatin B in gel slices from AD and HC acid saliva pools.

		Coverage [%]		# PSMs		# Unique Peptides		Score Sequest HT		Abundance	
Protein	Gel position	AD	HC	AD	HC	AD	HC	AD	HC	AD	HC
Cystatin B	1	12	-	1	-	1	-	3.3	-	5.98E+04	-
Cystatin B	2	12	-	1	-	1	-	3.8	-	6.34E+04	-
Cystatin B	3	24	12	2	1	2	1	7.8	3.6	7.65E+05	1.84E+05
Cystatin B	4	46	24	4	2	3	2	11.8	7.6	1.49E+06	6.74E+05
Cystatin B	5	46	53	10	6	3	4	26.9	16.6	2.28E+07	7.36E+06
Cystatin B	6	46	46	7	3	3	3	22.3	11.4	1.10E+07	1.51E+06

Co-IP assay and characterization of cystatin B interactome

Cystatin B Co-IP assay was performed on whole salivary pools from AD and HC (Table 3.1) and the protein complex pull-down obtained was submitted to SDS-PAGE followed by western blot (Fig. 3.3). Before SDS-PAGE, the IP samples were divided in two aliquots per group, to be analyzed in NR- and R- conditions. In NR-conditions, it was possible to detect a signal higher than 250kDa, corresponding to the same signal observed in the preliminary tests by western blot (Fig. 3.1). The most intense signal in NR-condition was at 150kDa, corresponding to the cystatin B antibody used for the Co-IP assay, which is hiding, if present, the signal observed in the same molecular weight by western-blot of whole salivary pools from 24 HC subjects (HC) and 24 AD patients (Fig. 3.1, panel D). The same signal was undetectable in R-conditions cause the antibody itself was reduced to its heavy (50 kDa) and light (25 kDa) chains. In NR-condition of both AD and HC IPs it is still possible to observe the signal of cystatin B in its monomeric and dimeric form (10-15 and 25 kDa, respectively). Two more signals in NR-condition were detected at 100 and 75 kDa, which seemed to be the result of antibody trail and probably unspecific. Finally, in R-conditions, only three signals were detected: those of reduced antibody and the monomeric form of cystatin B between 10kDa and 15kDa, which was the most intense one.

To characterize cystatin B protein partners the Co-IP assays was performed again in triplicate followed by in gel tryptic digestion and HR-MS/MS analysis. Fig. 3.4 shows the SDS-PAGE gel obtained on IP samples from AD and HC pools in NR- and R- conditions. In both conditions each sample has been subjected to an extensive excision along each lane. Among stained protein bands, it is possible to recognize in the NR-condition lanes those corresponding to the intact form of cystatin B antibody (NR-, band 5, 150 kDa) which have been reduced to its heavy and light chains in R-condition (R-, band 7, 50 kDa and band 9, 25 kDa, respectively). All the stained bands were manually excised from the gel and transferred

to fresh tubes, as well as the zones of the lanes that did not appear stained for a total of 13 gel slices for each lane (Fig. 3.4).

When performing Co-IP with magnetic beads, due to their specific size and shape, sample pre-clearing step is unnecessary and, thus, we skipped it. Anyways, being aware of the high risk of detecting unspecific interactors and contaminants in this kind of experimental procedure, we decided to perform the characterization of background and to compare it with cystatin B Co-IPs' results. The background has been characterized pooling together the whole saliva of 26 HC subjects (NEG pool, Table 3.1) and performing the Co-IP in duplicate using the normal mouse IgG polyclonal antibody instead of cystatin B mouse monoclonal antibody. Normal mouse Co-IPs have been performed starting from the same amount of whole saliva with respect to cystatin B Co-IPs. The IP samples obtained from NEG were submitted to SDS-PAGE in R- conditions and 10 gels pieces were cut for the bottom-up analysis (Fig. 3.5).

Fig. 3.3: Western blot immunodetection with cystatin B antibody after Co-IP assay of whole saliva pools from 24 AD patients and 24 HC subjects (AD and HC pools, Table 3.1) in NR- and R- conditions.

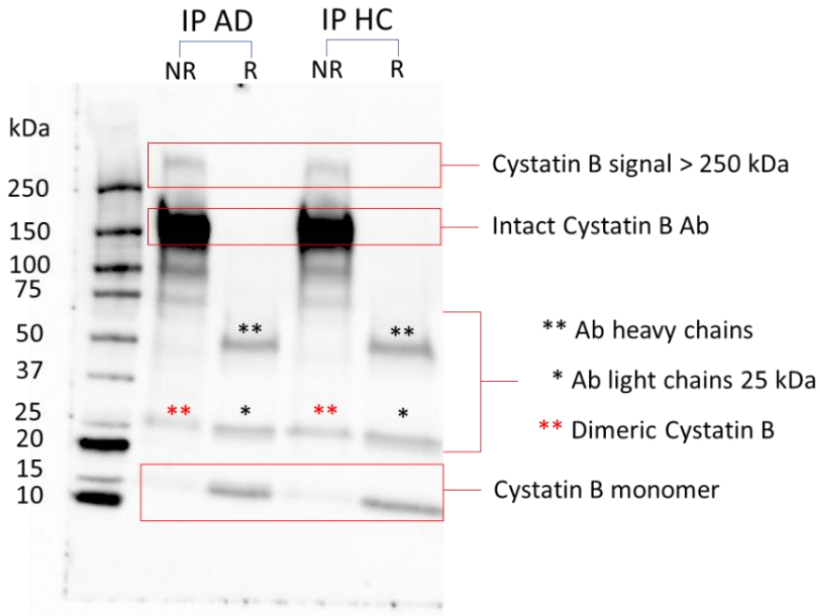


Fig. 3.4: SDS-PAGE in NR- and R- conditions of IP proteins after cystatin B Co-IP assay from AD and HC pools (Table 3.1) in triplicate. Black lines indicate how each lane, according to the condition, has been cut prior to tryptic digestion and nano-HPLC-HR-ESI-MS/MS analysis.

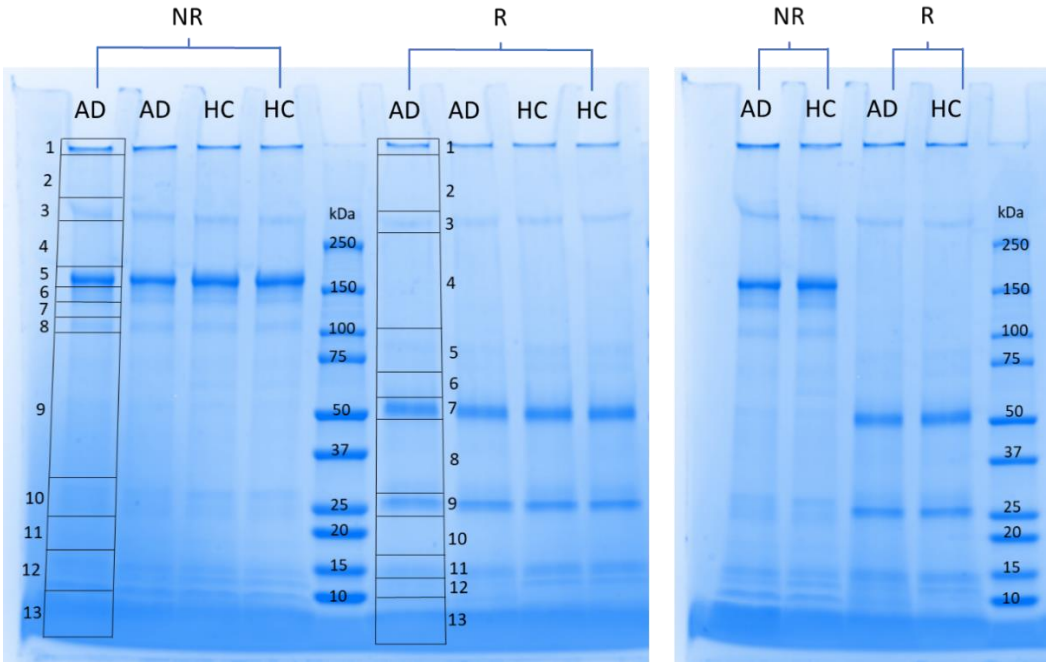
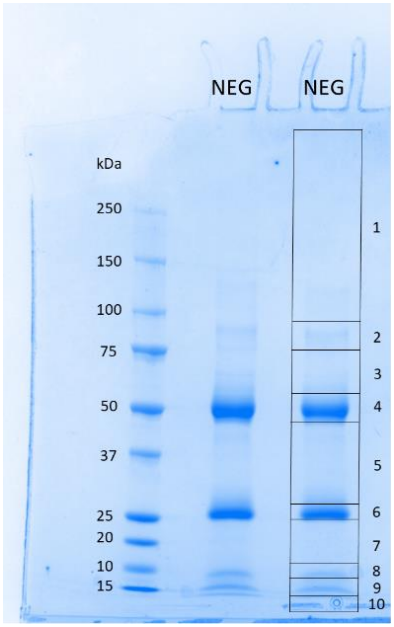


Fig. 3.5: SDS-PAGE in R- condition of IP proteins after normal mouse Co-IP assay from NEG pool in duplicate used to obtain a negative control and characterize the background. Black lines indicate how each lane has been cut prior to tryptic digestion and nano-HPLC-HR-ESI-MS/MS analysis.



PD software identified with high confidence and at least 2 unique peptides a total of 206 proteins in R-condition and a total of 254 proteins in NR-condition, which are common to both cystatin B Co-IPs from AD and HC in triplicate, and a total of 139 proteins in normal mouse CoIPs from NEG in duplicate. By excluding 74 contaminants proteins deriving from sample manipulation (*e.g.* skin proteins as keratins) and the sample itself (*e.g.* hemoglobin or immunoglobulins) (Supplementary Table S3.2), the number of identified proteins was reduced to 153 (R-condition) and 190 (NR-condition) in cystatin B IP from AD and HC and 86 (R-condition) in normal mouse IP from NEG (Supplementary Table S3.2).

The results obtained by PD have been exported to create a proper data matrix for Perseus statistics with the aim of detect which proteins were enriched in NEG or unchanged both between NEG and AD and between NEG and HC so we could mark them as unspecific interactors. Results of Student t-test were visualized by Volcano Plot as shown in Figure 3.6 where AD enriched proteins are represented by blue dots (panel A) and HC enriched proteins are represented by green dots (panel B), with respect to NEG ones. Grey dots in both panels represent unchanged proteins, while red dots are NEG enriched ones.

Protein partners of cystatin B were filtered out according to the following criteria. After the statistics based on LFQ abundances (R-condition), when the level of a protein resulted unchanged between NEG and HC group and between NEG and AD group or enriched in NEG, it was considered an unspecific interactor (Table S3.2). When a protein resulted unchanged with respect to NEG only in AD group or only in HC group, we did not exclude it to test differences between AD and HC in following analyses. Considering that the majority of proteins detectable after Co-IP assay can be contaminants and represent artificial interactors of the bait protein under study[184], we further filtered proteins using CRAPome database. All the proteins found in more than 25% of experiments reported in CRAPome database (threshold on the ratio found/total equal to 180/716) were excluded. The CRAPome score for

each protein identified in both R- and NR-condition, has been reported in Table S3.2 where proteins considered unspecific interactors are marked in red.

In total, we characterized 81 cystatin B interactors found in both AD Co-IP and HC Co-IP assays in R-condition. The interactors' list is reported in Table 3.4: for each protein it is indicated the UniProt-KB code, the molecular weight expressed in kDa, the LFQ abundance in log2 scale measured in each replicate and the values of fold change and -log10 *p* value calculated by Perseus in the comparisons of HC with respect to NEG and in AD with respect to NEG. In table S3.3 (Supplementary Material) it is reported the fold change and *p* value calculated by Perseus for those proteins which showed a good CRAPome score but were excluded because found enriched in NEG or unchanged both between NEG and HC and between NEG and AD.

A comparison between cystatin B and its 81 interactors identified in R-conditions and the list of proteins detected in NR-conditions after excluding all the proteins previously marked as unspecific interactors according to CRAPome database and NEG analysis (Supplementary Table S3.2) was performed. In NR-conditions the identification by a single unique peptide was only accepted for proteins previously characterized in R-conditions (Table S3.2). The overlapping between R- and NR- conditions is shown in Figure 3.7. More than 85% of proteins identified in R-conditions were also detected in NR-conditions (72 out of 82) which also represent more than 60% of all the proteins detected in NR-conditions (72 out of 113). The extra 41 proteins identified only in NR-conditions are available for consultation in supplementary Table S3.2. The reason why we considered this result only as a qualitative indication is because NR-condition is set in absence of reducing agents. Thus, covalent interactions between Co-IP proteins stayed intact prior SDS-PAGE resulting in a minor efficient gel-digestion and gel extraction of peptides. This is also the reason why NEG was

only performed in R-conditions, so that we could have direct high-quality evidence of unspecific interactions measured in the same experimental condition.

Moreover, the Venn analysis evidenced that 54 proteins of those identified as cystatin B interactors in R-conditions were detectable in the gel positions corresponding to a molecular range weight over 150 kDa in NR-conditions (gel position from 1 to 4, Fig. 3.8). These proteins are highlighted in bold in Table 3.4 and they represent proteins with a stronger interaction between each other being resistant to denaturation by 2% SDS when no reducing agent was added.

Fig. 3.6: Volcano Plots of significant enriched proteins in AD (blue dots) with respect to NEG (A) and significant enriched proteins in HC (green dots) with respect to NEG (B). In both panels red dots represent enriched proteins in NEG, while grey dots represent unchanged proteins.

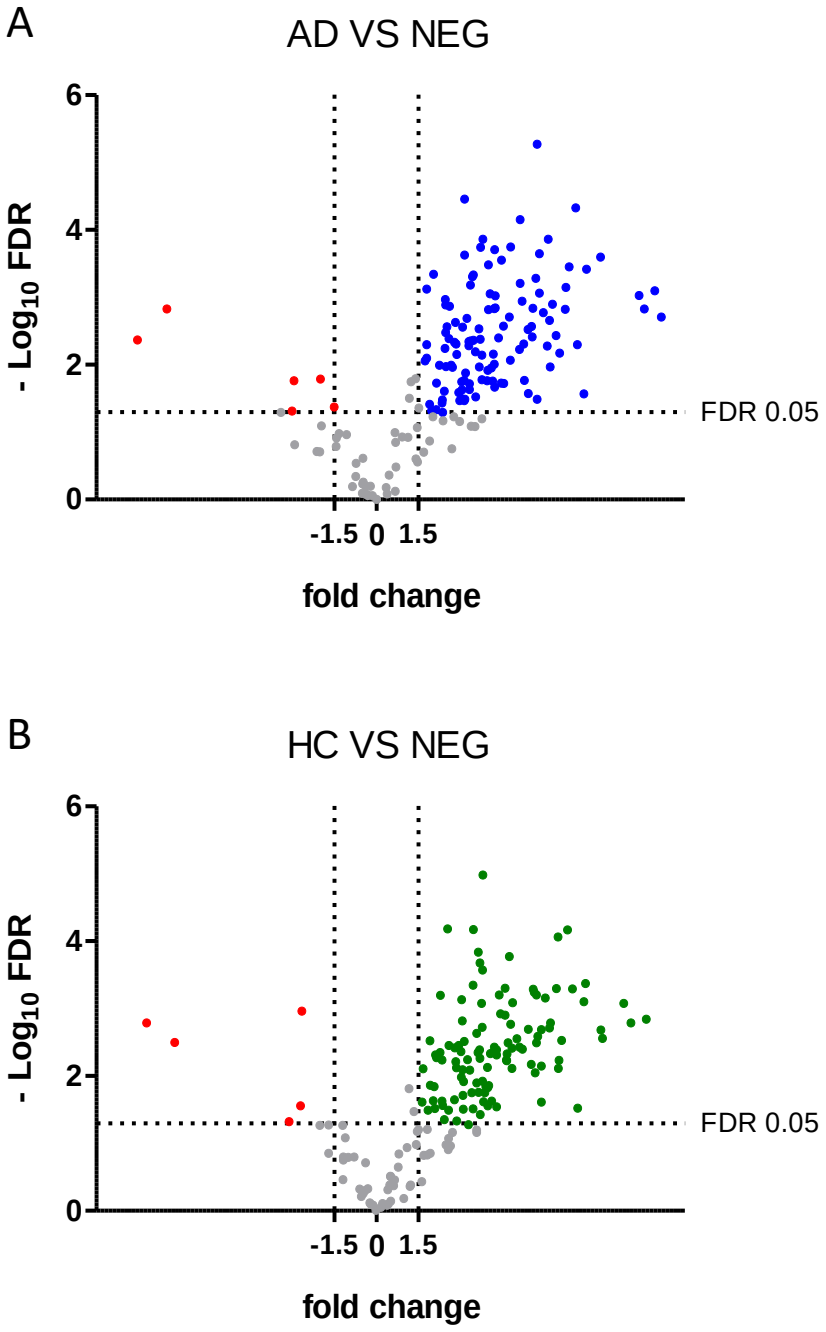


Table 3.4: Proteins found significantly enriched (↑) in AD group and HC group with respect to NEG which were identified as cystatin B interactors in R-condition. For each protein it is reported the UniProt-KB code, the LFQ abundance in log2 scale measured in each replicate, the *p* value and the fold change calculated by Perseus and the molecular weight (MW) expressed in kDa. In bold, proteins detected qualitatively also in NR-condition in both AD and HC Co-IP assay over the molecular weight of 150 kDa.

UniProt-KB Code	Protein name	Log2 LFQ abundances								AD VS NEG			HC VS NEG			MW
		AD			HC			NEG		-Log10 p value	Fold change		-Log10 p value	Fold change		
P04080	**Cystatin-B**	**26.6**	**25.9**	**25.6**	**26.4**	**24.8**	**24.0**	**19.3**	**19.0**	**3.0**	**9.4**	**↑AD**	**2.2**	**5.9**	**↑HC**	**11.1**
A8K2U0	**Alpha-2-macroglobulin-like protein 1**	**21.3**	**20.8**	**20.9**	**21.8**	**21.3**	**21.3**	**18.5**	**18.3**	**2.3**	**2.8**	**↑AD**	**3.1**	**3.0**	**↑HC**	**161.0**
O15144	**Actin-related protein 2/3 complex subunit 2**	**19.0**	**18.3**	**18.6**	**19.0**	**16.3**	**18.6**	**16.2**	**14.7**	**3.1**	**5.8**	**↑AD**	**0.9**	**2.6**	**-**	**34.3**
O15162	Phospholipid scramblase 1	18.2	19.0	18.6	19.1	18.1	18.7	15.8	15.8	2.0	4.1	↑AD	2.2	2.8	↑HC	35.0
O75131	Copine-3	17.7	17.4	17.6	18.7	18.5	18.2	16.9	15.4	1.1	1.4	-	1.6	2.4	↑HC	60.1
P01009	Alpha-1-antitrypsin	20.4	19.9	19.7	19.8	20.0	19.5	16.9	15.1	3.2	6.8	↑AD	1.9	3.8	↑HC	46.7
P01034	**Cystatin-C**	**21.1**	**22.1**	**21.5**	**21.6**	**20.3**	**20.6**	**19.3**	**19.4**	**3.3**	**3.4**	**↑AD**	**1.2**	**1.5**	**-**	**15.8**
P01037	**Cystatin-SN**	**24.3**	**24.4**	**24.7**	**24.0**	**22.9**	**24.2**	**20.2**	**19.8**	**3.1**	**9.9**	**↑AD**	**2.3**	**3.7**	**↑HC**	**16.4**
P02671	**Fibrinogen alpha chain**	**20.2**	**18.9**	**19.9**	**20.4**	**19.6**	**19.7**	**15.5**	**17.1**	**2.8**	**6.7**	**↑AD**	**1.9**	**3.6**	**↑HC**	**94.9**
P02675	**Fibrinogen beta chain**	**21.9**	**20.8**	**20.9**	**21.2**	**21.3**	**21.3**	**16.1**	**15.2**	**1.8**	**3.1**	**↑AD**	**3.2**	**5.6**	**↑HC**	**55.9**
P02679	**Fibrinogen gamma chain**	**22.4**	**22.0**	**22.0**	**22.9**	**22.8**	**22.4**	**19.9**	**19.4**	**1.7**	**4.2**	**↑AD**	**2.8**	**3.1**	**↑HC**	**51.5**
P04083	**Annexin A1**	**26.5**	**26.0**	**26.9**	**26.8**	**25.5**	**26.6**	**22.5**	**22.0**	**2.3**	**7.2**	**↑AD**	**2.3**	**4.1**	**↑HC**	**38.7**
P04792	**Heat shock protein beta-1**	**21.3**	**22.6**	**23.1**	**23.3**	**23.2**	**23.7**	**19.3**	**18.7**	**2.8**	**4.2**	**↑AD**	**3.2**	**4.4**	**↑HC**	**22.8**
P04899	**Guanine nucleotide-binding protein G(i) subunit alpha-2**	**20.3**	**20.1**	**21.0**	**21.0**	**20.2**	**20.9**	**16.1**	**16.5**	**2.3**	**6.1**	**↑AD**	**2.9**	**4.4**	**↑HC**	**40.4**
P05107	Integrin beta-2	21.8	20.2	21.1	21.5	21.4	21.3	17.0	14.5	3.1	4.1	↑AD	2.1	5.7	↑HC	84.7
P05109	**Protein S100-A8**	**26.6**	**26.4**	**27.0**	**26.9**	**26.4**	**27.3**	**24.7**	**25.1**	**3.5**	**6.9**	**↑AD**	**1.9**	**1.9**	**↑HC**	**10.8**
P05164	**Myeloperoxidase**	**26.3**	**26.6**	**26.3**	**26.2**	**26.3**	**26.2**	**24.9**	**24.2**	**3.6**	**4.5**	**↑AD**	**2.1**	**1.7**	**↑HC**	**83.8**
P06396	**Gelsolin**	**20.5**	**20.3**	**21.3**	**20.9**	**20.4**	**21.4**	**15.1**	**14.6**	**2.5**	**2.5**	**↑AD**	**3.2**	**6.0**	**↑HC**	**85.6**
P06702	**Protein S100-A9**	**25.7**	**25.0**	**25.8**	**25.8**	**24.8**	**24.5**	**22.5**	**22.9**	**4.5**	**3.2**	**↑AD**	**1.6**	**2.3**	**↑HC**	**13.2**

UniProt-KB Code	Protein name	Log2 LFQ abundances								AD VS NEG			HC VS NEG			MW
		AD			HC			NEG		-Log10 p value	Fold change		-Log10 p value	Fold change		
P06703	**Protein S100-A6**	**19.3**	**19.8**	**20.9**	**19.4**	**19.9**	**20.6**	**15.5**	**14.3**	**3.8**	**4.8**	**↑AD**	**2.4**	**5.1**	**↑HC**	**10.2**
P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	20.1	17.6	19.3	18.9	19.6	19.7	17.9	16.7	0.7	1.7	-	1.5	2.1	↑HC	76.7
P06744	**Glucose-6-phosphate isomerase**	**21.4**	**21.4**	**22.2**	**21.6**	**21.6**	**21.8**	**17.0**	**15.6**	**1.5**	**3.0**	**↑AD**	**2.7**	**5.4**	**↑HC**	**63.1**
P08246	**Neutrophil elastase**	**27.9**	**28.5**	**29.0**	**27.5**	**27.6**	**28.1**	**25.4**	**25.8**	**3.7**	**5.8**	**↑AD**	**2.3**	**2.1**	**↑HC**	**28.5**
P08311	**Cathepsin G**	**27.5**	**28.4**	**28.3**	**27.5**	**26.9**	**27.7**	**24.6**	**25.0**	**1.3**	**2.0**	**↑AD**	**2.5**	**2.6**	**↑HC**	**28.8**
P08493	**Matrix Gla protein**	**22.2**	**22.6**	**21.9**	**21.1**	**20.4**	**20.4**	**16.5**	**15.4**	**2.4**	**3.4**	**↑AD**	**2.5**	**4.7**	**↑HC**	**12.3**
P08758	Annexin A5	19.0	18.3	19.3	19.0	17.1	19.0	16.6	15.0	2.8	4.0	↑AD	1.1	2.6	-	35.9
P08962	CD63 antigen	20.8	18.3	19.4	20.0	18.7	19.2	15.3	14.9	2.8	4.2	↑AD	2.4	4.2	↑HC	25.6
P09211	**Glutathione S-transferase P**	**22.8**	**23.1**	**23.2**	**23.1**	**22.4**	**22.5**	**20.5**	**20.7**	**2.4**	**3.5**	**↑AD**	**2.3**	**2.1**	**↑HC**	**23.3**
P09228	Cystatin-SA	19.3	17.8	19.2	18.2	16.9	18.5	15.2	15.1	3.2	3.4	↑AD	1.7	2.8	↑HC	16.4
P0C0L4	Complement C4-A	19.4	17.6	17.7	18.9	18.9	18.7	16.2	14.8	1.2	2.7	-	2.1	3.3	↑HC	192.7
P0DTE7	**Alpha-amylase 1B**	**26.7**	**27.0**	**26.6**	**25.8**	**26.6**	**26.0**	**23.4**	**23.2**	**2.9**	**5.2**	**↑AD**	**2.4**	**2.8**	**↑HC**	**57.7**
P11215	Integrin alpha-M	22.0	19.9	20.4	21.9	21.7	20.9	17.8	16.7	2.6	4.5	↑AD	2.3	4.3	↑HC	127.1
P11413	**Glucose-6-phosphate 1-dehydrogenase**	**20.4**	**20.6**	**20.5**	**20.5**	**20.6**	**20.4**	**16.0**	**15.5**	**3.6**	**3.1**	**↑AD**	**3.8**	**4.7**	**↑HC**	**59.2**
P12273	**Prolactin-inducible protein**	**25.7**	**24.7**	**24.3**	**25.0**	**24.6**	**24.0**	**22.3**	**22.8**	**2.8**	**6.0**	**↑AD**	**1.8**	**2.1**	**↑HC**	**16.6**
P12429	Annexin A3	23.4	22.8	23.2	23.2	21.3	23.3	18.8	19.5	2.1	4.8	↑AD	1.5	3.5	↑HC	36.4
P13796	**Plastin-2**	**22.8**	**22.5**	**22.5**	**22.4**	**22.4**	**22.2**	**19.0**	**18.8**	**3.5**	**4.0**	**↑AD**	**4.2**	**3.5**	**↑HC**	**70.2**
P14780	Matrix metalloproteinase-9	19.3	17.9	18.3	19.1	19.8	18.8	16.2	14.5	2.6	2.5	↑AD	1.8	3.9	↑HC	78.4
P15104	Glutamine synthetase	18.5	17.3	18.5	18.5	15.3	17.7	16.5	15.4	2.5	5.4	↑AD	0.4	1.2	-	42.0
P17213	**Bactericidal permeability-increasing protein**	**24.7**	**24.9**	**24.7**	**23.4**	**23.9**	**23.4**	**21.6**	**21.6**	**5.3**	**5.7**	**↑AD**	**2.5**	**1.9**	**↑HC**	**53.9**
P17931	**Galectin-3**	**19.0**	**20.1**	**20.3**	**20.4**	**19.7**	**20.0**	**15.4**	**15.5**	**1.5**	**3.0**	**↑AD**	**3.3**	**4.6**	**↑HC**	**26.1**
P20292	Arachidonate 5-lipoxygenase-activating protein	21.3	21.7	21.9	22.6	22.3	22.1	16.0	15.7	3.3	3.5	↑AD	4.1	6.5	↑HC	18.1
P22079	**Lactoperoxidase**	**20.6**	**20.7**	**21.2**	**20.3**	**20.2**	**20.6**	**17.5**	**15.0**	**2.8**	**5.6**	**↑AD**	**1.6**	**4.1**	**↑HC**	**80.2**
P23280	**Carbonic anhydrase 6**	**24.0**	**24.4**	**24.1**	**23.2**	**22.9**	**23.9**	**21.5**	**21.9**	**2.3**	**5.3**	**↑AD**	**1.6**	**1.6**	**↑HC**	**35.3**
P24158	Myeloblastin	23.4	24.2	24.5	23.7	22.6	23.8	21.8	22.0	2.9	2.6	↑AD	1.2	1.5	-	27.8

UniProt-KB Code	Protein name	Log2 LFQ abundances								AD VS NEG			HC VS NEG			MW
		AD			HC			NEG		-Log10 p value	Fold change		-Log10 p value	Fold change		
P28325	**Cystatin-D**	**22.8**	**22.0**	**22.2**	**22.1**	**20.4**	**20.8**	**14.9**	**16.3**	**2.2**	**3.5**	**↑AD**	**2.2**	**5.5**	**↑HC**	**16.1**
P28676	Grancalcin	22.0	22.7	23.1	23.7	23.4	24.0	19.7	20.2	3.3	5.7	↑AD	3.1	3.8	↑HC	24.0
P29373	**Cellular retinoic acid-binding protein 2**	**20.7**	**20.1**	**20.6**	**20.3**	**19.6**	**19.3**	**15.5**	**14.3**	**2.3**	**3.3**	**↑AD**	**2.4**	**4.8**	**↑HC**	**15.7**
P29401	**Transketolase**	**20.9**	**20.8**	**21.4**	**20.9**	**20.8**	**21.5**	**15.7**	**15.0**	**2.8**	**9.6**	**↑AD**	**3.2**	**5.7**	**↑HC**	**67.8**
P30740	**Leukocyte elastase inhibitor**	**23.5**	**23.4**	**23.3**	**23.1**	**22.5**	**23.3**	**18.9**	**19.4**	**3.4**	**7.5**	**↑AD**	**2.7**	**3.8**	**↑HC**	**42.7**
P31146	**Coronin-1A**	**22.2**	**22.2**	**22.7**	**23.2**	**22.9**	**23.1**	**16.2**	**15.1**	**2.5**	**3.7**	**↑AD**	**3.4**	**7.5**	**↑HC**	**51.0**
P31930	Cytochrome b-c1 complex subunit 1	18.8	18.8	18.9	19.0	19.8	18.9	16.0	15.2	2.7	3.2	↑AD	2.3	3.6	↑HC	52.6
P32926	**Desmoglein-3**	**19.9**	**19.9**	**20.2**	**21.0**	**20.4**	**20.4**	**18.1**	**17.0**	**3.2**	**5.1**	**↑AD**	**2.1**	**3.1**	**↑HC**	**107.5**
P36952	**Serpin B5**	**20.0**	**19.4**	**20.7**	**20.6**	**19.6**	**20.1**	**17.0**	**15.2**	**3.0**	**2.5**	**↑AD**	**1.9**	**4.0**	**↑HC**	**42.1**
P41218	**Myeloid cell nuclear differentiation antigen**	**22.6**	**21.7**	**23.1**	**23.1**	**21.9**	**23.1**	**16.8**	**15.3**	**1.5**	**3.2**	**↑AD**	**2.5**	**6.6**	**↑HC**	**45.8**
P47929	**Galectin-7**	**20.2**	**20.4**	**20.2**	**21.7**	**20.9**	**20.1**	**15.4**	**15.0**	**4.3**	**7.1**	**↑AD**	**2.6**	**5.8**	**↑HC**	**15.1**
P48594	Serpin B4	21.0	20.6	21.4	21.2	20.7	21.1	15.5	14.1	2.0	2.7	↑AD	2.8	6.2	↑HC	44.8
P49913	Cathelicidin antimicrobial peptide	19.1	19.2	19.8	18.3	18.2	19.3	17.0	16.6	1.4	1.9	↑AD	1.5	1.8	↑HC	19.3
P50395	**Rab GDP dissociation inhibitor beta**	**19.2**	**18.8**	**19.7**	**19.4**	**18.6**	**19.4**	**16.1**	**16.2**	**1.6**	**5.4**	**↑AD**	**2.5**	**2.9**	**↑HC**	**50.6**
P50995	Annexin A11	18.7	18.9	19.4	19.0	19.8	19.6	16.9	15.2	2.7	6.2	↑AD	1.8	3.4	↑HC	54.4
P51159	Ras-related protein Rab-27A	19.2	19.5	20.4	20.7	19.4	20.4	16.1	14.9	3.7	4.2	↑AD	2.2	4.7	↑HC	24.9
P52209	**6-phosphogluconate dehydrogenase, decarboxylating**	**24.2**	**24.0**	**24.3**	**24.2**	**23.9**	**24.1**	**20.5**	**20.3**	**3.9**	**6.1**	**↑AD**	**3.7**	**3.7**	**↑HC**	**53.1**
P52566	**Rho GDP-dissociation inhibitor 2**	**20.9**	**19.6**	**20.6**	**20.0**	**18.8**	**19.6**	**15.5**	**14.2**	**1.5**	**2.3**	**↑AD**	**2.2**	**4.6**	**↑HC**	**23.0**
P54108	Cysteine-rich secretory protein 3	21.6	20.7	21.0	20.6	19.4	20.7	16.0	16.0	2.7	4.7	↑AD	2.4	4.3	↑HC	27.6
P59998	**Actin-related protein 2/3 complex subunit 4**	**18.9**	**19.4**	**18.8**	**19.6**	**19.8**	**17.1**	**16.0**	**14.6**	**3.6**	**8.0**	**↑AD**	**1.2**	**3.6**	**-**	**10.2**
P59665	Neutrophil defensin 1	27.3	27.6	28.1	27.6	28.6	28.7	26.3	26.3	1.8	1.4	-	1.6	2.0	↑HC	19.7
P60174	**Triosephosphate isomerase**	**17.8**	**18.2**	**18.9**	**19.7**	**19.3**	**20.1**	**17.4**	**17.5**	**0.9**	**0.9**	**-**	**2.4**	**2.3**	**↑HC**	**26.7**
P60763	Ras-related C3 botulinum toxin substrate 3	21.7	21.6	22.1	22.1	21.6	21.4	16.5	15.1	2.0	6.2	↑AD	2.7	5.9	↑HC	21.4
P61158	**Actin-related protein 3**	**18.3**	**18.5**	**18.7**	**18.7**	**17.4**	**18.9**	**16.3**	**14.6**	**2.9**	**2.5**	**↑AD**	**1.3**	**2.9**	**↑HC**	**47.3**
Q01469	**Fatty acid-binding protein 5**	**24.0**	**24.2**	**23.7**	**24.2**	**23.6**	**23.5**	**19.0**	**14.1**	**2.6**	**2.8**	**↑AD**	**1.5**	**7.2**	**↑HC**	**15.2**

UniProt-KB Code	Protein name	Log2 LFQ abundances								AD VS NEG			HC VS NEG			MW
		AD			HC			NEG		-Log10 p value	Fold change		-Log10 p value	Fold change		
Q08188	**Protein-glutamine gamma-glutamyltransferase E**	**26.0**	**24.3**	**24.9**	**25.4**	**24.7**	**25.1**	**23.0**	**22.5**	**2.6**	**3.1**	**↑AD**	**2.2**	**2.3**	**↑HC**	**76.6**
Q16610	**Extracellular matrix protein 1**	**19.8**	**20.2**	**20.8**	**20.7**	**20.0**	**20.2**	**16.1**	**15.3**	**1.8**	**3.8**	**↑AD**	**2.9**	**4.6**	**↑HC**	**60.6**
Q16851	UTP--glucose-1-phosphate uridylyltransferase	18.9	19.4	19.4	19.3	19.7	19.6	15.8	16.0	2.7	10.2	↑AD	3.8	3.6	↑HC	56.9
Q8N4F0	BPI fold-containing family B member 2	24.6	24.0	24.2	23.9	23.3	23.6	22.4	22.6	4.2	5.1	↑AD	1.8	1.2	-	49.1
Q8TAX7	**Mucin-7**	**25.0**	**24.2**	**25.0**	**23.6**	**23.7**	**23.4**	**17.0**	**16.5**	**3.1**	**1.8**	**↑AD**	**4.2**	**6.8**	**↑HC**	**39.1**
Q8TDL5	**BPI fold-containing family B member 1**	**23.8**	**24.0**	**24.0**	**24.3**	**25.0**	**25.3**	**22.0**	**21.8**	**3.0**	**4.2**	**↑AD**	**2.4**	**3.0**	**↑HC**	**52.4**
Q96DA0	**Zymogen granule protein 16 homolog B**	**27.0**	**26.4**	**26.9**	**26.3**	**24.9**	**26.1**	**24.0**	**24.7**	**3.9**	**3.8**	**↑AD**	**1.0**	**1.4**	**-**	**22.7**
Q96DR5	BPI fold-containing family A member 2	26.6	24.8	25.9	24.9	23.1	24.4	21.4	21.8	3.3	2.0	↑AD	1.5	2.6	↑HC	27.0
Q96HE7	**ERO1-like protein alpha**	**18.3**	**19.1**	**20.3**	**20.0**	**19.5**	**19.8**	**16.9**	**14.1**	**1.2**	**2.7**	**-**	**1.5**	**4.3**	**↑HC**	**54.4**
Q9HC84	**Mucin-5B**	**24.6**	**24.6**	**24.5**	**24.3**	**23.5**	**23.5**	**20.9**	**20.1**	**2.4**	**3.7**	**↑AD**	**2.2**	**3.2**	**↑HC**	**596.0**
Q9HDC9	Adipocyte plasma membrane-associated protein	20.8	19.4	20.3	20.6	18.9	20.4	16.9	15.1	2.9	6.3	↑AD	1.6	4.0	↑HC	46.5
Q9UBC9	**Small proline-rich protein 3**	**21.3**	**21.9**	**23.7**	**22.3**	**21.7**	**22.6**	**16.1**	**15.4**	**3.7**	**3.7**	**↑AD**	**3.3**	**6.4**	**↑HC**	**18.1**
Q9UBG3	**Cornulin**	**18.6**	**20.5**	**22.6**	**21.6**	**20.4**	**21.1**	**14.3**	**15.3**	**2.6**	**5.5**	**↑AD**	**2.7**	**6.2**	**↑HC**	**53.5**

Fig. 3.7: Venn diagram showing the overlap degree (green) obtained between proteins identified in AD and HC IP in R- (yellow) and NR- (sky-blue) conditions.

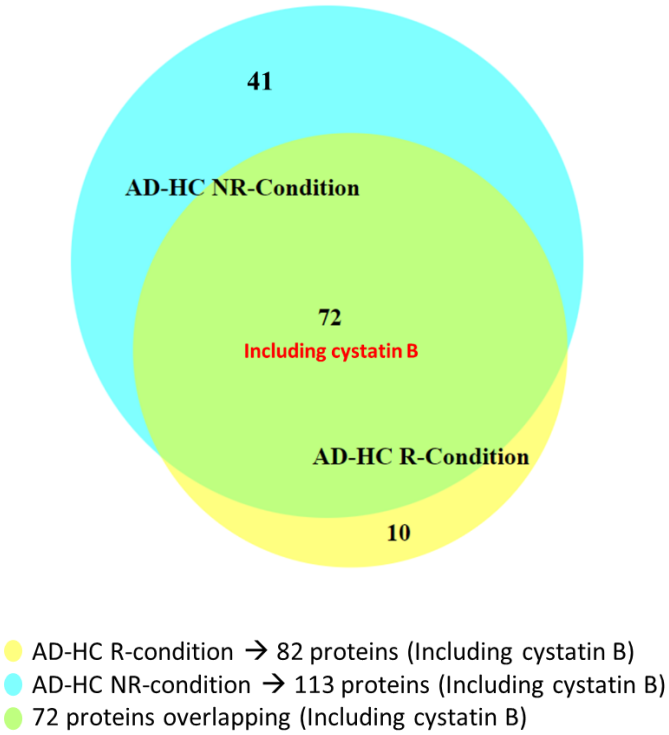
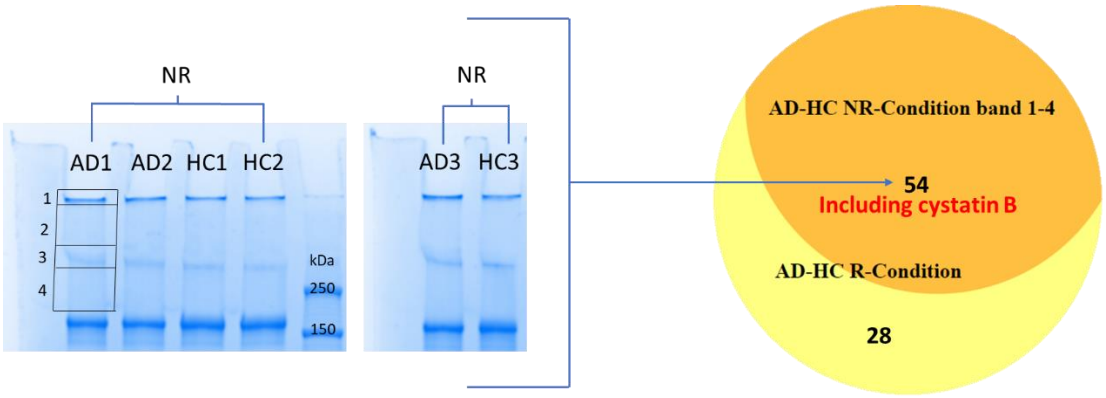


Fig. 3.8: Venn diagram showing the number of proteins among those identified in R-condition as cystatin B interactors were detectable in the molecular weight above 150 kDa in NR-conditions.



The list of identified cystatin B interactors in R-condition (Table 3.4) was submitted to STRING database. The known interactions among the proteins involved are represented in the network in Fig. 3.9 where proteins are identified by nodes marked with the proper UniProt-KB codes, while interactions are visualized by lines in cyan (if from curated databases) and/or magenta (if experimentally determined). The network shows the direct interaction between S100A8 and S100A9 proteins (UniProt-KB codes P05109 and P06702); MUC-7 and MUC-5 (UniProt-KB codes Q8TAX7 and Q9HC84); cathelicidin and myeloblastin (UniProt-KB codes P49913 and P49913). Another interacting cluster involves the actin-related protein 2/3 complex in the subunits 2, 3 and 4 (UniProt-KB codes O15144, P61158 and P59998, respectively); finally, two other interacting clusters involve 6-phosphogluconate dehydrogenase, glucose-6-phosphate 1-dehydrogenase, glucose-6-phosphate isomerase, transketolase and triosephosphate isomerase (UniProt-KB codes P52209, P11413, P06744, P29401, P60174, respectively); fibrinogen alpha, beta and gamma chains, integrin beta-2 and alpha-M, matrix metalloproteinase-9 and cathepsin G (UniProt-KB codes P02671, P02675, P02679, P05107, P11215, P14780, P08311, respectively).

Pathway analysis according to Reactome database was also performed by STRING on all cystatin B interactors (Table 3.4) revealing that half of proteins identified are involved in the immune system pathway, and specifically in the innate immune system response as shown in Table 3.5. The table reports the number of proteins counted in the submitted list with respect to the total number of proteins involved in each specific pathway according to the database (count in the network). For each pathway, identified by the specific Reactome code, the strength of result and FDR correction are also indicated. Even if involving a lower number of proteins, other two enriched pathways were found: neutrophil degranulation and the presence of antimicrobial peptides, both in strict connection to the immune response.

The functional enrichment analysis of cystatin B and its 81 interactors was performed by the functional annotation chart in DAVID tool, where a total of 112 enriched GO-terms were found (data not shown) for cellular component (CC), molecular function (MF) and biological process (BP). Among them, for each GO-term, we checked for the three most representative of cystatin B interactome. In Table 3.6, for each category, we reported the three most enriched GO-terms, the number and percentage of proteins associated among those listed in table 3.4, the *p*-value and the Benjamini post-test score. The analysis revealed that almost the total (more than 90%) of proteins identified are part of the extracellular compartment and that ca. 20% of them explicate a binding function calcium-dependent which can involve proteases. Finally, the biological process enrichment was in agreement with Reactome database outcome, since it evidenced roles related to the immune system as the defense against bacteria and fungi and the process of phagocytosis (Table 3.5 and 3.6). In Fig. S3.1 (supplementary section) it is shown a panoramic of all MF (panel A), CC (panel B) and BP (panel C) annotated GO-terms for cystatin B and all its interactors.

Fig. 3.9: Network of the known interactions (edges) occurring between proteins (nodes) identified as cystatin B interactors. Cyan lines represent interactions identified from curated databases, while magenta lines represent interactions experimentally determined. Each protein is identified by its UniProt-KB code.

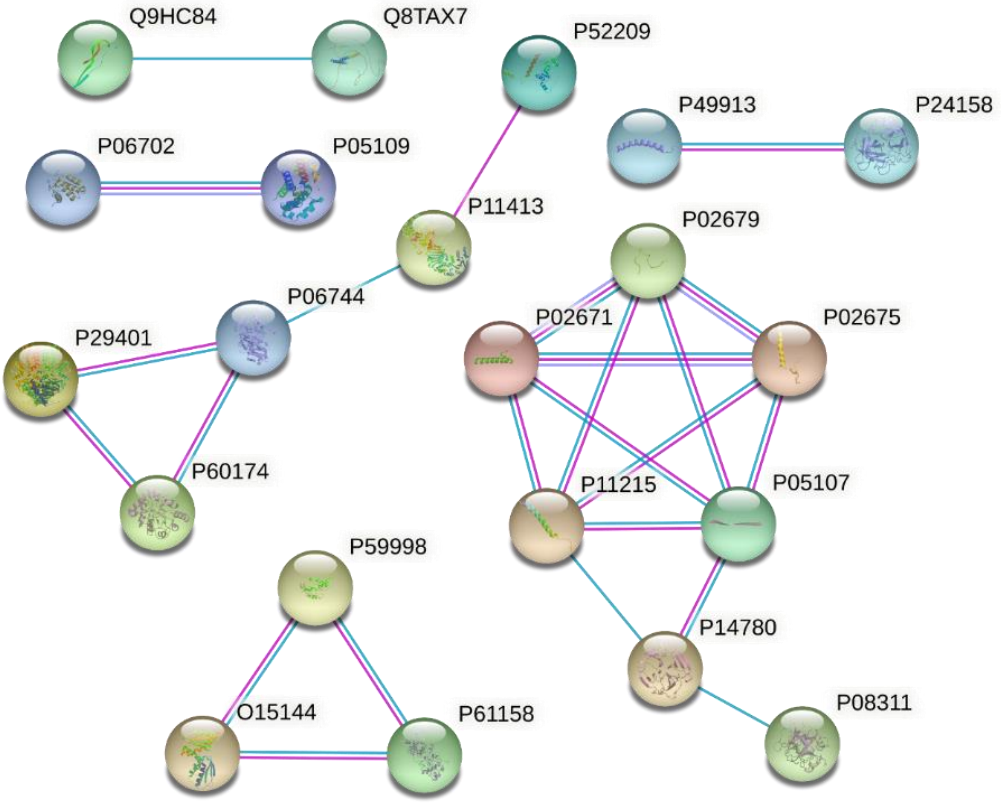


Table 3.5: Table result of pathway analysis performed by STRING and based on Reactome database on cystatin B and its 81 interactors. In the table the most representative outcomes of the network according to the number of proteins involved (count in network) are shown, together with the values of strength and FDR. The pathway code according to Reactome is also indicated.

Pathway	Reactome Pathways	Count in network	Strength	FDR
HSA-168256	Immune System	42 of 1956	0.71	2.95E-17
HSA-168249	Innate Immune System	40 of 1025	0.97	4.92E-25
HSA-6798695	Neutrophil degranulation	27 of 473	1.14	2.51E-19
HSA-6803157	Antimicrobial peptides	10 of 87	1.44	3.27E-08
HSA-6803157	Hemostasis	12 of 605	0.68	1.17E-02

Table 3.6: Three most significant GO-terms for cellular compartment (CC), molecular function (MF) and biological process (BP) categories obtained by DAVID tool through functional annotation chart analysis. For each term the *p*-value and Benjamini post-test score is reported. Count and percentage columns refers to the number of proteins found involved among those listed in Table 3.4.

Category	Term	Count	%	PValue	Benjamini
CC	GO:0070062 extracellular exosome	76 of 82	93.8	5.97E-55	8.78E-53
CC	GO:0005615 extracellular space	44 of 82	54.3	6.61E-28	4.86E-26
CC	GO:0005576 extracellular region	23 of 82	28.4	1.01E-06	4.97E-05
MF	GO:0002020 protease binding	9 of 82	11.1	1.86E-08	3.06E-06
MF	GO:0005509 calcium ion binding	15 of 82	18.5	3.84E-06	3.17E-04
MF	GO:0048306 calcium-dependent protein binding	6 of 82	7.4	6.57E-06	3.61E-04
BP	GO:0042742 defense response to bacterium	9 of 82	11.1	3.96E-07	2.40E-04
BP	GO:0006909 phagocytosis	6 of 82	7.4	2.96E-06	8.98E-04
BP	GO:0050832 defense response to fungus	5 of 82	6.2	7.33E-06	1.48E-03

After the characterization of cystatin B interactors found enriched in HC and AD with respect to NEG in R-condition (Table 3.4), we compared LFQ abundances of the same proteins between AD and HC groups. Firstly, we performed a quality check of the peptides extraction from each gel lane in AD and HC. In Figure 3.10 is reported the sample abundances chart prior (panel A) and after (panel B) the LFQ abundances normalization performed by PD software on the total peptide amount. Variances were comparable among each sample and further improved after normalization, so that any possible difference measurable between AD and HC subjects was not attributable to uneven extraction of peptides from SDS-PAGE gels. Perseus statistics was applied again and results of Student t-test were visualized by Volcano Plot as shown in Figure 3.11 where proteins considered significantly increased or decreased ($p < 0.05$ and s0 ±1) are marked in red. Specifically, triosephosphate isomerase (UniProt-KB code P60174) and grancalcin (UniProt-KB code P28676) resulted significantly decreased in AD group with respect to HC, while bactericidal permeability-increasing protein (UniProt-KB code P17213), matrix Gla protein (UniProt-KB code P08493) and MUC-7 (UniProt-KB Q8TAX7) resulted significantly increased in AD group with respect to HC. Perseus output matrix reporting *p*-values and fold change calculated after the Student t-test is shown in Table 3.7 for proteins resulted significantly variated in AD group with respect to HC. *p*-values and fold change of proteins resulted unmodified in the two groups are shown in Supplementary Table S3.4.

Fig. 3.10: PD sample abundances chart at protein level prior (panel A) and after (panel B) the normalization of LFQ abundances with respect to the total peptide amount in AD and HC groups (R-Condition).

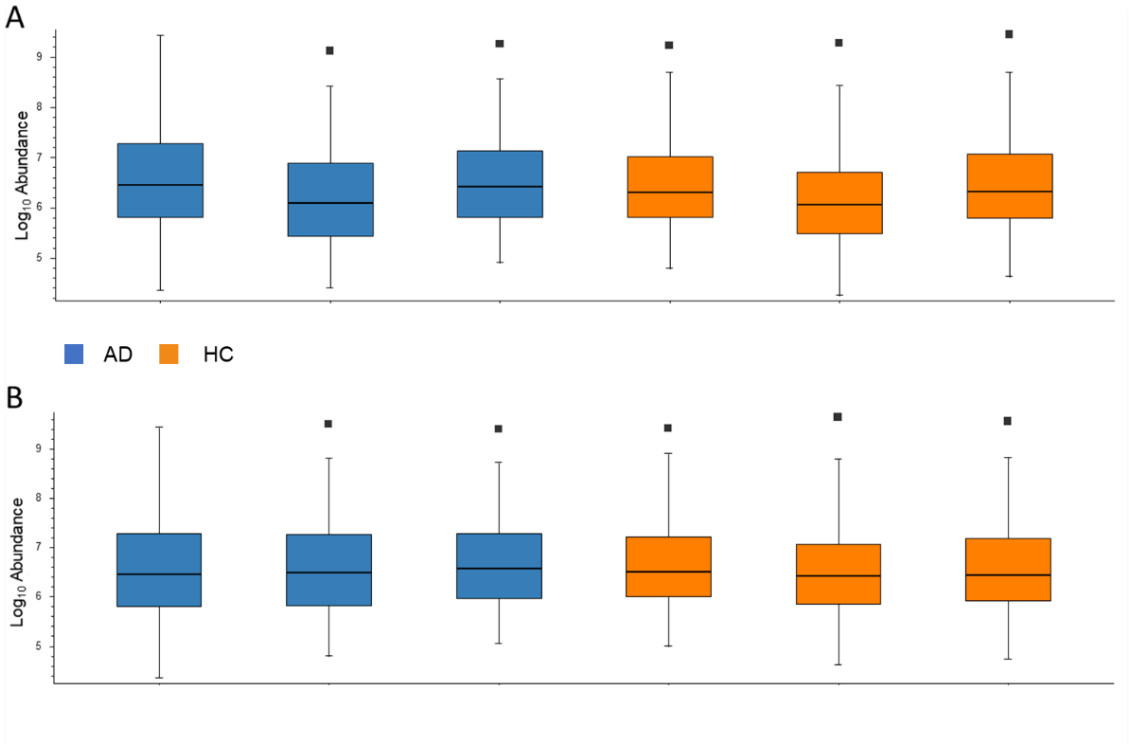
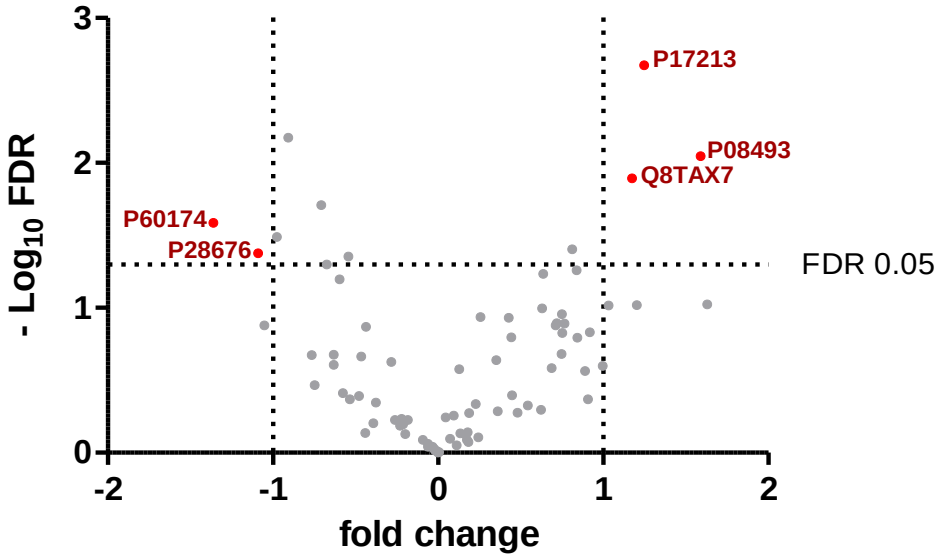


Fig. 3.11: Volcano Plot of significant increased (red dots on the right) or decreased (red dots on the left) proteins in AD with respect to HC group marked with their UniProt-KB codes.



Cystatin B was used as bait protein in the Co-IP assays. Its level was found unchanged between AD and HC with a *p*-value of 0.6 and a fold change of 1 (Table S3.4). We checked in PD software the quantification channels of cystatin B LFQ abundances in the triplicate of both groups when performing the Co-IP assays (Fig. 3.12) and we observed a grouped ratio AD/HC of 2.2 representing the mean of cystatin B LFQ measured in each group's replicate. This result agreed with Part I of the present thesis, where we showed how cystatin B levels in the acid soluble fraction of saliva from patients' group increased with respect to controls. Even if there was no significant change in the levels of cystatin B when performing the Co-IP assay of whole saliva pools from AD and HC subjects, considering the different ratio in the relative LFQ abundances observed, we performed again Perseus statistics to verify if there was any difference in the ratio bait/preys proteins between the patients and the controls. To do this, we calculated the ratio between LFQ abundances of cystatin B and the LFQ abundances of its 81 interactors in each replicate from AD and HC, then we performed again the statistical analysis with Perseus. Results of Student t-test are shown in Figure 3.13 where significantly variated proteins ($p < 0.05$ and s0 ±1) are marked in red. Specifically, three proteins resulted significantly decreased in their stoichiometric ratio with the bait protein in AD group: galectin-7 (UniProt-KB code P47929), copine-3 (UniProt-KB code O75131) and integrin alpha-M (UniProt-KB code P11215). Perseus output matrix reporting *p*-values and fold change calculated after the Student t-test is shown in Table 3.8 for proteins resulted significantly variated in AD group with respect to HC. *p*-values and fold change of proteins resulted unmodified in the two groups are shown in Supplementary Table S3.5.

Fig. 3.12: PD Quan channels of cystatin B showing the grouped relative abundance in AD and HC groups.

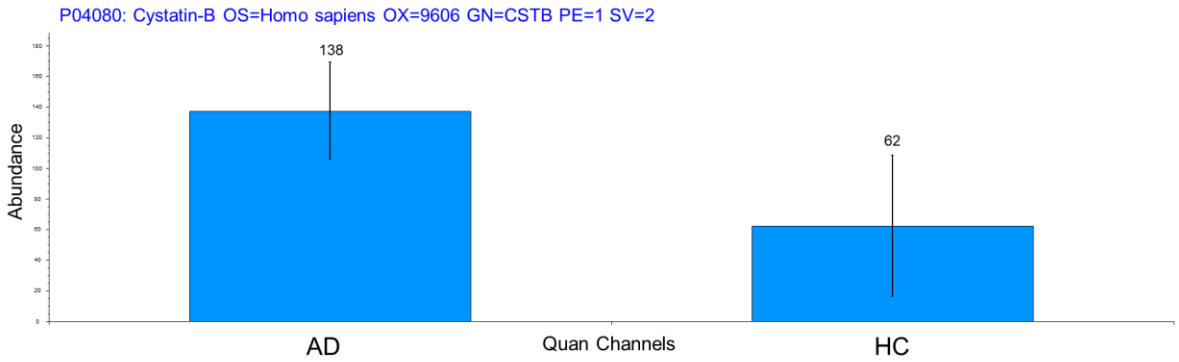


Fig. 3.13: Volcano Plot of significant decreased proteins (red dots on the left marked with relative UniProt-KB codes) in AD with respect to HC group when comparing the stoichiometric ratio of proteins with respect to cystatin B in the two groups.

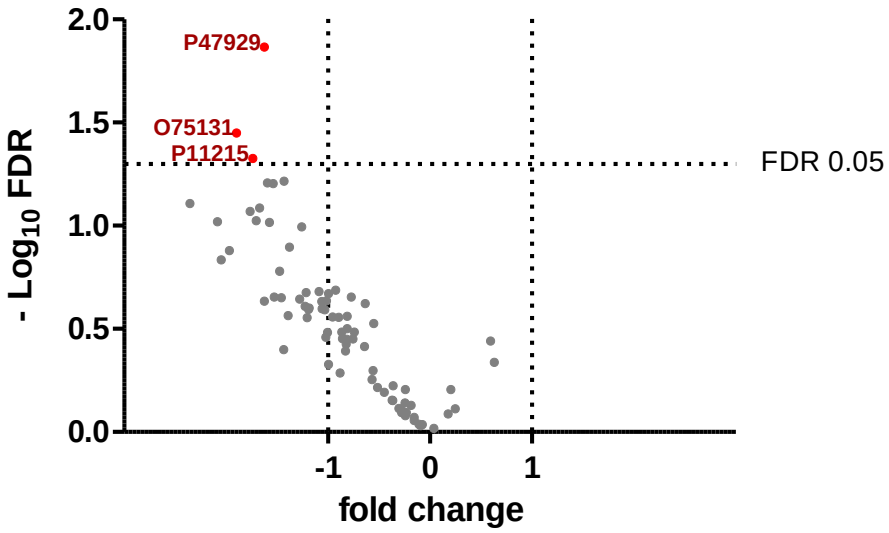


Table 3.7: Perseus matrix of proteins significantly changed between AD and HC groups after statistics of cystatin B and its 81 interactors. For each protein it is reported the UniProt-KB code, the LFQ abundance in log2 scale measured in each replicate, the *p*-value and the fold change calculated by Perseus.

		Log2 LFQ abundances						AD VS HC		
UniProt-KB Code	**Protein name**	**AD**			**HC**			**-Log10 p value**	**Fold change**	
P17213	Bactericidal permeability-increasing protein	24.7	24.9	24.7	23.3	23.8	23.3	2.7	1.2	↑AD
P08493	Matrix Gla protein	22.1	22.6	21.9	21.1	20.3	20.4	2.1	1.6	↑AD
Q8TAX7	Mucin-7	25.0	24.2	25.0	23.6	23.7	23.4	1.9	1.2	↑AD
P60174	Triosephosphate isomerase	17.8	18.2	18.9	19.6	19.3	20.1	1.6	-1.4	↓AD
P28676	Grancalcin	22.0	22.7	23.1	23.7	23.4	24.0	1.4	-1.1	↓AD

Table 3.8: Perseus matrix after statistics of stoichiometric ratios between the 81 cystatin B interactors and cystatin B itself between AD and HC groups. For each protein it is reported the UniProt-KB code, the LFQ abundance in log2 scale measured in each replicate, the *p*-value and the fold change calculated by Perseus.

		Log2 LFQ abundances						AD VS HC		
UniProt-KB Code	**Protein name**	**AD**			**HC**			**-Log10 p value**	**Fold change**	
P47929	Galectin-7	-6.3	-5.5	-5.4	-4.6	-3.9	-3.9	1.9	-1.6	↓AD
O75131	Copine-3	-8.8	-8.5	-8.0	-7.7	-6.2	-5.8	1.5	-1.9	↓AD
P11215	Integrin alpha-M	-4.6	-6.0	-5.2	-4.5	-3.0	-3.1	1.3	-1.7	↓AD

DISCUSSION

Salivary cystatin B is involved in a multiprotein complex

Cystatin B is a small protein of 98 amino acids with a single cysteine in position 3 of its sequence, whose crystal structure has been determined in complex with papain[185]. In human saliva from adult individuals (age range 20-50 years old), cystatin B has been mainly detected as S-derivatives, being the unique cysteine glutathionylated, cysteinylated or engaged in a disulfide bridge to form a protein homo-dimer, while the unmodified monomeric cystatin B resulted almost undetectable[186]. Non-covalent mechanisms of protein oligomerization have also been proposed *in vitro* in the model of domain-swapping for the dimeric form and the isomerization from *trans* to *cis* of a proline residue to form a cystatin B tetrameric structure[187]. Cipollini and colleagues studied the ability of cystatin B to form aggregates using cellular systems and they found that cystatin B has a polymeric structure prone to aggregate when overexpressed, but with a different mechanism compared to *in vitro* studies, considered that the formation of polymers seemed to be determined by the addition of a single cystatin B monomer at a time[175]. In the present study, we evidenced the presence of cystatin B in molecular weight ranges considerably higher than those expected for its monomeric or dimeric forms. To investigate the presence of cystatin B polymers or multiprotein complexes cystatin B-related in saliva, we performed Co-IP assays using cystatin B as bait protein and we characterized for the first time, in R-condition, 81 protein partners defining cystatin B interactome. Among these, 54 proteins including cystatin B where also characterized when working in NR-condition in the molecular weight range over 150 kDa, through SDS-PAGE followed by nanoHPLC-HR-MS/MS analyses. The signal of cystatin B was also specifically revealed by western blot in the range of the SDS-PAGE gel corresponding to more than 250 kDa, while in the natural occurring molecular weight ranges it was still possible to detect

cystatin B in its monomeric and dimeric form, evidencing how this protein exists in different conformations in human saliva. Our results showed that a subset of proteins found co-immunoprecipitated together with cystatin B were resistant to SDS 2% denaturation, while others showed a weaker interaction and were efficiently separated also in absence of a reducing agent. We did not interpret this result as the presence of covalent interactions involving all the 54 proteins interested, neither cystatin B could be the core protein structure of the covalent interactions of the multiprotein complex if we consider the presence of a unique cysteine in its sequence. We, instead, looked for known interactions among identified proteins using STRING database and we observed few proteins' clusters having different binding sites and interacting between each other (Fig. 3.9) which more likely might be able to aggregate with different protein partners, forming a core structure resistant to SDS denaturation that includes cystatin B. Indeed, with the exception of cathelicidin, myeloblastin, integrins alpha-M/beta-2 and matrix metalloproteinase-9, all the proteins shown in the network in Fig. 3.9 are part of the 54 proteins resistant to SDS denaturation. Our experimental design did not provide structural information about the multiprotein-complex, so we could not determine how proteins were directly or indirectly interacting between each other. Anyways, our results agreed with Cipollini and colleagues hypothesis: cystatin B is able to form polymeric structures SDS-resistant *in vivo* and those polymers are able to interact with different proteins[175]. In fact, we cannot exclude the presence of long cystatin B polymers in the medium/high molecular ranges of our assays (over 150 kDa), formed by the addition of one cystatin B monomer at the time as described by Cipollini *et al.*[175] or the result of the assembly of monomer and disulfide dimer proteoforms of cystatin B.

In this study, considered the limited volume of whole saliva available from AD patients, we performed the cystatin B Co-IP assay in a single pool from AD patients and a single pool from healthy subjects. Working on pools prevented the manifestation of individual differences, thus

the 81 cystatin B interactors characterized in reducing condition were always detectable in both groups and in each replicate. Due to its structure, cystatin B is able to interact and inhibit cathepsins B, L and H, the lysosomal cysteine proteases of the papain family[188]. None of them was found co-immunoprecipitated with cystatin B in the present study. The same evidence was found by Di Giaimo and colleagues when studying partners of cystatin B in rat cerebellum: through the two hybrid system they found that cystatin B was part of a tissue-specific multiprotein complex together with cytoplasmic proteins involved in the regulation of cytoskeletal functions but not with any protease[177]. This result agrees with some of the proteins identified in this study, like actin-related proteins (Arps). The presence of proteins capable of remodeling the actin cytoskeleton, like Arps[189], is also coherent with neutrophil degranulation pathway, which was found enriched by Reactome database[190,191]. Neutrophil granulocytes are the most abundant white blood cells in human circulation which show a key role in the host defense against invading pathogens and tissue injury. The reorganization of actin and cytoskeleton is required for degranulation of all types of granules and release of granule proteins to the phagosome or the extracellular milieu[192]. The translocation and exocytosis of granules in neutrophils require the increase in intracellular levels of Ca^{2+}, mediated by numerous target proteins which include Ca^{2+}-binding proteins as annexins and calmodulin[193]; interestingly, calcium ion binding and calcium-dependent protein binding were two of the three most enriched GO-terms for molecular function of cystatin B and its interactors, among which we identified different annexins (A1, A3, A5, A11), but also other proteins like grancalcin, gelsolin, plastin-2, S100A6, which can participate in the adhesion of neutrophils to fibronectin[194,195] and also play a role in the reorganization of actin cytoskeleton[196–198]. The involvement of proteins able to modulate the dynamics in the cytoskeleton is very interesting considering the results of the first and second Part of this thesis, particularly as far as Tβ4 is concerned. Indeed, Tβ4, which we found up-regulated in

saliva of patients with AD, is the main actin assembly modulator[105], and this lead to think that also the peptide could be involved in the cystatin B multi-protein complex identified in this study. Our experimental evidence did not allow either to confirm or deny its presence in the Co-IP of cystatin B, probably due to a technical limitation of our experimental protocol. However, the identification of proteins functionally correlated with Tβ4, which was candidate as AD salivary biomarker through the studies here presented, was amazing and suggested the need to further investigate.

Among the other proteins participating in the cystatin B interactome, we identified myeloperoxidase and CD63, which are specific markers of primary granules of the neutrophils[191]. These granules contain many other different antimicrobial proteins and peptides[191], some of which were identified as cystatin B interactors like elastase and cathepsin G (both serine proteases), bactericidal permeability-increasing protein, and α-defensin 1. α-defensins 1, 2, 3 and 4 were found increased in AD group with respect to elderly healthy subjects in our previous investigation of the acid soluble fraction of saliva, as described in the Part I of the thesis. Although the strong correlation of these anti-microbial peptides in human saliva, as demonstrated in Part II of the thesis, α-defensin 1 was the only identified in the multi-protein complex co-immuno-precipitated with cystatin B.

The results found on S100A8 and S100A9 participation in cystatin B interactome were interesting for two reasons: i) they are indicated by our studies described in Part I e Part II, and by other research groups[164], as biomarker candidates for diagnosis of AD and for classification; ii) they represent the main protein content of neutrophils[199], by which are constitutively expressed as a Ca^{2+} sensor, participating in cytoskeleton rearrangement and arachidonic acid metabolism[164]. During inflammation, S100A8 and S100A9 are released and exerts a critical role in modulating the inflammatory response by stimulating leukocyte recruitment and inducing cytokine secretion.

Other cystatin B interactors found in this study were integrin alpha-M/beta-2 complex, which can further interact with matrix metalloproteinase-9, that shows itself a role in the regulation of neutrophil migration[200], and fibrinogen alpha, beta and gamma chains. Integrin alpha-M/beta-2 complex can indeed promote neutrophil adhesion ligand *in vivo* and can interact both with fibrinogen and complement system[201,202]. Additionally, a cluster of proteins involved in metabolism was also part of the protein complex cystatin B-linked including 6-phosphogluconate dehydrogenase, glucose-6-phosphate 1-dehydrogenase, glucose-6-phosphate isomerase, transketolase and triosephosphate isomerase (TPI), actively involved in the pentose phosphate pathway and glycolysis[219]. Moreover, several protease inhibitors have been identified as cystatin B interactors, such as alpha-1-antitrypsin that is a SERPIN principally involved in the regulation of the protease activity in inflammatory response[203], serpin B4, which is a serine protease inhibitor[204], cystatin C, D, SA and SN, inhibitors of endogenous cathepsin[161], and extracellular matrix protein 1, which specifically inhibits the proteolytic activity of matrix metalloproteinase-9[205].

Differences between AD patients and HC subjects in salivary cystatin B interactome

Even if patients with AD and healthy controls exhibited the same protein pattern of the cystatin B interactome, some proteins were found differentially expressed between the two groups. TPI, which is a glycolytic enzyme that catalyzes the interconversion between glyceraldehyde 3-phosphate and dihydroxyacetonephosphate, resulted down-regulated in the patients. Inefficiency in glycolysis is a phenomena characteristic of AD pathology[206,207]. Aβ aggregates, typical of AD pathology, increase the levels of H_2O_2 and advanced glycation end products[91]. As a result of oxidative stress, irreversible nytrosilation of proteins in tyrosine residues can occur and one of the proteins found more nitrotyrosinated in AD is TPI, whose functional deficiency is associated with neurodegeneration[208,209]. Nitrotyrosinated TPI was

not measured in our study, but the reduced level in AD of the native enzyme could be the result of the increased levels of the oxidized form.

Also grancalcin was down-regulated in AD group. It is a Ca^{2+} binding protein specifically expressed in neutrophils, for which a functional role in the adhesion to fibronectin has been proposed via the functional interaction with integrins[195]. Neutrophil granules also contain bactericidal permeability-increasing protein[191], an antibacterial polypeptide cytotoxic to Gram-negative bacteria[210], which in this study was found up-regulated in AD patients. Mucin-7 showed significantly higher levels in cystatin B interactome of AD patients with respect to healthy controls. This is a small secreted MUC involved in the antimicrobial humoral immune response of the oral cavity, where it participate in the clearance of bacteria [211]. The increasing level of antimicrobial proteins/peptides in cystatin B interactome of AD patients with respect to HC observed in the present study is coherent with our previous findings in the acid soluble fraction of saliva[146], where we observed the establishment of several protective mechanisms in the AD patients against pathogen-targeting agents generating infections and inflammatory conditions. Finally, matrix Gla protein was found up-regulated in the cystatin B interactome of AD patients. To our knowledge, to date, there is no evidence of its implication in AD and the role this protein may exert in the interactome is unclear. Matrix Gla protein is a Ca^{2+} binding protein and a structural constituent of the extracellular matrix[212]. Recently, Freitas and colleagues evidenced that the study of modifications in the extracellular proteins of CNS with a focus on neurodegeneration disorders like Alzheimer, Parkinson and Huntington's diseases, can reveal the involvement of specific pathways and biological processes disease-related[213]. The functional enrichment analysis of cystatin B interactome revealed that more than 90% of proteins identified are part of the extracellular compartment, but the specific variations of extracellular proteins in the oral cavity of AD patients, which are independently expressed in saliva with respect to cystatin B interactome, require further investigations.

When measuring the ratio of cystatin B interactors and cystatin B itself in AD and HC groups, we found that three proteins were significantly down-regulated in the patients' interactome as a result of a different stoichiometric ratio with the bait protein. One is galectin-7, a beta-galactoside-binding protein implicated in modulating cell-cell and cell-matrix interactions[214]. The ability of galectin-7 to induce the expression of matrix metalloproteinase-9 has been observed in different kinds of cancer, including oral squamous cell carcinoma[215], but we did not observed such activity or interaction between the two proteins in our investigation, considering that the levels of matrix metalloproteinase-9 were found unchanged between AD and HC subjects. The other two are copine-3, a Ca^{2+} binding protein[216], and integrin alpha-M. Integrin alpha-M is a protein with Aβ binding activity, whose gene mutations have been related to activated microglia in transgenic mouse model and indicated as risk factors in response to Aβ deposition in AD[217]. Noteworthy, the lower level of stoichiometric ratio detected between integrin alpha-M and cystatin B in AD patients interactome appears specific, considering that no difference was measured in the levels of integrin beta-2, a direct protein partner of integrin alpha-M[202]. Furthermore, other proteins involved in AD and in Aβ binding as cystatin C[218], which was also characterized among cystatin B interactors in both AD and HC, did not show any difference in its level between the two groups. These results suggested that in saliva from AD patients, galectin-7, copine-3 and integrin alpha-M were less efficiently incorporated onto cystatin B interactome. If this was due to their lower expression in saliva of AD patients or if the degree of interactions is modified by a different arrangement of proteins within the complex, remains to be clarified.

To our knowledge, this is the first-time cystatin B interactome is investigated and described in saliva. Our study revealed that cystatin B exists in different forms and its prone to aggregate with other proteins in human saliva, more likely exhibiting a polymeric structure itself. Among cystatin B interactors, a strong evidence of proteins participating in neutrophil

composition and degranulation was observed. Moreover, the cystatin B interactome appeared similar between patients and controls except for the abundance of some partners, some of which were already found implicated in the disease, as TPI, while others have unknown function disease-related as matrix Gla protein. Among the future perspectives, it is of high relevance to demonstrate through a reverse Co-IP followed by western blot and/or HR-MS/MS analysis, the same interactions obtained when using cystatin B as bait protein. Finally, to better understand how proteins interact within each other from a stochiometric perspective and which functions this multiprotein complex could explicate in saliva, it might be helpful to isolate it in a native form and to submit it to structural proteomic investigations, as well as investigating the involvement of cystatin B and its partners in the formation of larger circulating immune-complexomes (CICs). Indeed, CICs are detectable in different body fluids, as serum and CSF[219,220], but also in saliva[221]. Immune-complexomes (ICs) form naturally *in vivo* to complete a humoral immune response of the body to an antigen and they are generally effectively removed by the system of mononuclear phagocytes[222]. ICs formation constantly occurs in healthy organisms: denatured proteins, antigens of bacteria, and dead cells undergo binding to natural antibodies followed by subsequent elimination by macrophages, but in some cases they can reflect immunological abnormalities caused by disease, especially for autoimmune diseases[223], also in the CNS[220]. Furthermore, because CICs contains many different antigens that could reflect a pathological condition and/or differences between individuals, it has been suggested that in some contexts their study might be more effective than the immunoblotting detection of individual antigens[223]. ICs in saliva have been recently described for the first time in the context of Sjögren's syndrome[221] where among the most representative antigens incorporated in CICs there where α-defensin 1, small proline-rich protein, myeloperoxidase, neutrophil elastase and cathepsin G. The authors evidenced the involvement of neutrophil degranulation, suggesting that a repeated destruction

of neutrophils due to abnormal autoimmunity could be in part responsible of the Sjögren's syndrome[221]. Considering half of the proteins found interacting with cystatin B were involved in the immune system pathway, with a clear participation of proteins deriving from neutrophils and neutrophil degranulation, their potential involvement in CICs in saliva of healthy subjects and/or AD patients stays open, and it will be the object of future investigations.

GENERAL CONCLUSIONS

The results obtained in the present thesis open several future perspectives on the study of the salivary protein profile of AD subjects and the application of proteomics studies to human saliva in general. In Part I and II few proteins and peptides, namely S100A8 and S100A9 proteins and their oxidized proteoforms, cystatin A and B, Tβ4 and α-defensins, appeared to be potential biomarker candidates for AD in the acid-soluble fraction of human saliva. In the future, it will be relevant to carry out salivary proteomic studies onto cohorts of AD patients at different disease severity status (mild, moderate, severe) to investigate the levels of these specific proteins and to individuate potential biomarkers useful for the disease early diagnosis. Additionally, the application of Random Forest multivariate analyses to larger groups including different neurological disorders will be tested for the classification of subjects according to their salivary protein profiles in a more complex statistical environment.

Finally, in the Part III of the present thesis, the participation of cystatin B into a multiprotein complex was highlighted with few differences at quantity level between AD patients and healthy subjects. In the future, the isolation of the multiprotein complex in its native form will help to study the physical connections between proteins and to individuate probable biological functions related to the complex and to cystatin B specifically.

PART III SUPPLEMENTARY SECTION

Table S3.1: Demographic data of 28 HC subjects and 28 AD patients included in the acid salivary pools used to perform western blot analysis of cystatin B signals. For each subject it is indicated the TPC of the acid-soluble fraction of saliva determined by BCA assay.

HC subjects	Sex and age	TPC μg/ml	AD Patients	Sex and age	TPC μg/ml
#1	M, 70	1310.0	#1	M, 82	1860.0
#2	M, 85	1443.0	#3	M, 85	1142.0
#3	F, 84	1011.0	#5	F, 63	1116.0
#4	M, 82	1226.0	#6	M, 78	1853.0
#5	F, 81	1580.0	#7	M, 85	2316.0
#6	F, 81	810.0	#8	F, 81	277.1
#7	F, 79	1772.0	#9	F, 78	564.7
#8	M,74	1050.0	#10	M, 85	580.3
#9	M, 71	1366.0	#11	F, 80	781.9
#11	M, 76	1382.0	#13	F, 79	1136.0
#12	F, 77	1828.0	#14	F, 82	2051.0
#13	M, 74	891.7	#16	F, 83	827.0
#14	M, 87	1376.0	#17	F, 63	442.7
#15	M, 73	638.6	#18	F, 80	756.0
#16	F, 81	1082.0	#19	F, 80	867.8
#17	F, 82	1108.0	#20	M, 87	302.4
#18	F, 72	1047.0	#21	M, 81	472.7
#19	F, 86	1115.0	#22	M, 87	484.0
#20	F, 73	2121.0	#23	F, 75	529.1
#22	F, 78	792.5	#25	F, 83	499.7
#24	F, 78	648.9	#26	F, 84	791.5
#25	F, 75	1636.0	#27	F, 81	873.5
#26	M, 75	897.8	#30	M, 86	721.2
#28	M, 78	625.8	#34	F, 77	863.9
#29	M, 73	655.5	#35	M, 87	1307.0
#30	F, 76	694.7	#36	M, 84	576.4
#31	F, 81	455.6	#37	F, 77	1400.0
#35	M, 80	722.2	#38	F, 78	801.8

S3.2: Proteins identified by PD software with high confidence and at least 2 unique peptides. The symbol "●" indicates in which group and condition the protein was identified; "●*" indicates proteins identified in AD and HC (NR-condition) by a unique peptide. At the bottom, the 74 proteins excluded as contaminants deriving from both sample manipulation and the sample itself. For each protein, the score from CRAPome is reported; "unmapped" refers to proteins not recognized by the database. Proteins marked in red represent proteins identified as unspecific interactors according to one or both following conditions: i) a CRAPome score ratio found/total over the 25% of total experiments deposited in the database; ii) enrichment (↑) in NEG or "unchanged" LFQ levels between NEG and AD and HC groups (R-Condition only). Except for proteins considered contaminants, the multiply-charged ions (m/z) used by PD for HR-MS/MS characterization are indicated.

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P31946	14-3-3 protein beta/alpha	●		●	368 / 716	533.58 (+3); 452.26 (+2); 454.26 (+2); 591.78 (+2)	↑AD, ↑HC
P61981	14-3-3 protein gamma	●			332 / 716	822.39 (+2); 412.89 (+3); 452.26 (+2); 454.26 (+2)	Unchanged
P31947	14-3-3 protein sigma	●		●	314 / 716	489.26 (+3); 531.58 (+3); 591.29 (+2); 452.26 (+2); 454.26 (+2); 535.77 (+2); 462.89 (+3)	↑AD, ↑HC
P63104	14-3-3 protein zeta/delta	●		●	410 / 716	774.86 (+2); 427.22 (+3); 652.84 (+2); 576.28 (+2); 454.26 (+2); 1020.99 (+2); 502.27 (+3)	↑AD, ↑HC
P08865	40S ribosomal protein SA	●			335 / 716	602.33 (+2); 456.78 (+2)	↑AD, ↑HC
P52209	6-phosphogluconate dehydrogenase, decarboxylating	●	●	●	122 / 716	714.39 (+2); 796.40 (+2); 596.79 (+2); 474.78 (+2); 676.29 (+2); 561.28 (+2); 568.77 (+2); 512.93 (+3); 464.59 (+3); 440.23 (+2)	↑AD, ↑HC
P62906	60S ribosomal protein L10a	●			280 / 716	495.60 (+3); 483.75 (+2)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
O15144	Actin-related protein 2/3 complex subunit 2	●		●	71 / 716	551.25 (+2); 515.31 (+2); 587.60 (+3)	↑AD
P59998	Actin-related protein 2/3 complex subunit 4	●		●	111 / 716	593.84 (+2); 505.28 (+2); 474.77 (+2)	↑AD
P61158	Actin-related protein 3	●		●	137 / 716	649.33 (+2); 622.30 (+2); 529.27 (+2); 474.24 (+2)	↑AD, ↑HC
P23526	Adenosylhomocysteinase	●		●	232 / 716	567.81 (+2); 628.85 (+2); 442.78 (+2)	Unchanged
Q01518	Adenylyl cyclase-associated protein 1	●		●	182 / 716	618.30 (+2); 404.90 (+3); 747.80 (+2); 442.75 (+2); 436.26 (+2)	↑AD, ↑HC
Q9HDC9	Adipocyte plasma membrane-associated protein	●		●	38 / 716	524.25 (+2); 555.29 (+2); 643.86 (+2); 549.78 (+2); 622.83 (+2); 456.75 (+2)	↑AD, ↑HC
P61204	ADP-ribosylation factor 3	●			261 / 716	552.31 (+2); 545.32 (+2)	↑AD, ↑HC
P05141	ADP/ATP translocase 2	●			525 / 716	610.30 (+2); 476.24 (+2); 610.34 (+2); 451.74 (+2)	↑AD, ↑HC
P02768	Albumin	●	●	●	239 / 716	547.32 (+3); 829.38 (+2); 637.65 (+3); 575.31 (+2); 722.32 (+2); 679.82 (+2); 569.75 (+2); 686.29 (+2); 492.75 (+2); 500.80 (+2); 440.72 (+2); 543.25 (+3); 386.72 (+2); 480.78 (+2); 467.26 (+2); 395.24 (+2); 470.73 (+2); 432.90 (+3); 725.77 (+2); 756.42 (+2); 464.25 (+2); 409.54 (+3); 507.30 (+2);376.90 (+3); 387.45 (+4); 476.22 (+2); 682.37 (+3)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P01009	Alpha-1-antitrypsin	●		●	2 / 716	504.75 (+2); 508.31 (+2); 917.46 (+2); 555.80 (+2)	↑AD, ↑HC
A8K2U0	Alpha-2-macroglobulin-like protein 1	●	●	●	12 / 716	530.31 (+2); 501.28 (+2); 726.85 (+2); 690.34 (+2); 590.80 (+2); 494.95 (+3); 539.78 (+2); 466.76 (+2); 690.38 (+2)	↑AD, ↑HC
P0DTE7	Alpha-amylase 1B	●	●	●	5 / 716	501.72 (+2); 544.98 (+3); 550.34 (+2); 724.32 (+3); 870.39 (+2); 862.40 (+3); 777.82 (+2); 524.23 (+3); 645.66 (+3); 645.81 (+2); 907.90 (+2); 472.24 (+3); 822.38 (+3); 476.57 (+3); 577.28 (+3); 768.33 (+3); 429.90 (+3); 863.90 (+2); 539.27 (+3); 500.24 (+2); 399.86 (+3); 442.73 (+2); 521.78 (+2), 644.35 (+2); 714.35 (+2); 580.60 (+3)	↑AD, ↑HC
P06733	Alpha-enolase	●	●	●	478 / 716	514.59(+3); 572.31 (+2); 817.41 (+2); 713.37 (+2); 450.28 (+2); 457.47 (+4); 778.89 (+2); 785.06 (+3); 770.89 (+2); 703.86 (+2); 760.42 (+2); 915.13 (+3); 381.88 (+3); 504.25 (+2)	↑AD, ↑HC
P04083	Annexin A1	●	●	●	68 / 716	775.91 (+2); 694.39 (+2); 515.29 (+3); 560.31 (+3); 689.97 (+3); 851.95 (+2); 786.06 (+3); 612.34 (+3); 870.37 (+2); 491.75 (+2); 452.89 (+3); 477.01 (+4) 631.80 (+2); 547.60 (+3); 386.56 (+3); 607.27 (+2); 457.93 (+3); 415.26 (+2); 540.75 (+2); 454.73 (+2); 402.74 (+2); 458.75 (+2); 532.75 (+2); 415.25 (+2);	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
						560.60 (+3); 573.27 (+3); 568.30 (+3); 382.21 (+2)	
P50995	Annexin A11	●		●	55 / 716	531.28 (+2); 526.76 (+2); 668.81 (+2); 483.74 (+2)	↑AD, ↑HC
P07355	Annexin A2	●	●	●	398 / 716	771.93 (+2); 615.64 (+3); 647.33 (+3); 556.28 (+2); 544.30 (+2); 564.79 (+2); 611.80 (+2); 622.81 (+2); 689.00 (+3); 711.35 (+2); 543.75 (+2); 954.94 (+2); 381.23 (+2); 556.61 (+3); 889.43 (+2); 451.89 (+3) 423.22 (+3)	↑AD, ↑HC
P12429	Annexin A3	●	●	●	3 / 716	611.81 (+2); 484.26 (+2); 571.93 (+3); 721.36 (+2); 675.83 (+2); 837.44 (+2); 546.31 (+2); 793.35 (+2); 739.88 (+2); 537.29 (+2); 472.23 (+2); 465.27 (+2); 509.77 (+2); 891.41 (+2); 489.77 (+2)	↑AD, ↑HC
P08758	Annexin A5	●			80 / 716	501.30 (+2); 507.76 (+2)	↑AD
P03973	Antileukoproteinase	●	●	●	4 / 716	692.83 (+2); 514.26 (+3); 555.72 (+2); 826.03 (+3); 540.57 (+3); 637.62 (+3); 524.61 (+3); 545.79 (+2); 984.40 (+3)	Unchanged
P02647	Apolipoprotein A-I	●	●		17 / 716	434.55 (+3); 405.88 (+3); 529.27 (+3); 693.86 (+2); 806.90 (+2); 700.84 (+2)	Unchanged
P20292	Arachidonate 5-lipoxygenase-activating protein	●		●*	0 / 716	599.81 (+2); 447.74 (+2); 552.28 (+2)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P25705	ATP synthase subunit alpha, mitochondrial	●		●	496 / 716	513.80 (+2); 438.75 (+2); 586.32 (+2); 518.58 (+3); 720.33 (+2); 570.98 (+3); 500.79 (+2)	↑AD, ↑HC
P06576	ATP synthase subunit beta, mitochondrial	●		●	427 / 716	488.28 (+2); 611.29 (+3); 718.38 (+2); 469.56 (+3); 544.82 (+2); 599.66 (+3); 701.35 (+2)	↑AD, ↑HC
P20160	Azurocidin	●	●	●	0 / 716	413.54 (+3); 753.40 (+2); 574.77 (+2); 978.81 (+3); 420.20 (+2); 744.88 (+2); 619.82 (+2); 807.88 (+4); 1024.81 (+3)	Unchanged
P17213	Bactericidal permeability-increasing protein	●	●	●	1 / 716	438.23 (+3); 775.89 (+2); 628.34 (+2); 497.27 (+2); 464.54 (+3); 995.52 (+2); 536.25 (+2); 521.82 (+2); 795.74 (+3); 478.27 (+2); 930.81 (+3); 464.54 (+3); 417.87 (+3)	↑AD, ↑HC
Q96DR5	BPI fold-containing family A member 2	●	●	●	0 / 716	666.38 (+2); 551.33 (+2); 567.83 (+2); 481.78 (+2); 556.86 (+2); 436.26 (+2); 445.74 (+2); 450.78 (+2); 643.02 (+3); 450.75 (+2)	↑AD, ↑HC
Q8TDL5	BPI fold-containing family B member 1	●	●	●	6 / 716	480.77 (+2); 815.95 (+2); 505.92 (+3); 662.34 (+2); 836.44 (+2); 410.91 (+3); 443.27 (+2); 654.00 (+3); 423.26 (+2); 401.24 (+3); 840.99 (+2); 589.82 (+2); 692.61 (+4); 544.30 (+3); 538.33 (+2); 615.86 (+2)	↑AD, ↑HC
Q8N4F0	BPI fold-containing family B member 2	●	●	●	0 / 716	659.70 (+3); 569.31 (+2); 568.32 (+3); 671.38 (+2); 387.74 (+2); 953.83 (+3);	↑AD

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
						474.29 (+2); 433.75 (+2); 650.85 (+2); 789.08 (+3); 399.24 (+2); 715.63 (+4)	
P12830	Cadherin-1	●	●	●	0 / 716	494.79 (+2); 549.24 (+2); 537.27 (+2)	Unchanged
Q14CN2	Calcium-activated chloride channel regulator 4	●			0 / 716	493.92 (+3); 480.28 (+2)	Unchanged
P23280	Carbonic anhydrase 6	●	●	●	1 / 716	582.32 (+2); 764.73 (+3); 788.40 (+2); 478.27 (+2); 548.79 (+2); 1146.59 (+2); 469.76 (+2)	↑AD, ↑HC
P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	●	●		0 / 716	733.35 (+2); 502.31 (+2)	↑HC
P49913	Cathelicidin antimicrobial peptide	●		●	0 / 716	528.57 (+3); 633.80 (+2)	↑AD, ↑HC
P08311	Cathepsin G	●	●	●	1 / 716	588.31 (+2); 780.92 (+2); 772.93 (+2); 481.94 (+3); 417.24 (+3); 560.81 (+2); 643.66 (+3); 552.81 (+2); 705.68 (+3); 633.82 (+4); 477.74 (+2); 429.91 (+3); 404.71 (+2); 572.98 (+3); 555.79 (+2); 490.29 (+2); 648.99 (+3)	↑AD, ↑HC
P08962	CD63 antigen	●		●*	0 / 716	506.23 (+3); 637.29 (+2)	↑AD, ↑HC
P29373	Cellular retinoic acid-binding protein 2	●		●	7 / 716	554.78 (+2); 426.73 (+2); 627.29 (+3); 456.61 (+3); 550.77 (+2); 475.23 (+2)	↑AD, ↑HC
P23528	Cofilin-1	●		●	496 / 716	572.81 (+2); 437.23 (+3); 669.32 (+2); 458.26 (+2)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P02747	Complement C1q subcomponent subunit C	●	●	●	0 / 716	629.35 (+2); 486.93 (+3); 964.45 (+2); 542.79 (+2); 822.05 (+3)	Unchanged
P01024	Complement C3	●	●	●	45 / 716	542.28 (+2); 576.81 (+2); 501.78 (+2); 595.81 (+2); 633.81 (+2); 612.80 (+2); 532.28 (+2); 401.25 (+2); 403.24 (+2); 542.96 (+3); 908.95 (+2); 756.41 (+2); 555.82 (+2)	Unchanged
P0C0L4	Complement C4-A	●		●	87 / 716	557.81 (+2); 771.41 (+2); 525.76 (+2); 638.32 (+2)	↑HC
O75131	Copine-3	●			91 / 716	560.81 (+2); 555.29 (+2)	↑HC
Q9UBG3	Cornulin	●		●	2 / 716	501.91 (+3); 585.28 (+3); 472.25 (+2); 684.12 (+4); 497.80 (+2)	↑AD, ↑HC
P31146	Coronin-1A	●		●	25 / 716	513.81 (+2); 569.75 (+2); 430.23 (+3); 480.76 (+2); 579.33 (+2); 447.75 (+2); 449.23 (+3); 569.75 (+2); 578.24 (+2); 751.86 (+2)	↑AD, ↑HC
P04080	Cystatin-B	●	●	●	164 / 716	663.86 (+2); 877.06 (+3); 474.91 (+3); 615.32 (+4)	↑AD, ↑HC
P01034	Cystatin-C	●	●	●	4 / 716	606.30 (+3); 613.81 (+2)	↑AD
P28325	Cystatin-D	●		●	0 / 716	713.37 (+2); 887.40 (+2); 425.75 (+2); 475.91 (+3)	↑AD, ↑HC
P09228	Cystatin-SA	●		●	1 / 716	846.91 (+2); 568.28 (+4); 714.67 (+3); 523.72 (+2); 440.90 (+3)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P01037	Cystatin-SN	●	●	●	2 / 716	691.68 (+3); 646.83 (+2); 709.34 (+3); 638.98 (+3); 568.28 (+4); 714.67 (+3); 957.97 (+2); 1037.02 (+2)	↑AD, ↑HC
P54108	Cysteine-rich secretory protein 3	●		●	0 / 716	1042.94 (+2); 596.25 (+2); 556.24 (+2)	↑AD, ↑HC
P31930	Cytochrome b-c1 complex subunit 1	●			157 / 716	529.79 (+2); 662.31 (+2)	↑AD, ↑HC
Q9UGM3	Deleted in malignant brain tumors 1 protein	●	●	●	9 / 716	1207.49 (+2); 796.33 (+3); 830.67 (+3); 710.83 (+2); 718.83 (+2); 503.29 (+2); 558.28 (+2); 521.24 (+3); 773.36 (+2); 796.36 (+2); 1002.14 (+3); 809.39 (+2); 672.97 (+3); 798.37 (+2); 756.35 (+2); 729.33 (+2); 427.22 (+3); 431.89 (+3); 730.38 (+2); 496.29 (+2); 538.28 (+2); 1235.90 (+3); 441.25 (+2); 396.73 (+2); 787.84 (+2) 805.33 (+3); 781.36 (+2); 837.04 (+3); 632.33 (+2); 515.92 (+3); 764.85 (+2)	Unchanged
P81605	Dermcidin	●	●	●	279 / 716	487.26 (+3); 581.28 (+2); 564.77 (+2)	Unchanged
Q02413	Desmoglein-1	●	●	●	116 / 716	1023.06 (+2); 867.93 (+2); 818.42 (+2); 614.30 (+2); 619.99 (+3); 763.71 (+3); 479.02 (+4); 669.67 (+3); 875.92 (+2); 554.93 (+3); 614.66 (+3); 545.28 (+2); 639.34 (+3); 769.04 (+3);	Unchanged

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
						715.61 (+3); 560.28 (+2); 644.67 (+3); 682.37 (+3)	
P32926	Desmoglein-3	●	●	●	0 / 716	647.34 (+2); 580.95 (+3); 641.38 (+3); 708.39 (+2); 777.45 (+2)	↑AD, ↑HC
P15924	Desmoplakin	●	●	●	328 / 716	565.31 (+2); 636.36 (+2); 694.36 (+2); 511.77 (+2); 522.79 (+2); 601.33 (+2); 570.34 (+2); 573.82 (+2); 587.82 (+2); 454.26 (+2); 425.72 (+2); 802.42 (+2); 752.69 (+3); 843.47 (+2); 756.92 (+2); 631.30 (+2); 554.62 (+3); 811.74 (+3); 469.92 (+3); 713.70 (+3); 708.37 (+3); 572.97 (+3); 531.61 (+3); 676.86 (+2); 599.27 (+2); 439.90 (+3); 619.66 (+3); 543.28 (+2); 513.74 (+2); 503.27 (+2); 627.80 (+2); 610.79 (+2); 579.82 (+2); 484.59 (+3); 432.55 (+3); 429.23 (+3); 491.60 (+3); 387.22 (+3); 469.54 (+3); 434.90 (+3); 515.28 (+3); 921.46 (+2); 396.72 (+4); 498.24 (+2); 580.75 (+2); 583.30 (+2); 494.28 (+2); 398.20 (+2); 511.29 (+3); 564.29 (+2); 707.86 (+2); 769.94 (+2); 529.81 (+2); 665.98 (+3); 728.36 (+3); 599.27 (+2); 502.27 (+2); 645.84 (+2)	Unchanged
P68104	Elongation factor 1-alpha 1	●	●	●	653 / 716	513.31 (+2); 438.92 (+3); 468.91 (+3); 457.79 (+2); 560.80 (+2); 488.28 (+2); 657.87 (+2); 435.77 (+2); 420.23 (+2);	Unchanged

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
						970.48 (+3); 975.82 (+3); 839.13 (+3); 1189.89 (+3)	
P26641	Elongation factor 1-gamma	●		●	461 / 716	414.55 (+3); 674.37 (+2); 524.95 (+3); 381.71 (+2)	Unchanged
P12724	Eosinophil cationic protein	●	●	●	0 / 716	677.37 (+2); 890.44 (+2); 572.27 (+2); 504.57 (+3); 677.37 (+2); 446.73 (+2)	Unchanged
Q96HE7	ERO1-like protein alpha	●		●	97 / 716	753.37 (+2); 645.82 (+2); 482.92 (+3); 565.33 (+2); 521.76 (+2); 573.24 (+2)	↑HC
Q16610	Extracellular matrix protein 1	●		●	14 / 716	813.91 (+2); 538.26 (+3); 562.81 (+2); 523.91 (+3); 659.34 (+2); 473.91 (+3); 540.33 (+2)	↑AD, ↑HC
P15311	Ezrin	●		●	264 / 716	591.80 (+2); 552.80 (+2); 457.77 (+2); 472.90 (+3); 494.26 (+2); 501.76 (+2); 488.78 (+2); 437.57 (+3); 447.78 (+2)	↑HC
P52907	F-actin-capping protein subunit alpha-1	●			318 / 716	574.94 (+3); 599.35 (+2); 853.41 (+2)	↑AD
P47756	F-actin-capping protein subunit beta	●		●	366 / 716	586.30 (+2); 677.32 (+2); 421.55 (+3)	↑AD, ↑HC
Q01469	Fatty acid-binding protein 5	●	●	●	106 / 716	522.28 (+2); 855.90 (+2); 464.28 (+2); 600.60 (+3); 636.30 (+2); 760.01 (+3); 802.71 (+3); 847.90 (+2); 812.04 (+3); 416.58 (+3)	↑AD, ↑HC
P02671	Fibrinogen alpha chain	●		●	2 / 716	514.76 (+2); 464.78 (+2); 503.91 (+3); 486.60 (+3)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P02675	Fibrinogen beta chain	●		●	15 / 716	601.05 (+4); 512.58 (+3); 490.72 (+2); 809.90 (+2); 802.34 (+2); 884.90 (+2); 474.73 (+2); 651.01 (+3); 654.81 (+2)	↑AD, ↑HC
P02679	Fibrinogen gamma chain	●	●	●	2 / 716	575.76 (+2); 597.75 (+2); 497.92 (+3); 559.27 (+2); 757.37 (+2); 431.92 (+3); 736.35 (+3); 561.65 (+3); 388.88 (+3); 440.72 (+2)	↑AD, ↑HC
P04075	Fructose-bisphosphate aldolase A	●	●	●	301 / 716	666.85 (+2); 496.94 (+3); 549.61 (+3); 566.79 (+2); 522.79 (+2); 564.62 (+3); 448.24 (+3); 382.24(+2); 708.37 (+3); 1019.82 (+3)	↑AD, ↑HC
P17931	Galectin-3	●		●	20 / 716	425.21 (+3); 552.99 (+3); 431.74 (+2); 413.22 (+4); 442.25 (+3)	↑AD, ↑HC
Q08380	Galectin-3-binding protein	●	●	●	102 / 716	590.27 (+3); 603.78 (2); 796.90 (+2); 678.39 (+2); 488.29 (+2); 442.89 (+3); 413.73 (+2); 377.89 (+3); 531.60 (+3)	Unchanged
P47929	Galectin-7	●		●	43 / 716	619.82 (+2); 467.23 (+3); 494.62 (+3); 429.23 (+2)	↑AD, ↑HC
P06396	Gelsolin	●		●	49 / 716	441.73 (+2); 660.35 (+2); 425.91 (+3); 378.24 (+2); 450.55 (+3)	↑AD, ↑HC
P11413	Glucose-6-phosphate 1-dehydrogenase	●		●	89 / 716	596.29 (+2); 421.54 (+3); 566.78 (+2); 458.25 (+2); 394.24 (+2); 583.26 (+3); 432.22 (+2); 380.21 (+3); 529.82 (+2); 574.80 (+2)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P06744	Glucose-6-phosphate isomerase	●		●	141 / 716	609.83 (+2); 828.41 (+2); 570.30 (+3); 522.29 (+2); 916.51 (+2); 403.23 (+2); 546.77 (+2); 496.76 (+2); 406.89 (+3); 534.95 (+3)	↑AD, ↑HC
P15104	Glutamine synthetase	●			58 / 716	441.23 (+2); 497.76 (+2)	↑AD
P09211	Glutathione S-transferase P	●	●	●	164 / 716	669.37 (+2); 568.79 (+2); 942.48 (+2); 376.22 (+2); 646.82 (+2); 709.39 (+3); 955.80 (+3)	↑AD, ↑HC
P04406	Glyceraldehyde-3-phosphate dehydrogenase	●	●	●	458 / 716	706.40 (+2); 917.46 (+2); 882.40 (+2); 454.88 (+3); 765.90 (+2); 403.22 (+2); 749.04 (+3); 406.21 (+2); 440.58 (+3); 807.45 (+2); 435.26 (+2); 580.65 (+3); 405.89 (+3); 611.98 (+3)	↑AD, ↑HC
P28676	Grancalcin	●	●	●	1 / 716	512.77 (+2); 452.94 (+3); 600.86 (+2); 435.23 (+2); 768.85 (+2); 580.01 (+4); 415.70 (+2)	↑AD, ↑HC
P04899	Guanine nucleotide-binding protein G(i) subunit alpha-2	●		●	159 / 716	529.32 (+2); 873.95 (+2); 546.24 (+3); 529.58 (+3)	↑AD, ↑HC
P00738	Haptoglobin	●	●		15 / 716	490.75 (+2); 673.33 (+2); 602.32 (+2); 460.73 (+2)	Unchanged
P0DMV9	Heat shock 70 kDa protein 1B	●	●	●	698 / 716	744.35 (+2); 559.25 (+3); 614.82 (+2); 599.35 (+2); 844.45 (+2); 489.27 (+3); 559.25 (+3); 555.29 (+2); 421.22 (+3); 755.72 (+3); 829.93 (+2); 994.49 (+3)	Unchanged

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P04792	Heat shock protein beta-1	●	●	●	128 / 716	953.50 (+2); 582.31 (+2); 595.31 (+3); 459.25 (+2); 548.61 (+3); 548.61 (+3); 490.25 (+3); 578.27 (+2); 382.88 (+3)	↑AD, ↑HC
P16402	Histone H1.3	●	●	●	600 / 716	423.26 (+2); 599.84 (+2); 554.29 (+2); 526.93 (+3); 420.92 (+3); 487.31 (+2)	↑AD, ↑HC
P16401	Histone H1.5	●	●	●	243 / 716	447.60 (+3); 547.28 (+2); 522.26 (+3); 420.92 (+3); 606.85 (+2)	↑AD, ↑HC
Q16778	Histone H2B type 2-E	●		●	513 / 716	487.94 (+3); 640.33 (+2); 390.20 (+3); 888.41 (+2); 576.28 (+2); 477.30 (+2); 503.62 (+3)	↑AD, ↑HC
Q5QNW6	Histone H2B type 2-F	●		●	515 / 716	487.94 (+3); 633.32 (+2); 390.20 (+3); 888.41 (+2); 569.28 (+2); 477.30 (+2); 503.62 (+3)	↑AD, ↑HC
P84243	Histone H3.3	●		●	384 / 716	668.35 (+2); 416.25 (+2); 417.57 (+3); 425.72 (+2); 394.74 (+2); 467.60 (+3)	↑AD, ↑HC
P62805	Histone H4	●	●	●	489 / 716	590.81 (+2); 526.63 (+3); 378.85 (+3); 494.94 (+3); 663.85 (+2); 442.59 (+3); 495.29 (+2); 485.60 (+3); 567.77 (+2)	↑AD, ↑HC
Q86YZ3	Hornerin	●	●	●	223 / 716	583.26 (+3); 643.62 (+3); 528.91 (+3); 480.55 (+3); 487.73 (+4); 773.66 (+3); 620.01 (+4); 509.56 (+3); 783.67 (+3); 482.97 (+4); 644.30 (+3)	Unchanged
P11215	Integrin alpha-M	●	●	●	0 / 716	606.84 (+2); 585.31 (+2); 805.39 (+2); 543.82 (+2); 681.86 (+2); 775.94 (+2); 497.24 (+2); 548.79 (+2); 676.32 (+2);	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
						520.32 (+2); 996.93 (+2); 720.45 (+2); 715.69 (+3); 480.92 (+3); 557.29 (+3); 783.06 (+3); 607.81 (+4); 568.61 (+3); 450.90 (+3); 860.45 (+2); 447.91 (+3); 442.28 (+3); 520.93 (+3); 399.98 (+4); 553.28 (+3); 982.04 (+2)	
P05107	Integrin beta-2	●		●	0 / 716	890.95 (+2); 579.30 (+2); 545.28 (+2); 556.26 (+2); 563.81 (+2); 407.26 (+2); 856.35 (+2); 494.94 (+3); 548.25 (+3); 485.26 (+2)	↑AD, ↑HC
P14923	Junction plakoglobin	●	●	●	244 / 716	671.37 (+2); 677.02 (+3); 618.84 (+2); 577.62 (+3); 544.33 (+2); 611.00 (+3); 676.81 (+2); 495.26 (+3); 719.05 (+3); 499.80 (+2); 714.41 (+2); 651.85 (+2); 456.88 (+3); 725.41 (+2); 406.78 (+2); 645.71 (+3); 501.78 (+2); 435.26 (+2); 438.75 (+2); 545.80 (+2); 800.40 (+2); 533.93 (+3); 783.11 (+3); 742.38 (+2)	Unchanged
Q9P0G3	Kallikrein-14	●	●		1 / 716	636.99 (+3); 565.80 (+2); 665.86 (+2); 687.34 (+3); 741.40 (+2)	Unchanged
P00338	L-lactate dehydrogenase A chain	●	●	●	304 / 716	567.78 (+2); 624.80 (+2); 632.84 (+2); 499.26 (+3); 457.29 (+2); 465.29 (+2); 559.79 (+2); 664.03 (+3)	↑AD, ↑HC
P22079	Lactoperoxidase	●		●	0 / 716	485.93 (+3); 561.27 (+2); 474.90 (+3); 618.81 (+2); 471.26 (+2); 570.84 (+2); 439.24 (+3)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P02788	Lactotransferrin	●	●	●	61 / 716	417.20 (+4); 426.71 (+2); 426.91 (+3); 429.23 (+2); 435.94 (+3); 465.24 (+3); 468.27 (+2); 483.91 (+3); 487.60 (+3); 490.26 (+3); 496.24 (+2); 506.95 (+3); 510.76 (+2); 525.28 (+3); 549.26 (+2); 555.65 (+3); 565.30 (+2); 575.82 (+2); 582.96 (+3); 598.31 (+2); 604.30 (+3); 619.31 (+2); 643.30 (+2); 668.35 (+2); 681.86 (+2); 695.87 (+2); 696.70 (+3); 697.34 (+2); 723.36 (+2); 753.36 (+3); 768.92 (+2); 790.38 (+2); 795.88 (+2); 807.90 (+2); 841.90 (+2); 904.40 (+2); 910.90 (+2); 931.88 (+2); 953.46 (+3); 462.71 (+2); 538.94 (+3); 538.96 (+3); 603.27 (+3); 604.66 (+3); 621.59 (+3); 659.29 (+3); 715.35 (+4); 720.38 (+3); 782.38 (+3); 846.46 (+3); 474.91 (+3); 497.94 (+3); 678.30 (+3); 691.02 (+3); 988.43 (+2); 994.97 (+2); 423.58 (+3); 1036.02 (+2)	Unchanged
P30740	Leukocyte elastase inhibitor	●	●	●	17 / 716	401.45 (+4); 435.90 (+3); 825.91 (+2); 551.80 (+2); 604.32 (+2); 609.81 (+2); 546.01 (+4); 503.75 (+2); 590.36 (+2); 429.90 (+3); 843.94 (+2); 575.30 (+2); 684.98 (+3); 376.22 (+2); 595.97 (+3)	↑AD, ↑HC
P61626	Lysozyme C	●	●	●	74 / 716	700.84 (+2); 690.34 (+2); 976.46 (+3); 506.73 (+2); 394.71 (+3); 640.81 (+3); 406.19 (+2); 788.37 (+3); 682.35 (+2)	Unchanged

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P08493	Matrix Gla protein	●		●	0 / 716	786.35 (+2); 638.33 (+2)	↑AD, ↑HC
P14780	Matrix metalloproteinase-9	●			2 / 716	542.77 (+2); 489.26 (+2)	↑AD, ↑HC
P26038	Moesin	●	●		230 / 716	591.80 (+2); 488.78 (+2); 494.26 (+2); 473.26 (+3); 552.79 (+2); 502.26 (+3)	Unchanged
Q9HC84	Mucin-5B	●	●	●	4 / 716	628.32 (+2); 840.40 (+2); 575.64 (+3); 774.36 (+2); 711.88 (+2); 955.84 (+2); 554.28 (+2); 629.86 (+2); 798.02 (+3); 754.87 (+2); 380.90 (+3); 705.00 (+3); 539.93 (+3); 565.94 (+3); 538.79 (+2); 616.63 (+3); 694.87 (+2); 967.15 (+3); 557.27 (+3); 662.29 (+2); 844.40 (+2); 663.33 (+2); 801.86 (+2); 574.75 (+2); 990.98 (+2); 690.79 (+2); 595.93 (+3); 402.53 (+3); 919.93 (+2); 500.24 (+3); 568.77 (+2); 467.21 (+3); 758.63 (+3); 713.82 (+2); 478.73 (+2); 521.60 (+3); 480.75 (+2); 518.80 (+2); 767.69 (+3); 874.36 (+2); 790.99 (+3); 566.75 (+2); 532.94 (+3); 475.22 (+3); 539.24 (+2); 720.31 (+3); 487.26 (+2); 386.21 (+3); 611.30 (+4); 416.24 (+2); 494.59 (+3); 466.21 (+2); 754.87 (+2); 554.27 (+2); 663.33 (+2); 774.36 (+2); 660.99 (+3); 629.86 (+2); 628.32 (+2); 380.90 (+3); 704.98 (+3); 690.79 (+2); 710.97 (+3); 531.47 (+4); 1004.48 (+3); 1065.95 (+2); 1079.97 (+2); 616.63 (+3); 1065.95 (+2)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
Q8TAX7	Mucin-7	●		●	0 / 716	530.79 (+2); 415.73 (+2); 407.73 (+2); 526.53 (+4)	↑AD, ↑HC
P24158	Myeloblastin	●	●	●	1 / 716	567.82 (+2); 430.93 (+3); 521.77 (+2); 645.89 (+2)	↑AD
P41218	Myeloid cell nuclear differentiation antigen	●		●	2 / 716	657.82 (+2); 383.20 (+3); 400.75 (+2); 475.27 (+2); 424.75 (+2); 715.33 (+3); 376.23 (+3); 970.44 (+2)	↑AD, ↑HC
P05164	Myeloperoxidase	●	●	●	3 / 716	679.81 (+2); 634.30 (+2); 396.56 (+3); 490.92 (+3); 593.32 (+2); 711.37 (+2); 548.83 (+2); 575.29 (+2); 576.81 (+2); 678.87 (+2); 640.88 (+2); 723.85 (+2); 986.52 (+2); 772.93 (+2); 876.76 (+3); 549.26 (+3); 469.74 (+2); 574.26 (+2); 456.74 (+2); 731.39 (+2); 710.05 (+3); 540.31 (+2); 509.26 (+3); 871.43 (+3); 658.01 (+3); 474.90 (+3); 444.25 (+2); 414.58 (+3); 429.2 (+3); 626.31 (+3); 503.59 (+3); 510.25 (+3)	↑AD, ↑HC
O14950	Myosin regulatory light chain 12B	●		●	295 / 716	630.80 (+2); 627.28 (+2); 614.81 (+2); 708.32 (+2)	↑AD, ↑HC
P35579	Myosin-9	●		●	448 / 716	385.89 (+3); 398.90 (+3); 408.55 (+3); 415.73 (+2); 420.58 (+3); 426.56 (+3); 440.25 (+3); 452.26 (+2); 459.24 (+2); 459.93 (+3); 462.75 (+2); 464.47 (+4); 480.91 (+3); 481.25 (+2); 481.25 (+3); 486.77 (+2); 489.76 (+2); 493.60 (+3); 493.79 (+2); 499.99 (+4); 509.26 (+2);	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
						520.29 (+3); 524.62 (+3); 531.29 (+2); 534.59 (+3); 546.78 (+2); 550.82 (+2); 554.94 (+3); 560.26 (+3); 572.91 (+3); 590.62 (+3); 597.31 (+2); 603.32 (+2); 605.97 (+3); 608.34 (+2); 610.83 (+2); 610.96 (+3); 623.62 (+3); 623.64 (+3); 629.34 (+2); 630.31 (+3); 633.76 (+2); 637.85 (+2); 638.63 (+3); 642.86 (+2); 653.34 (+2); 666.31 (+2); 765.89 (+2); 824.73 (+3); 827.40 (+2); 935.49 (+2); 973.51 (+2); 405.89 (+3); 407.56 (+3); 831.73 (+3); 791.34 (+2)	
P59665	Neutrophil defensin 1	●	●	●	21 / 716	381.54 (+3); 493.76 (+2); 559.2 (+2); 425.21 (+3); 637.31 (+2)	↑HC
P08246	Neutrophil elastase	●	●	●	2 / 716	1023.50 (+3); 747.36 (+4); 940.13 (+3); 909.46 (+2); 552.80 (+2); 533.31 (+2); 387.88 (+3); 474.75 (+2); 424.24 (+2); 705.35 (+4); 743.36 (+4); 464.61 (+3); 758.63 (+4); 608.99 (+3); 606.64 (+3); 496.28 (+3)	↑AD, ↑HC
O75594	Peptidoglycan recognition protein 1	●	●		0 / 716	778.42 (+2); 441.26 (+2); 555.96 (+3)	Unchanged
Q06830	Peroxiredoxin-1	●		●	549 / 716	554.30 (+2); 453.94 (+3); 598.99 (+3); 606.34 (+2); 590.79 (+2); 598.82 (+2); 460.76 (+2); 481.52 (+4); 447.72 (+2); 470.73 (+2)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P32119	Peroxiredoxin-2	●		●	482 / 716	512.27 (+2); 606.34 (+2); 489.77 (+2); 643.66 (+3); 431.76 (+2); 667.87 (+2); 393.88 (+3)	↑AD, ↑HC
Q13162	Peroxiredoxin-4	●		●	394 / 716	453.94 (+3); 613.35 (+2); 460.76 (+2); 606.82 (+2); 431.57 (+3)	↑AD, ↑HC
P30041	Peroxiredoxin-6	●		●	388 / 716	379.22 (+3); 453.74 (+2); 698.33 (+2); 791.84 (+2); 596.34 (+2); 543.30 (+2)	↑AD, ↑HC
P00558	Phosphoglycerate kinase 1	●		●	200 / 716	545.60 (+3); 377.88 (+3); 442.73 (+2); 590.34 (+3); 877.90 (+2); 483.24 (+2)	↑AD, ↑HC
O15162	Phospholipid scramblase 1	●			0 / 716	804.33 (+2); 567.78 (+2)	↑AD, ↑HC
Q13835	Plakophilin-1	●	●	●	41 / 716	762.90 (+2); 454.56 (+3); 450.74 (+2); 626.27 (+2); 647.33 (+2); 594.32 (+2); 917.45 (+2)	Unchanged
P13796	Plastin-2	●	●	●	167 / 716	568.31 (+2); 442.23 (+3); 759.88 (+2); 947.43 (+2); 506.81 (+2); 652.31 (+2); 793.43 (+2); 838.42 (+2); 513.26 (+3); 583.77 (+2); 559.29 (+2); 497.76 (+2); 499.76 (+2); 707.79 (+2); 535. 32 (+2)	↑AD, ↑HC
P01833	Polymeric immunoglobulin receptor	●	●	●	12 / 716	397.22 (+2); 438.72 (+2); 450.57 (+3); 452.24 (+2); 461.75 (+2); 470.78 (+2); 506.79 (+2); 522.74 (+2); 530.93 (+3); 540.28 (+2); 549.96 (+3); 584.30 (+3); 606.26 (+2); 614.88 (+2); 619.65 (+3); 621.64 (+3); 625.84 (+2); 675.99 (+3); 685.80 (+2); 701.86 (+2); 703.87 (+2); 724.33 (+3); 724.74 (+3); 747.73 (+3);	Unchanged

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
						760.39 (+2); 770.34 (+2); 829.41 (+2); 663.63 (+3); 681.63 (+3); 690.68 (+3); 702.35 (+3); 711.05 (+3); 845.39 (+2); 908.44 (+4); 469.58 (+3); 477.24 (+3); 853.90 (+2); 615.96 (+3); 477.50 (+4); 1021.94 (+2); 1210.91 (+3); 551.78 (+4)	
P07737	Profilin-1	●		●	308 / 716	690.36 (+2); 821.38 (+2); 735.88 (+2); 607.31 (+2); 412.24 (+3); 639.03 (+3); 822.47 (+2)	↑AD, ↑HC
P35232	Prohibitin	●			269 / 716	722.83 (+2); 575.30 (+2)	Unchanged
P12273	Prolactin-inducible protein	●	●	●	59 / 716	690.38 (+3); 642.39 (+2); 907.99 (+2); 678.78 (+2); 627.85 (+2); 665.68 (+3); 805.93 (+2); 513.80 (+2); 400.92 (+3); 706.82 (+2); 600.87 (+2)	↑AD, ↑HC
P60900	Proteasome subunit alpha type-6	●			193 / 716	643.36 (+2); 580.80 (+2)	↑AD, ↑HC
P07237	Protein disulfide-isomerase	●		●	247 / 716	484.57 (+3); 611.81 (+2)	↑AD, ↑HC
P06703	Protein S100-A6	●		●	72 / 716	458.25 (+2); 454.21 (+2)	↑AD, ↑HC
P05109	Protein S100-A8	●	●	●	79 / 716	711.36 (+2); 517.27 (+3); 424.90 (+3); 482.24 (+2); 440.23 (+2); 419.71 (+2); 432.23 (+2); 411.71 (+2); 686.35 (+2); 475.76 (+2)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P06702	Protein S100-A9	●	●	●	116 / 716	728.36 (+2); 503.29 (+2); 602.98 (+3); 586.94 (+3); 486.25 (+2); 489.01 (+4); 503.29 (+2); 439.24 (+2); 485.91 (+3)	↑AD, ↑HC
Q08188	Protein-glutamine gamma-glutamyltransferase E	●	●	●	47 / 716	727.86 (+2); 831.42 (+2); 546.81 (+2); 566.80 (+2); 644.00 (+3); 964.48 (+3); 598.29 (+2); 537.61 (+3); 470.28 (+2); 909.47 (+2); 692.98 (+3); 923.47 (+2); 457.88 (+3); 431.25 (+2); 656.33 (+2); 629.32 (+2); 672.30 (+3); 497.24 (+3); 612.31 (+3); 501.25 (+2); 424.24 (+2); 478.25 (+4); 375.20 (+2); 978.45 (+2); 520.84 (+2); 551.82 (+2); 928.96 (+2); 726.91 (+3); 453.89 (+3); 564.62 (+3); 1020.05 (+2); 815.04 (+3); 1013.99 (+2); 584.64 (+3); 986.98 (+2); 1088.52 (+2)	↑AD, ↑HC
P14618	Pyruvate kinase PKM	●	●	●	536 / 716	599.33 (+2); 680.35 (+2); 607.29 (+2); 571.31 (+2); 498.25 (+2); 398.55 (+3); 420.77 (+2); 611.32 (+2); 510.26 (+2); 607.29 (+2); 465.60 (+3); 613.31 (+3); 510.26 (+2); 890.44 (+2); 471.73 (+4); 725.71 (+3)	↑AD, ↑HC
P50395	Rab GDP dissociation inhibitor beta	●		●	166 / 716	515.26 (+2); 469.75 (+2); 546.90 (+3); 571.30 (+2); 393.89 (+3); 473.57 (+3)	↑AD, ↑HC
P60763	Ras-related C3 botulinum toxin substrate 3	●			42 / 716	620.80 (+2); 530.78 (+2); 475.75 (+2)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P51159	Ras-related protein Rab-27A	●			0 / 716	615.78 (+2); 532.29 (+2); 476.94 (+3); 471.75 (+2); 522.94 (+3)	↑AD, ↑HC
P52566	Rho GDP-dissociation inhibitor 2	●		●	4 / 716	656.36 (+2); 428.25 (+2)	↑AD, ↑HC
Q96P63	Serpin B12	●	●	●	49 / 716	838.45 (+2); 467.76 (+2); 549.80 (+2); 439.23 (+3); 577.30 (+2); 774.71 (+3)	Unchanged
P48594	Serpin B4	●			34 / 716	432.74 (+2); 431.26 (+2)	↑AD, ↑HC
P36;952	Serpin B5	●		●	3 / 716	611.28 (+2); 509.27 (+2); 490.75 (+2); 713.87 (+2); 551.32 (+2); 645.27 (+3); 531.26 (+3); 472.76 (+2)	↑AD, ↑HC
Q9UBC9	Small proline-rich protein 3	●		●	4 / 716	585.95 (+3); 430.75 (+2); 429.74 (+2); 413.74 (+2); 590.62 (+3); 656.82 (+4); 452.25 (+2)	↑AD, ↑HC
P29401	Transketolase	●		●	168 / 716	422.22 (+3); 473.26 (+2); 394.24 (+2); 469.74 (+2); 600.82 (+2); 471.94 (+3); 942.96 (+2); 458.28 (+2)	↑AD, ↑HC
P60174	Triosephosphate isomerase	●	●	●	177 / 716	486.91 (+3); 617.80 (+2); 425.56 (+3); 420.88 (+3); 801.95 (+2)	↑HC
P0DPH8	Tubulin alpha-3D chain	●			690 / 716	508.29 (+2); 573.63 (+3)	↑AD, ↑HC
Q16851	UTP--glucose-1-phosphate uridylyltransferase	●		●*	14 / 716	515.30 (+2); 498.81 (+2)	↑AD, ↑HC
P08670	Vimentin	●		●	543 / 716	547.27 (+2); 375.87 (+3); 568.94 (+3)	↑AD, ↑HC

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P21796	Voltage-dependent anion-selective channel protein 1	●			213 / 716	607.31 (+2); 587.31 (+2)	↑AD, ↑HC
P45880	Voltage-dependent anion-selective channel protein 2	●			256 / 716	470.73 (+2); 647.34 (+2)	↑AD, ↑HC
Q96DA0	Zymogen granule protein 16 homolog B	●	●	●	11 / 716	730.39 (+3); 649.63 (+3); 599.34 (+2); 1075.56 (+3); 1095.07 (+2); 460.23 (+2); 464.31 (+2); 468.23 (+2); 548.27 (+2); 412.70 (+2); 458.72 (+2)	↑AD
P30050	60S ribosomal protein L12		●		397 / 716	709.85 (+2); 441.28 (+2)	↑NEG
P60709	Actin, cytoplasmic 1		●		667 / 716	488.73 (+2); 400.24 (+3); 744.36 (+3); 581.31 (+2); 589.31 (+2); 566.77 (+2); 543.95 (+3); 850.73 (+3); 856.06 (+3); 506.24 (+3); 895.95 (+2);652.03 (+3); 547.28 (+3);507.74 (+2); 800.66 (+4)	↑NEG
P12814	Alpha-actinin-1		●	●	314 / 716	715.38 (+2); 769.39 (+2); 493.24 (+3); 431.92 (+3); 432.74 (+2); 587.80 (+2); 464.89 (+3)	Unchanged
P05089	Arginase-1		●	●	57 / 716	557.30 (+2); 429.22 (+3); 552.33 (+2)	↑NEG
Q9NP55	BPI fold-containing family A member 1		●		3 / 716	692-04 (+3); 740.94 (+2)	Unchanged
Q9NZT1	Calmodulin-like protein 5		●	●	150 / 716	755.04 (+3); 609.83 (+2); 465.75 (+2); 867.10 (+3); 817.90 (+2); 753.85 (+2)	Unchanged
P04040	Catalase		●	●	57 / 716	747.35 (+2); 487.75 (+2); 669.33 (+2); 741.37 (+2); 646.81 (+2)	Unchanged

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
Q08554	Desmocollin-1		●	●	71 / 716	561.94 (+3); 668.34 (+2); 542.74 (+2); 578.92 (+3); 493.58 (+3); 752.39 (+3); 396.89 (+3)	↑NEG
P13639	Elongation factor 2		●	●	488 / 716	562.29 (+2); 797.89 (+2); 554.32 (+2); 425.58 (+3); 485.28 (+2); 859.44 (+3)	Unchanged
Q5D862	Filaggrin-2		●	●	153 / 716	660.79 (+2); 784.67 (+3); 507.91 (+3); 635.36 (+2)	Unchanged
Q9H4G4	Golgi-associated plant pathogenesis-related protein 1		●		7 / 716	627.31 (+2); 727.35 (+2)	Unchanged
Q16777	Histone H2A type 2-C		●		552 / 716	472.77 (+2); 425.57 (+3); 644.39 (+3)	↑NEG
P23284	Peptidyl-prolyl cis-trans isomerase B		●		228 / 716	682.86 (+2); 524.28 (+2)	Unchanged
Q6P4A8	Phospholipase B-like 1		●		4 / 716	691.85 (+2); 573.32 (+2)	↑NEG
P28799	Progranulin		●		16 / 716	740.35 (+2); 631.66 (+3); 501.78 (+2)	↑NEG
P22735	Protein-glutamine gamma-glutamyltransferase K		●	●	35 / 716	700.35 (+2); 835.96 (+2); 597.79 (+2); 572.29 (+2); 424.56 (+3); 413.90 (+3)	Unchanged
P61160	Actin-related protein 2			●	140 / 716	675.81 (+2); 538.58 (+3); 685.37 (+2)	-
O15143	Actin-related protein 2/3 complex subunit 1B			●	44 / 716	497.77 (+2); 543.60 (+3)	-
P40394	All-trans-retinol dehydrogenase			●	6 / 716	529.76 (+2); 675.82 (+2); 452.26 (+2)	-

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
O43707	Alpha-actinin-4			●	335 / 716	516.93 (+3); 715.39 (+2); 769.39 (+2); 433.89 (+3); 432.74 (+2); 587.80 (+2); 483.91 (+3); 464.89 (+3)	-
Q9UBD6	Ammonium transporter Rh type C			●	0 / 716	874.45 (+3); 437.73 (+2)	-
P07384	Calpain-1 catalytic subunit			●	30 / 716	642.81 (+2); 411.89 (+3); 503.78 (+2)	-
P16152	Carbonyl reductase			●	143 / 716	422.77 (+2); 603.83 (+2)	-
P31944	Caspase-14			●	62 / 716	602.80 (+2); 579.81 (+2)	-
O00299	Chloride intracellular channel protein 1			●	147 / 716	641.34 (+2); 539.76 (+2); 479.24 (+2); 664.82 (+2)	-
Q15517	Corneodesmosin			●	50 / 716	827.92 (+2); 741.39 (+3)	-
P01040	Cystatin-A			●	57 / 716	548.61 (+3); 988.50 (+2); 678.86 (+2)	-
Q14574	Desmocollin-3			●	15 / 716	553.30 (+2); 564.97 (+3); 599.99 (+3)	-
P61803	DAD1			●	27 / 716	523.79 (+2); 560.28 (+2)	-
Q6ZVX7	F-box only protein 50			●	21 / 716	386.54 (+3); 616.32 (+2)	-
P21333	Filamin-A			●	508 / 716	484.26 (+3); 613.29 (+2); 455.57 (+3); 549.63 (+3); 595.62 (+3); 550.29 (+2)	-
P06737	Glycogen phosphorylase, liver form			●	74 / 716	432.21 (+3); 527.29 (+2); 458.74 (+2)	-
P62826	GTP-binding nuclear protein Ran			●	440 / 716	607.81 (+2); 508.29 (+2); 622.82 (+2)	-
Q6ZN66	Guanylate-binding protein 6			●	1 / 716	474.75 (+2); 429.90 (+3)	-

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P11142	Heat shock cognate 71 kDa protein			●	703 / 716	596.67 (+3); 600.34 (+2); 744.35 (+2); 614.82 (+2); 564.58 (+3); 494.61 (+3)	-
P07900	Heat shock protein HSP 90-alpha			●	565 / 716	621.86 (+2); 618.30 (+2); 757.40 (+2); 612.82 (+2); 576.28 (+2)	-
P08238	Heat shock protein HSP 90-beta			●	573 / 716	621.86 (+2); 625.31 (+2); 757.40 (+2); 580.80 (+2); 576.28 (+2)	-
P52790	Hexokinase-3			●	23 / 716	525.80 (+2); 422.89 (+3); 531.29 (+2)	-
Q9UHA7	Interleukin-36 alpha			●	1 / 716	682.36 (+2); 549.81 (+2)	-
O75874	Isocitrate dehydrogenase			●	62 / 716	719.86 (+2); 544.29 (+2)	-
Q9UKR3	Kallikrein-13			●	0 / 716	608.83 (+2); 488.25 (+3); 398.23 (+3)	-
P07195	L-lactate dehydrogenase B chain			●	340 / 716	624.80 (+2); 480.28 (+2); 457.29 (+2); 504.26 (+3)	-
P09960	Leukotriene A-4 hydrolase			●	66 / 716	604.33 (+2); 524.93 (+3); 769.34 (+2); 664.85 (+2)	-
P40926	Malate dehydrogenase, mitochondrial			●	198 / 716	383.22 (+3); 669.86 (+2); 496.77 (+2); 735.85 (+2); 577.28 (+2)	-
P60660	Myosin light polypeptide 6			●	411 / 716	596.28 (+3); 677.87 (+2); 580.29 (+3)	-
P43490	Nicotinamide phosphoribosyltransferase			●	135 / 716	395.55 (+3); 542.81 (+2); 386.25 (+2)	-
Q6UX06	Olfactomedin-4			●	1 / 716	726.86 (+2); 731.38 (+2); 500.29 (+3)	-
Q96G03	Phosphoglucomutase-2			●	25 / 716	603.99 (+3); 499.24 (+2); 599.81 (+2)	-

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
P18669	Phosphoglycerate mutase 1			●	239 / 716	384.23 (+3); 438.20 (+3); 530. 28 (+2)	-
P55058	Phospholipid transfer protein			●	0 / 716	424.56 (+3); 664.33 (+2)	-
P13797	Plastin-3			●	208 / 716	759.88 (+2); 552.94 (+3); 513.26 (+3); 516.28 (+2); 628.63 (+3)	-
Q8WUM4	Programmed cell death 6-interacting protein			●	227 / 716	532.80 (+2); 487.29 (+2); 450.23 (+2); 766.86 (+2)	-
Q99623	Prohibitin-2			●	243 / 716	608.31 (+2); 497.74 (+2)	-
Q9UM07	Protein-arginine deiminase type-4			●	0 / 716	565.81 (+2); 385.25 (+2)	-
P61026	Ras-related protein Rab-10			●	176 / 716	658.83 (+2); 430.21 (+3)	-
P62491	Ras-related protein Rab-11A			●	125 / 716	637.81 (+2); 522.31 (+2); 540.79 (+2)	-
P08134	Rho-related GTP-binding protein RhoC			●	77 / 716	544.79 (+2); 499.73 (+2)	-
P13489	Ribonuclease inhibitor			●	70 / 716	575.26 (+2); 605.35 (+2)	-
P02787	Serotransferrin			●	11 / 716	863.39 (+2); 489.75 (+2)	-
Q9UIV8	Serpin B13			●	1 / 716	733.37 (+2); 819.36 (+2); 481.75 (+2); 532.28 (+3); 529.58 (+3); 420.24 (+2); 419.24 (+2)	-
P29508	Serpin B3			●	45 / 716	511.82 (+2); 431.26 (+2); 432.74 (+2)	-
P22531	Small proline-rich protein 2E			●	28 / 716	753.36 (+2); 732.85 (+2); 715.32 (+3)	-
P11169	Solute carrier family 2, facilitated glucose transporter member 3			●	1 / 716	441.22 (+2); 552.82 (+2)	-

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)	m/z (charge) for MS/MS	Enrichment
O00391	Sulfhydryl oxidase 1			●	1 / 716	473.75 (+2); 678.84 (+2); 557.81 (+2)	-
Q6UWP8	Suprabasin			●	68 / 716	691.36 (+2); 859.90 (+2); 434.70 (+4); 441.71 (+4)	-
P10599	Thioredoxin			●	432 / 716	668.82 (+2); 493.93 (+3); 611.27 (+2)	-
P07996	Thrombospondin-1			●	14 / 716	623.85 (+2); 515.75 (+2); 515.60 (+3); 625.97 (+3)	-
Q86T26	Transmembrane protease serine 11B			●	0 / 716	572.82 (+2); 415.72 (+2)	-
P06753	Tropomyosin alpha-3 chain			●	354 / 716	454.73 (+4); 576.63 (+3); 491.59 (+3)	-
Q71U36	Tubulin alpha-1A chain			●	694 / 716	508.29 (+2); 573.63 (+3); 633.28 (+2)	-
Q13885	Tubulin beta-2A chain			●	675 / 716	623.30 (+2); 580.32 (+2)	-
P09769	Tyrosine-protein kinase Fgr			●	157 / 716	431.74 (+2); 471.77 (+2); 433.74 (+2)	-
P62987	Ubiquitin-60S ribosomal protein L40			●	462 / 716	894.47 (+2); 508.60 (+3); 541.28 (+2)	-
P18206	Vinculin			●	156 / 716	585.83 (+2); 553.31 (+2)	-
O75083	WD repeat-containing protein 1			●	89 / 716	513.29 (+3); 551.76 (+2)	-

CONTAMINANTS FROM SAMPLE MANIPULATION AND THE SAMPLE ITSELF

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)
P69905	Hemoglobin subunit alpha	●		●	90 / 716
P68871	Hemoglobin subunit beta	●		●	113 / 716
Q9Y6R7	IgGFc-binding protein	●	●	●	4 / 716
P0DOX2	Immunoglobulin alpha-2 heavy chain	●	●	●	Unmapped
P0DOX3	Immunoglobulin delta heavy chain	●		●	Unmapped

CONTAMINANTS FROM SAMPLE MANIPULATION AND THE SAMPLE ITSELF

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)
P0DOX5	Immunoglobulin gamma-1 heavy chain	●	●	●	Unmapped
P01876	Immunoglobulin heavy constant alpha 1	●	●	●	51 / 716
P01859	Immunoglobulin heavy constant gamma 2	●	●	●	142 / 716
P01860	Immunoglobulin heavy constant gamma 3	●	●	●	54 / 716
P01861	Immunoglobulin heavy constant gamma 4	●	●	●	52 / 716
P01871	Immunoglobulin heavy constant mu	●	●	●	2 / 716
A0A0A0MS14	Immunoglobulin heavy variable 1-45	●			43 / 716
A0A0B4J1V0	Immunoglobulin heavy variable 3-15	●	●	●	1 / 716
P01768	Immunoglobulin heavy variable 3-30	●		●	1 / 716
A0A0A0MS15	Immunoglobulin heavy variable 3-49	●	●	●	0 / 716
A0A0J9YX35	Immunoglobulin heavy variable 3-64D	●	●	●	2 / 716
P01780	Immunoglobulin heavy variable 3-7	●	●	●	0 / 716
A0A0B4J1X5	Immunoglobulin heavy variable 3-74	●	●	●	0 / 716
P06331	Immunoglobulin heavy variable 4-34	●	●	●	0 / 716
A0A0B4J1U7	Immunoglobulin heavy variable 6-1	●	●	●	6 / 716
P01591	Immunoglobulin J chain	●	●	●	16 / 716
P01834	Immunoglobulin kappa constant	●	●	●	36 / 716
P0DOX7	Immunoglobulin kappa light chain	●	●	●	Unmapped
P01599	Immunoglobulin kappa variable 1-17	●	●	●	0 / 716
P06310	Immunoglobulin kappa variable 2-30	●	●		218 / 716

CONTAMINANTS FROM SAMPLE MANIPULATION AND THE SAMPLE ITSELF

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)
A0A087WW87	Immunoglobulin kappa variable 2-40	●	●	●	220 / 716
P01619	Immunoglobulin kappa variable 3-20	●	●	●	1 / 716
A0A0A0MRZ8	Immunoglobulin kappa variable 3D-11	●	●	●	1 / 716
P06312	Immunoglobulin kappa variable 4-1	●	●	●	29 / 716
P0DOY2	Immunoglobulin lambda constant 2	●	●	●	33 / 716
P01700	Immunoglobulin lambda variable 1-47	●		●	0 / 716
P01701	Immunoglobulin lambda variable 1-51	●	●	●	0 / 716
A0A075B6K5	Immunoglobulin lambda variable 3-9	●			0 / 716
A0A075B6I0	Immunoglobulin lambda variable 8-61	●	●		0 / 716
B9A064	Immunoglobulin lambda-like polypeptide 5	●	●	●	20 / 716
P13645	Keratin, type I cytoskeletal 10	●	●	●	616 / 716
P13646	Keratin, type I cytoskeletal 13	●	●	●	441 / 716
P02533	Keratin, type I cytoskeletal 14	●	●	●	523 / 716
P08779	Keratin, type I cytoskeletal 16	●	●	●	516 / 716
Q04695	Keratin, type I cytoskeletal 17	●	●	●	465 / 716
P08727	Keratin, type I cytoskeletal 19	●		●	506 / 716
P35527	Keratin, type I cytoskeletal 9	●	●	●	577 / 716
Q14533	Keratin, type II cuticular Hb1	●		●	131 / 716
P04264	Keratin, type II cytoskeletal 1	●	●	●	671 / 716
P35908	Keratin, type II cytoskeletal 2 epidermal	●	●	●	628 / 716
P19013	Keratin, type II cytoskeletal 4	●	●	●	391 / 716
P13647	Keratin, type II cytoskeletal 5	●	●	●	508 / 716

CONTAMINANTS FROM SAMPLE MANIPULATION AND THE SAMPLE ITSELF

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)
P02538	Keratin, type II cytoskeletal 6A	●	●	●	515 / 716
P04259	Keratin, type II cytoskeletal 6B	●	●	●	570 / 716
Q8N1N4	Keratin, type II cytoskeletal 78	●	●	●	191 / 716
Q5T749	Keratinocyte proline-rich protein	●	●	●	81 / 716
A0A0C4DH36	Probable non-functional immunoglobulin heavy variable 3-38	●			0 / 716
P27105	Stomatin OS=Homo sapiens	●	●	●	10 / 716
A0A075B6K4	Immunoglobulin lambda variable 3-10		●	●	0 / 716
A0A075B6R2	Immunoglobulin heavy variable 4-4			●	0 / 716
A0A075B6R9	Probable non-functional immunoglobulin kappa variable 2D-24			●	25 / 716
A0A075B6S5	Immunoglobulin kappa variable 1-27			●	16 / 716
A0A0B4J1Y9	Immunoglobulin heavy variable 3-72			●	0 / 716
A0A0C4DH34	Immunoglobulin heavy variable 4-28			●	0 / 716
A0A0C4DH38	Immunoglobulin heavy variable 5-51		●	●	0 / 716
P01593	Immunoglobulin kappa variable 1D-33			●	10 / 716
P01714	Immunoglobulin lambda variable 3 -19			●	0 / 716
P01718	Immunoglobulin lambda variable 3-27			●	0 / 716
P04430	Immunoglobulin kappa variable 1-16		●	●	0 / 716
P23083	Immunoglobulin heavy variable 1-2		●	●	0 / 716

CONTAMINANTS FROM SAMPLE MANIPULATION AND THE SAMPLE ITSELF

UniProt-KB Code	Protein name	AD and HC (R-)	NEG (R-)	AD and HC (NR-)	CRAPome (found/total)
Q01546	Keratin, type II cytoskeletal 2 oral			●	426 / 716
Q6KB66	Keratin, type II cytoskeletal 80		●	●	281 / 716
Q7Z794	Keratin, type II cytoskeletal 1b			●	592 / 716
Q86Y46	Keratin, type II cytoskeletal 73			●	231 / 716
A0A0B4J1V2	Immunoglobulin heavy variable 2-26		●		0 / 716
A0A0B4J1Y8	Immunoglobulin lambda variable 9-49		●		0 / 716
A0M8Q6	Immunoglobulin lambda constant 7		●		14 / 716
P0DP04	Immunoglobulin heavy variable 3-43D		●		0 / 716
P15814	Immunoglobulin lambda-like polypeptide 1		●		3 / 716

Fig. S3.1: Charts of all MF (panel A), CC (panel B), and BP (panel C) annotated for cystatin B and all its interactors according to GO-terms. Proteins are identified by their Uniprot-KB codes (x-axis) and are in the same order of Table 3.4 (from left to right), while GO-terms are annotated in the y-axis. The legend of colors is reported at the top of the figure.

Legend:

The colors for different evidence codes in the table:

Gene Ontology

- Inferred from experiment [IDA, IPI, IMP, IGI, IEP]
- Direct assay [IDA], Mutant phenotype [IMP]
- Genetic interaction [IGI], Physical interaction [IPI]
- Inferred from High Throughput Experiment [HDA, HMP, HGI, HEP]
- High Throughput Direct Assay [HDA], High Throughput Mutant Phenotype [HMP]
- High Throughput Genetic interaction [HGI], High Throughput Expression pattern [HEP]
- Traceable author [TAS], Non-traceable author [NAS], Inferred by curator [IC]
- Expression pattern [IEP], Sequence or structural similarity [ISS], Genomic context [IGC]
- Sequence Model [ISM], Sequence Alignment [ISA], Sequence Orthology [ISO]
- Biological aspect of ancestor [IBA], Rapid divergence [IRD]
- Reviewed computational analysis [RCA], Electronic annotation [IEA]
- No biological data [ND], Not annotated or not in background [NA]

Fig. S3.1, panel A: Molecular Function (MF) according to GO-terms.

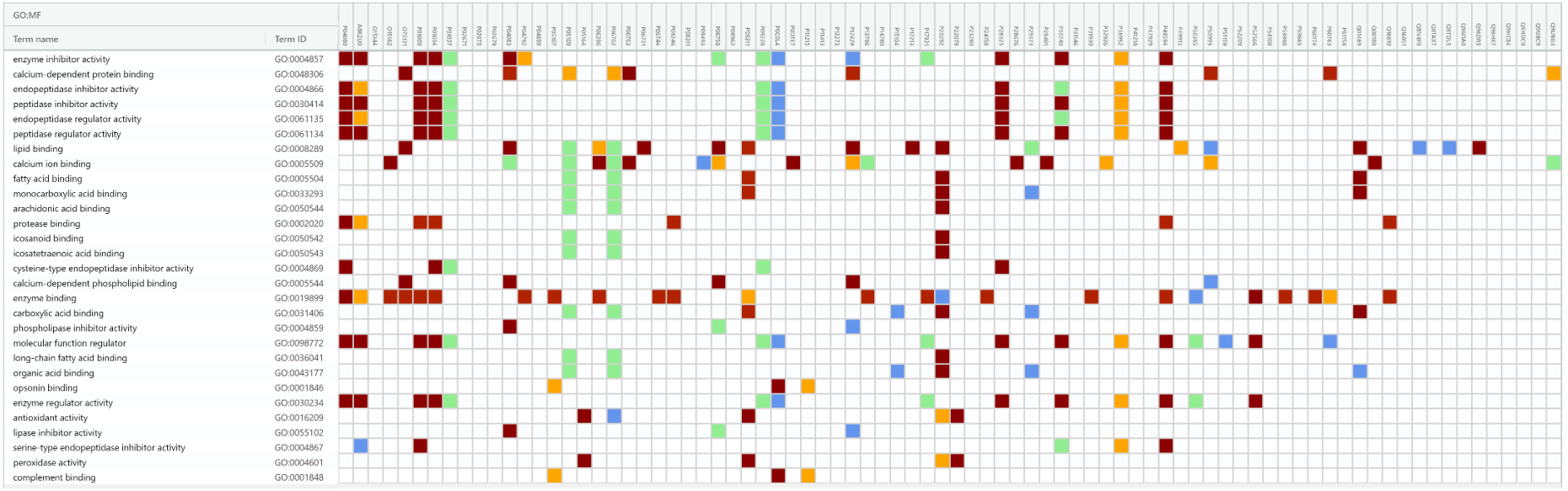


Fig S3.1, panel B: Cellular Component (CC) according to GO-terms.

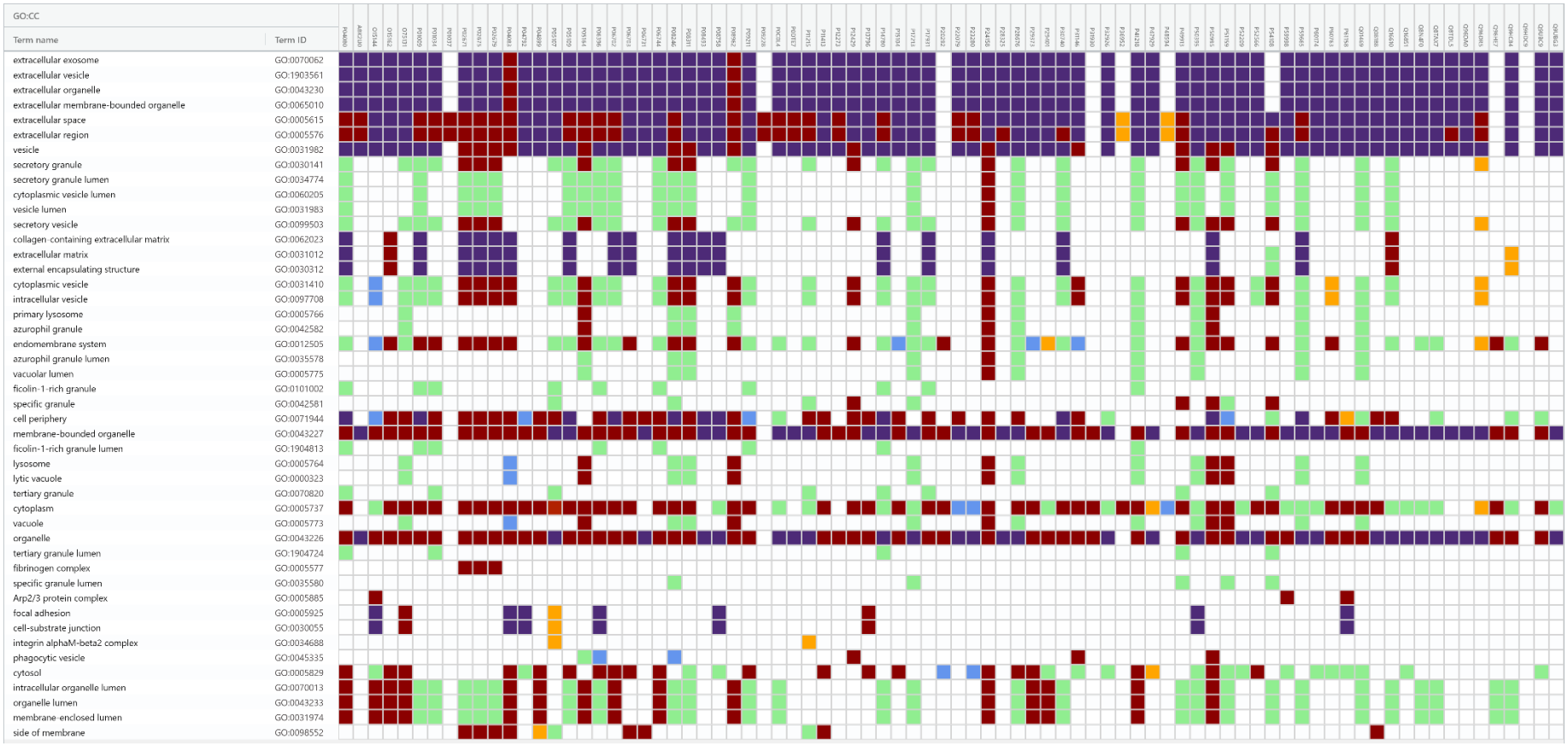
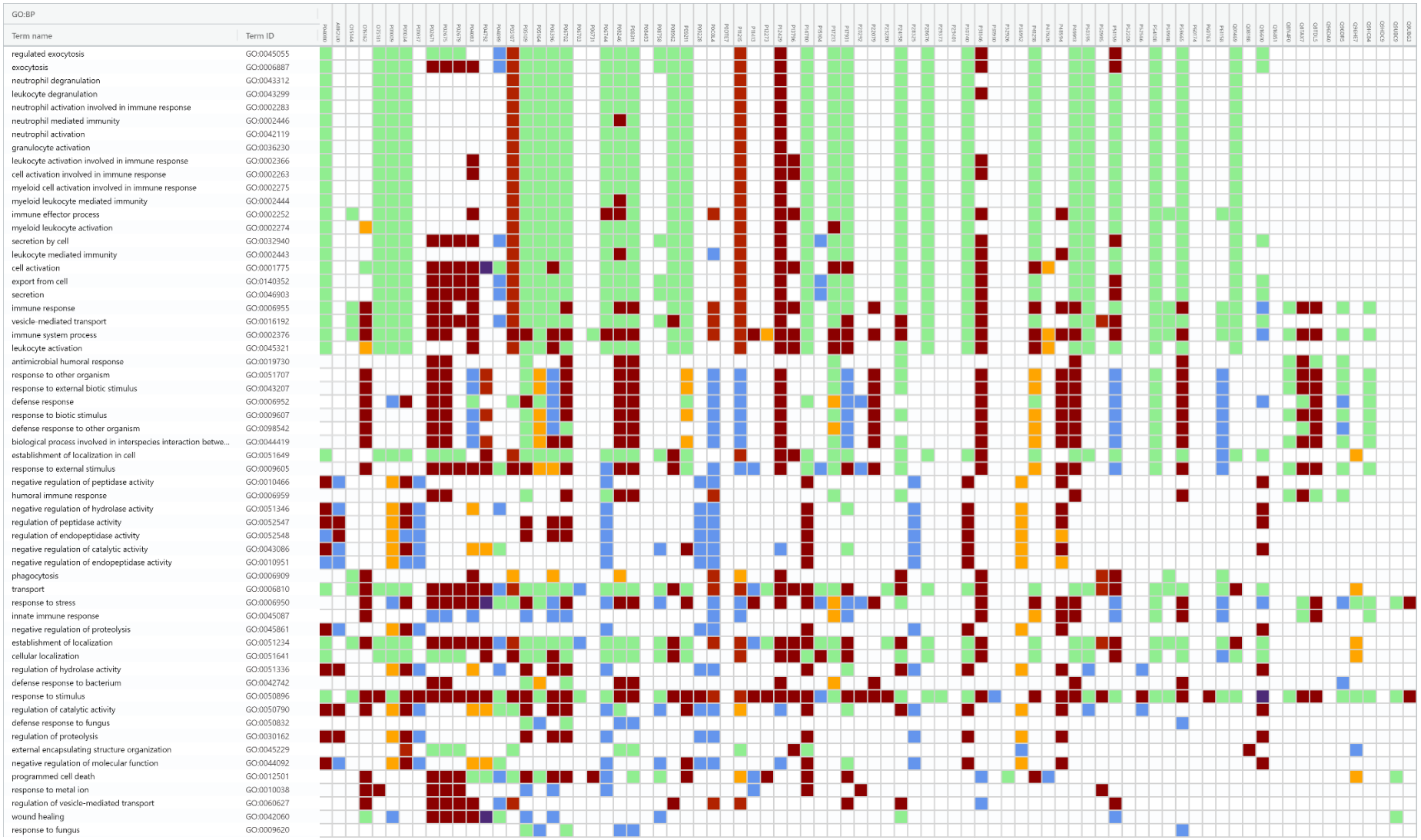
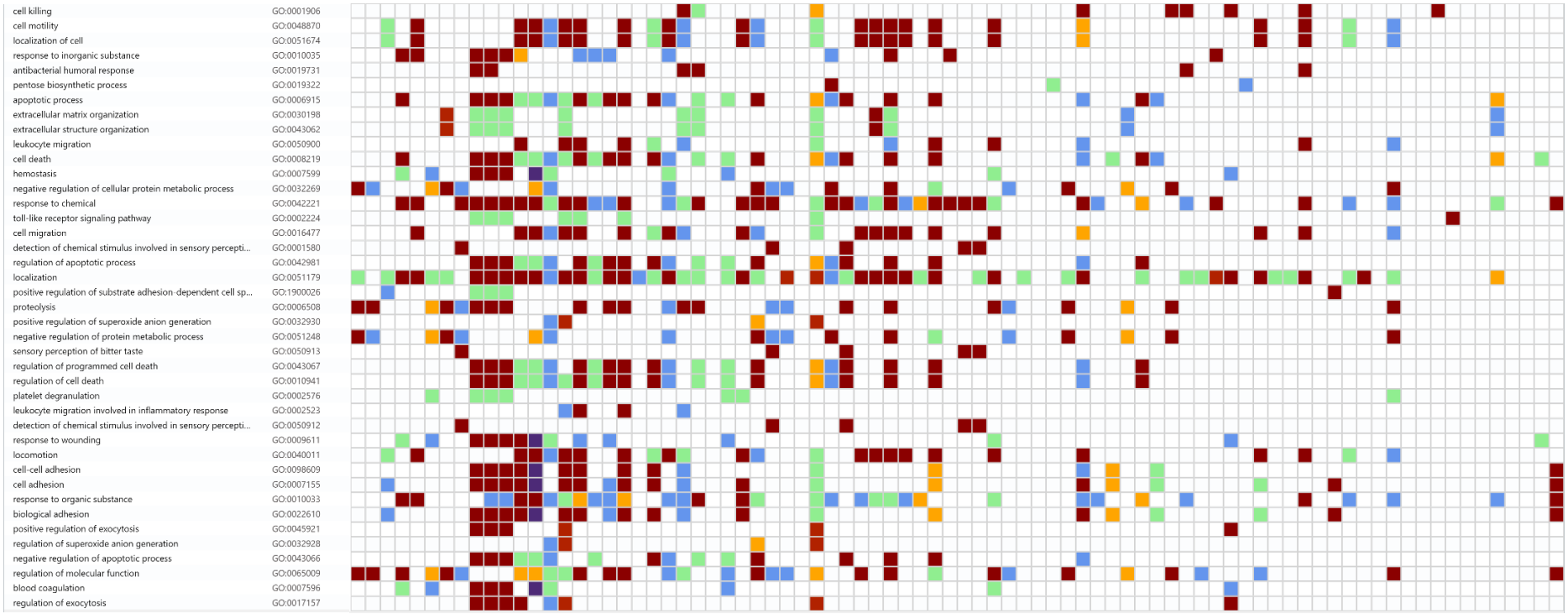


Fig. S3.1, panel C: Biological Process (BP) according to GO-terms




cell killing GO:0001906
cell motility GO:0048870
localization of cell GO:0051674
response to inorganic substance GO:0010035
antibacterial humoral response GO:0019731
pentose biosynthetic process GO:0019322
apoptotic process GO:0006915
extracellular matrix organization GO:0030198
extracellular structure organization GO:0043062
leukocyte migration GO:0050900
cell death GO:0008219
hemostasis GO:0007599
negative regulation of cellular protein metabolic process GO:0032269
response to chemical GO:0042221
toll-like receptor signaling pathway GO:0002224
cell migration GO:0016477
detection of chemical stimulus involved in sensory percepti... GO:0001580
regulation of apoptotic process GO:0042981
localization GO:0051179
positive regulation of substrate adhesion-dependent cell sp... GO:1900026
proteolysis GO:0006508
positive regulation of superoxide anion generation GO:0032930
negative regulation of protein metabolic process GO:0051248
sensory perception of bitter taste GO:0050913
regulation of programmed cell death GO:0043067
regulation of cell death GO:0010941
platelet degranulation GO:0002576
leukocyte migration involved in inflammatory response GO:0002523
detection of chemical stimulus involved in sensory percepti... GO:0050912
response to wounding GO:0009611
locomotion GO:0040011
cell-cell adhesion GO:0098609
cell adhesion GO:0007155
response to organic substance GO:0010033
biological adhesion GO:0022610
positive regulation of exocytosis GO:0045921
regulation of superoxide anion generation GO:0032928
negative regulation of apoptotic process GO:0043066
regulation of molecular function GO:0065009
blood coagulation GO:0007596
regulation of exocytosis GO:0017157

S3.3: Proteins excluded because found enriched in NEG (↑) or unchanged (-) between NEG and AD or HC groups despite showing a good CRAPome score. For each protein it is reported the UniProt-KB code, the LFQ abundance in log2 scale measured in each replicate, the *p* value and the fold change calculated by Perseus.

		Log2 LFQ abundances								AD vs NEG		HC vs NEG		
UniProt-KB Code	**Protein Name**	**AD**			**HC**			**NEG**		**-Log10 p value**	**Fold change**	**-Log10 p value**	**Fold change**	
P28799	Progranulin	16.1	17.7	17.3	17.2	17.6	17.1	20.1	19.9	1.8	-3.0	3.0	-2.7	↑ NEG
P05089	Arginase-1	15.7	17.5	17.0	16.9	17.4	16.9	20.6	18.9	1.3	-3.0	1.6	-2.7	↑ NEG
Q6P4A8	Phospholipase B-like 1	15.8	16.8	16.2	16.9	16.8	16.1	18.1	17.5	1.4	-1.5	1.3	-1.2	↑ NEG
Q08554	Desmocollin-1	15.9	17.0	17.1	17.0	17.0	17.0	21.4	18.8	1.3	-3.4	1.3	-3.1	↑ NEG
Q9NP55	BPI fold-containing family A member 1	15.5	16.5	16.0	16.7	16.4	15.9	16.7	17.9	1.0	-1.3	0.8	-1.0	-
Q9NZT1	Calmodulin-like protein 5	16.9	16.9	17.9	17.9	16.8	17.8	18.5	17.0	0.3	-0.5	0.1	-0.2	-
P04040	Catalase	16.8	16.6	17.3	17.8	16.5	17.2	19.8	17.9	1.1	-2.0	0.9	-1.7	-
Q5D862	Filaggrin-2	15.9	16.4	17.2	17.0	16.4	17.1	17.9	16.2	0.2	-0.5	0.1	-0.2	-
Q9H4G4	Golgi-associated plant pathogenesis-related protein 1	16.8	16.3	17.4	17.8	16.2	17.2	17.0	18.1	0.5	-0.7	0.3	-0.5	-
P22735	Protein-glutamine gamma-glutamyltransferase K	17.3	15.9	16.7	18.2	15.9	16.6	18.9	17.3	0.8	-1.5	0.5	-1.2	-
P02647	Apolipoprotein A-I	17.7	16.0	19.2	18.8	17.4	19.2	19.6	19.7	0.7	-2.0	0.8	-1.2	-
P20160	Azurocidin	25.1	25.6	25.7	25.0	24.0	25.1	24.0	24.4	1.8	1.2	0.4	0.5	-
P02747	Complement C1q subcomponent subunit C	20.0	21.1	20.7	22.5	21.6	22.3	21.4	21.9	1.0	-1.1	0.5	0.5	-
P01024	Complement C3	20.0	18.5	18.8	19.7	18.9	18.9	19.3	18.2	0.2	0.3	0.3	0.4	-
P12830	Cadherin-1	20.1	19.5	19.3	19.4	18.3	17.4	19.1	18.8	0.9	0.7	0.3	-0.6	-
Q14CN2	Calcium-activated chloride channel regulator 4	17.9	16.9	16.8	17.3	16.9	15.8	16.4	15.7	0.9	1.1	0.4	0.6	-
Q9UGM3	Deleted in malignant brain tumors 1 protein	26.7	26.2	26.9	26.4	26.4	26.5	25.6	24.6	1.4	1.5	1.5	1.3	-

UniProt-KB Code	Protein Name	Log2 LFQ abundances								AD vs NEG		HC vs NEG		
		AD			HC			NEG		-Log10 p value	Fold change	-Log10 p value	Fold change	
Q02413	Desmoglein-1	15.8	18.4	19.5	19.0	19.1	18.3	21.6	20.0	0.8	-2.9	1.3	-2.0	↑ NEG
P00738	Haptoglobin	20.5	20.3	19.8	20.1	17.9	19.6	19.2	20.3	0.4	0.5	0.2	-0.5	-
Q9P0G3	Kallikrein-14	19.8	20.1	19.7	20.0	19.2	20.2	18.4	19.1	1.5	1.2	0.9	1.1	-
Q08380	Galectin-3-binding protein	21.8	20.8	22.1	21.4	21.0	21.7	21.3	20.4	0.5	0.7	0.5	0.5	-
P02788	Lactotransferrin	26.0	24.5	25.2	26.6	26.2	26.5	26.3	25.0	0.2	-0.4	0.7	0.8	-
P61626	Lysozyme C	24.5	24.9	24.2	25.2	23.5	23.6	24.0	23.7	1.0	0.7	0.1	0.2	-
O75594	Peptidoglycan recognition protein 1	18.7	18.4	18.5	18.1	17.7	18.0	18.2	19.3	0.2	-0.2	0.8	-0.8	-
P01833	Polymeric immunoglobulin receptor	27.2	25.1	25.6	25.9	25.6	25.6	26.3	25.8	0.1	-0.1	0.7	-0.4	-
Q13835	Plakophilin-1	17.0	18.9	19.5	19.8	19.2	18.8	19.8	16.4	0.1	0.4	0.4	1.2	-
P12724	Eosinophil cationic protein	22.7	22.6	20.9	22.3	23.2	21.6	19.8	21.6	0.6	1.4	0.8	1.7	-
Q96P63	Serpin B12	16.4	18.4	19.2	19.4	19.3	17.9	19.9	16.9	0.1	-0.4	0.1	0.5	-
P03973	Antileukoproteinase	25.9	26.8	25.4	26.2	25.4	25.4	25.8	26.2	0.0	0.0	0.3	-0.3	-

Table S3.4: Perseus matrix of proteins unchanged between AD and HC groups after statistics of cystatin B and its interactors. For each protein it is reported the UniProt-KB code, the LFQ abundance in log2 scale measured in each replicate, the *p* value and the fold change calculated by Perseus.

		Log2 LFQ abundances						AD VS HC	
UniProt-KB Code	**Protein name**	**AD**			**HC**			**-Log10 p value**	**Fold change**
O75131	Copine-3	17.7	17.4	17.6	18.7	18.5	18.2	2.2	-0.9
P31146	Coronin-1A	22.2	22.2	22.7	23.2	22.9	23.1	1.7	-0.7
Q8TDL5	BPI fold-containing family B member 1	23.8	24.0	24.0	24.3	25.0	25.3	1.5	-1.0
Q9HC84	Mucin-5B	24.6	24.6	24.5	24.3	23.5	23.5	1.4	0.8
P02679	Fibrinogen gamma chain	22.4	22.0	22.0	22.9	22.8	22.4	1.4	-0.5
P20292	Arachidonate 5-lipoxygenase-activating protein	21.3	21.7	21.9	22.6	22.3	22.1	1.3	-0.7
P23280	Carbonic anhydrase 6	24.0	24.4	24.1	23.2	22.9	23.9	1.3	0.8
Q8N4F0	BPI fold-containing family B member 2	24.6	24.0	24.2	23.9	23.3	23.6	1.2	0.6
P32926	Desmoglein-3	19.9	19.9	20.2	21.0	20.4	20.4	1.2	-0.6
Q96DR5	BPI fold-containing family A member 2	26.6	24.8	25.9	24.9	23.1	24.4	1.0	1.6
P28325	Cystatin-D	22.8	22.0	22.2	22.1	20.4	20.8	1.0	1.2
Q96DA0	Zymogen granule protein 16 homolog B	27.0	26.4	26.9	26.3	24.9	26.1	1.0	1.0
P0DTE7	Alpha-amylase 1B	26.7	27.0	26.6	25.8	26.6	26.0	1.0	0.6
P29373	Cellular retinoic acid-binding protein 2	20.7	20.1	20.6	20.3	19.6	19.3	1.0	0.8
P13796	Plastin-2	22.8	22.5	22.5	22.4	22.4	22.2	0.9	0.3
P22079	Lactoperoxidase	20.6	20.7	21.2	20.3	20.2	20.6	0.9	0.4
P08246	Neutrophil elastase	27.9	28.5	29.0	27.5	27.6	28.1	0.9	0.7
P01037	Cystatin-SN	24.3	24.4	24.7	24.0	22.9	24.2	0.9	0.8
P04792	Heat shock protein beta-1	21.3	22.6	23.1	23.3	23.2	23.7	0.9	-1.1
P08311	Cathepsin G	27.5	28.4	28.3	27.5	26.9	27.7	0.9	0.7
A8K2U0	Alpha-2-macroglobulin-like protein 1	21.3	20.8	20.9	21.8	21.3	21.3	0.9	-0.4
P52566	Rho GDP-dissociation inhibitor 2	20.9	19.6	20.6	20.0	18.8	19.6	0.8	0.9
P49913	Cathelicidin antimicrobial peptide	19.1	19.2	19.8	18.3	18.2	19.3	0.8	0.8

UniProt-KB Code	Protein name	Log2 LFQ abundances						AD VS HC	
		AD			HC			-Log10 p value	Fold change
P30740	Leukocyte elastase inhibitor	23.5	23.4	23.3	23.1	22.5	23.3	0.8	0.4
P54108	Cysteine-rich secretory protein 3	21.6	20.7	21.0	20.6	19.4	20.7	0.8	0.8
P01034	Cystatin-C	21.1	22.1	21.5	21.6	20.3	20.6	0.7	0.8
P59665	Neutrophil defensin 1	27.4	27.6	28.1	27.6	28.6	28.7	0.7	-0.6
P14780	Matrix metalloproteinase-9	19.3	17.9	18.3	19.1	19.8	18.8	0.7	-0.8
P50995	Annexin A11	18.7	18.9	19.4	19.0	19.8	19.6	0.7	-0.5
P31930	Cytochrome b-c1 complex subunit 1	18.8	18.8	18.9	19.0	19.8	18.9	0.7	-0.4
P09211	Glutathione S-transferase P	22.8	23.1	23.2	23.1	22.4	22.5	0.6	0.4
Q16851	UTP--glucose-1-phosphate uridylyltransferase	18.9	19.4	19.4	19.3	19.7	19.6	0.6	-0.3
P47929	Galectin-7	20.2	20.4	20.2	21.7	20.9	20.1	0.6	-0.6
P04080	Cystatin-B	26.6	25.9	25.6	26.4	24.8	24.0	0.6	1.0
P24158	Myeloblastin	23.4	24.2	24.5	23.7	22.6	23.8	0.6	0.7
P05164	Myeloperoxidase	26.3	26.6	26.3	26.2	26.3	26.2	0.6	0.1
P09228	Cystatin-SA	19.3	17.8	19.2	18.2	16.9	18.5	0.6	0.9
P11215	Integrin alpha-M	22.0	19.9	20.4	21.9	21.7	20.9	0.5	-0.8
P0C0L4	Complement C4-A	19.4	17.6	17.7	18.9	18.9	18.7	0.4	-0.6
P06702	Protein S100-A9	25.7	25.0	25.8	25.8	24.8	24.5	0.4	0.5
P51159	Ras-related protein Rab-27A	19.2	19.5	20.4	20.7	19.4	20.4	0.4	-0.5
P15104	Glutamine synthetase	18.5	17.3	18.5	18.5	15.3	17.7	0.4	0.9
Q96HE7	ERO1-like protein alpha	18.3	19.2	20.3	20.0	19.5	19.8	0.4	-0.5
P05107	Integrin beta-2	21.8	20.2	21.1	21.5	21.4	21.3	0.4	-0.4
Q01469	Fatty acid-binding protein 5	24.0	24.2	23.7	24.2	23.6	23.5	0.3	0.2
P12429	Annexin A3	23.4	22.8	23.2	23.2	21.3	23.3	0.3	0.5
O15144	Actin-related protein 2/3 complex subunit 2	19.0	18.3	18.6	19.0	16.4	18.6	0.3	0.6
P12273	Prolactin-inducible protein	25.7	24.7	24.3	25.0	24.6	24.0	0.3	0.4
P08758	Annexin A5	19.0	18.3	19.3	19.0	17.1	19.0	0.3	0.5
P01009	Alpha-1-antitrypsin	20.4	19.9	19.7	19.8	20.0	19.5	0.3	0.2
P52209	6-phosphogluconate dehydrogenase, decarboxylating	24.2	24.0	24.3	24.2	23.9	24.1	0.3	0.1

UniProt-KB Code	Protein name	Log2 LFQ abundances						AD VS HC	
		AD			HC			-Log10 p value	Fold change
P11413	Glucose-6-phosphate 1-dehydrogenase	20.4	20.6	20.5	20.5	20.6	20.4	0.2	0.0
P04899	Guanine nucleotide-binding protein G(i) subunit alpha-2	20.3	20.1	21.0	21.0	20.2	20.9	0.2	-0.2
P02671	Fibrinogen alpha chain	20.2	18.9	19.9	20.4	19.6	19.7	0.2	-0.3
P05109	Protein S100-A8	26.6	26.4	27.0	26.9	26.4	27.3	0.2	-0.2
P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	20.1	17.6	19.3	18.9	19.6	19.7	0.2	-0.4
P06396	Gelsolin	20.5	20.3	21.3	20.9	20.4	21.4	0.2	-0.2
P17931	Galectin-3	19.0	20.1	20.3	20.4	19.7	20.0	0.2	-0.2
P04083	Annexin A1	26.5	26.0	26.9	26.8	25.5	26.6	0.1	0.2
Q9UBG3	Cornulin	18.6	20.5	22.6	21.6	20.4	21.1	0.1	-0.4
P50395	Rab GDP dissociation inhibitor beta	19.2	18.8	19.7	19.4	18.6	19.4	0.1	0.1
P41218	Myeloid cell nuclear differentiation antigen	22.6	21.7	23.1	23.1	21.9	23.1	0.1	-0.2
P61158	Actin-related protein 3	18.3	18.5	18.7	18.7	17.4	18.9	0.1	0.2
P08962	CD63 antigen	20.8	18.3	19.4	20.0	18.7	19.2	0.1	0.2
P60763	Ras-related C3 botulinum toxin substrate 3	21.7	21.6	22.1	22.1	21.6	21.4	0.1	0.1
Q9HDC9	Adipocyte plasma membrane-associated protein	20.8	19.4	20.3	20.6	18.9	20.4	0.1	0.2
P02675	Fibrinogen beta chain	21.9	20.8	20.9	21.2	21.3	21.3	0.1	-0.1
P59998	Actin-related protein 2/3 complex subunit 4	18.9	19.4	18.8	19.6	19.8	17.1	0.1	0.2
Q16610	Extracellular matrix protein 1	19.8	20.2	20.8	20.7	20.0	20.2	0.1	-0.1
Q9UBC9	Small proline-rich protein 3	21.3	21.9	23.7	22.3	21.7	22.6	0.1	0.1
P36952	Serpin B5	20.0	19.4	20.7	20.6	19.6	20.1	0.0	-0.1
P48594	Serpin B4	21.0	20.6	21.4	21.2	20.7	21.1	0.0	0.0
P29401	Transketolase	20.9	20.8	21.4	20.9	20.8	21.5	0.0	0.0
O15162	Phospholipid scramblase 1	18.2	19.0	18.6	19.1	18.1	18.7	0.0	0.0
P06744	Glucose-6-phosphate isomerase	21.4	21.4	22.2	21.6	21.6	21.8	0.0	0.0
Q08188	Protein-glutamine gamma-glutamyltransferase E	26.0	24.3	24.9	25.4	24.7	25.1	0.0	0.0
P06703	Protein S100-A6	19.3	19.8	20.9	19.4	19.9	20.6	0.0	0.0

Table S3.5: Perseus matrix of proteins unchanged between AD and HC groups after statistics of stoichiometric ratios between cystatin B interactors and cystatin B itself. For each protein it is reported the UniProt-KB code, the LFQ abundance in log2 scale measured in each replicate, the *p* value and the fold change calculated by Perseus.

		Log2 LFQ abundances						AD VS HC	
UniProt-KB Code	**Protein name**	**AD**			**HC**			**-Log10 p value**	**Fold change**
A8K2U0	Alpha-2-macroglobulin-like protein 1	-5.2	-5.1	-4.8	-4.6	-3.5	-2.8	1.2	-1.4
P32926	Desmoglein-3	-6.6	-6.0	-5.4	-5.3	-4.4	-3.6	1.2	-1.6
P02679	Fibrinogen gamma chain	-4.2	-3.9	-3.6	-3.5	-2.0	-1.6	1.2	-1.5
P60174	Triosephosphate isomerase	-8.7	-7.7	-6.7	-6.7	-5.4	-3.9	1.1	-2.4
P20292	Arachidonate 5-lipoxygenase-activating protein	-5.3	-4.2	-3.7	-3.8	-2.4	-1.9	1.1	-1.7
P14780	Matrix metalloproteinase-9	-7.3	-8.0	-7.4	-7.3	-4.9	-5.2	1.1	-1.8
P31146	Coronin-1A	-4.4	-3.7	-2.9	-3.2	-1.9	-0.9	1.0	-1.7
P28676	Grancalcin	-4.6	-3.2	-2.5	-2.6	-1.4	0.0	1.0	-2.1
P0C0L4	Complement C4-A	-7.2	-8.3	-8.0	-7.5	-5.9	-5.3	1.0	-1.6
P02671	Fibrinogen alpha chain	-6.4	-7.0	-5.8	-6.0	-5.1	-4.4	1.0	-1.3
P05107	Integrin beta-2	-4.8	-5.7	-4.6	-4.9	-3.4	-2.7	0.9	-1.4
Q8TDL5	BPI fold-containing family B member 1	-2.8	-1.9	-1.7	-2.0	0.2	1.3	0.9	-2.0
P04792	Heat shock protein beta-1	-5.3	-3.3	-2.6	-3.1	-1.6	-0.4	0.8	-2.1
P51159	Ras-related protein Rab-27A	-7.4	-6.4	-5.3	-5.7	-5.3	-3.6	0.8	-1.5
P31930	Cytochrome b-c1 complex subunit 1	-7.8	-7.1	-6.7	-7.3	-5.0	-5.1	0.8	-1.4
P60763	Ras-related C3 botulinum toxin substrate 3	-4.9	-4.3	-3.5	-4.2	-3.1	-2.6	0.7	-0.9
P02675	Fibrinogen beta chain	-4.7	-5.1	-4.8	-5.1	-3.5	-2.7	0.7	-1.1
P04899	Guanine nucleotide-binding protein G(i) subunit alpha-2	-6.2	-5.8	-4.6	-5.4	-4.6	-3.1	0.7	-1.2
Q08188	Protein-glutamine gamma-glutamyltransferase E	-0.6	-1.6	-0.7	-0.9	-0.1	1.1	0.7	-1.0
Q01469	Fatty acid-binding protein 5	-2.6	-1.7	-1.9	-2.1	-1.2	-0.6	0.7	-0.8
Q96HE7	ERO1-like protein alpha	-8.3	-6.7	-5.4	-6.3	-5.3	-4.2	0.7	-1.5
P50995	Annexin A11	-7.9	-7.0	-6.2	-7.3	-5.0	-4.5	0.7	-1.5
Q16851	UTP--glucose-1-phosphate uridylyltransferase	-7.7	-6.5	-6.2	-7.0	-5.1	-4.5	0.6	-1.3

UniProt-KB Code	Protein name	Log2 LFQ abundances						AD VS HC	
		AD			HC			-Log10 p value	Fold change
O15162	Phospholipid scramblase 1	-8.3	-6.9	-7.1	-7.2	-6.7	-5.4	0.6	-1.0
P59665	Neutrophil defensin 1	0.8	1.7	2.5	1.2	3.9	4.7	0.6	-1.6
P36952	Serpin B5	-6.5	-6.5	-4.9	-5.7	-5.2	-3.9	0.6	-1.1
P12273	Prolactin-inducible protein	-0.8	-1.2	-1.3	-1.4	-0.1	0.0	0.6	-0.6
P17931	Galectin-3	-7.6	-5.8	-5.3	-6.0	-5.1	-4.0	0.6	-1.2
P05109	Protein S100-A8	0.0	0.5	1.4	0.5	1.6	3.3	0.6	-1.2
Q16610	Extracellular matrix protein 1	-6.8	-5.7	-4.8	-5.7	-4.7	-3.8	0.6	-1.1
P41218	Myeloid cell nuclear differentiation antigen	-4.0	-4.2	-2.5	-3.3	-2.8	-0.9	0.6	-1.2
P48594	Serpin B4	-5.6	-5.3	-4.2	-5.1	-4.0	-2.9	0.6	-1.0
P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	-6.5	-8.3	-6.4	-7.4	-5.2	-4.4	0.6	-1.4
P01009	Alpha-1-antitrypsin	-6.2	-6.0	-6.0	-6.5	-4.7	-4.5	0.6	-0.8
P11413	Glucose-6-phosphate 1-dehydrogenase	-6.2	-5.3	-5.1	-5.9	-4.2	-3.6	0.6	-1.0
P52209	6-phosphogluconate dehydrogenase, decarboxylating	-2.4	-1.9	-1.3	-2.1	-0.9	0.1	0.6	-0.9
P06396	Gelsolin	-6.1	-5.6	-4.4	-5.4	-4.4	-2.7	0.6	-1.2
P06702	Protein S100-A9	-0.9	-0.9	0.2	-0.5	0.1	0.5	0.5	-0.6
P59998	Actin-related protein 2/3 complex subunit 4	-7.7	-6.6	-6.8	-6.7	-5.0	-6.9	0.5	-0.8
P13796	Plastin-2	-3.8	-3.4	-3.2	-3.9	-2.4	-1.8	0.5	-0.7
P05164	Myeloperoxidase	-0.3	0.7	0.7	-0.2	1.6	2.2	0.5	-0.9
P06744	Glucose-6-phosphate isomerase	-5.2	-4.5	-3.5	-4.8	-3.1	-2.2	0.5	-1.0
P29401	Transketolase	-5.7	-5.1	-4.2	-5.5	-4.0	-2.5	0.5	-1.0
P50395	Rab GDP dissociation inhibitor beta	-7.4	-7.1	-5.9	-7.0	-6.2	-4.6	0.5	-0.9
P08962	CD63 antigen	-5.7	-7.6	-6.3	-6.4	-6.1	-4.9	0.5	-0.8
P04083	Annexin A1	0.0	0.1	1.3	0.5	0.8	2.6	0.5	-0.8
P08493	Matrix Gla protein	-4.4	-3.3	-3.8	-5.2	-4.4	-3.6	0.4	0.6
Q9HDC9	Adipocyte plasma membrane-associated protein	-5.8	-6.5	-5.3	-5.7	-5.9	-3.6	0.4	-0.8
P09211	Glutathione S-transferase P	-3.8	-2.8	-2.4	-3.3	-2.3	-1.5	0.4	-0.6
Q9UBG3	Cornulin	-7.9	-5.4	-3.1	-4.7	-4.4	-2.9	0.4	-1.4
P61158	Actin-related protein 3	-8.3	-7.4	-6.9	-7.6	-7.4	-5.1	0.4	-0.8

		Log2 LFQ abundances						AD VS HC	
UniProt-KB Code	**Protein name**	**AD**			**HC**			**-Log10 p value**	**Fold change**
Q96DR5	BPI fold-containing family A member 2	0.0	-1.1	0.3	-1.5	-1.6	0.4	0.3	0.6
P06703	Protein S100-A6	-7.3	-6.1	-4.8	-6.9	-4.9	-3.4	0.3	-1.0
P30740	Leukocyte elastase inhibitor	-3.1	-2.5	-2.3	-3.2	-2.3	-0.7	0.3	-0.6
Q9UBC9	Small proline-rich protein 3	-5.3	-4.0	-1.9	-4.1	-3.1	-1.5	0.3	-0.9
P22079	Lactoperoxidase	-6.0	-5.2	-4.5	-6.1	-4.5	-3.4	0.3	-0.6
Q8N4F0	BPI fold-containing family B member 2	-2.0	-1.9	-1.4	-2.5	-1.4	-0.4	0.2	-0.4
P08758	Annexin A5	-7.6	-7.6	-6.4	-7.3	-7.7	-5.0	0.2	-0.5
P29373	Cellular retinoic acid-binding protein 2	-5.9	-5.8	-5.1	-6.0	-5.2	-4.8	0.2	-0.2
P28325	Cystatin-D	-3.8	-3.9	-3.5	-4.3	-4.3	-3.2	0.2	0.2
P12429	Annexin A3	-3.1	-3.1	-2.5	-3.1	-3.5	-0.7	0.2	-0.5
O15144	Actin-related protein 2/3 complex subunit 2	-7.6	-7.6	-7.0	-7.3	-8.4	-5.4	0.2	-0.4
P0DTE7	Alpha-amylase 1B	0.1	1.1	0.9	-0.6	1.9	2.0	0.2	-0.4
P01034	Cystatin-C	-5.4	-3.8	-4.2	-4.7	-4.5	-3.4	0.1	-0.3
Q9HC84	Mucin-5B	-2.0	-1.3	-1.1	-2.1	-1.3	-0.5	0.1	-0.2
P24158	Myeloblastin	-3.2	-1.7	-1.1	-2.7	-2.2	-0.2	0.1	-0.3
P17213	Bactericidal permeability-increasing protein	-1.8	-1.0	-1.0	-3.0	-0.9	-0.7	0.1	0.3
P08311	Cathepsin G	0.9	2.6	2.7	1.1	2.2	3.6	0.1	-0.3
P08246	Neutrophil elastase	1.3	2.6	3.3	1.1	2.9	4.1	0.1	-0.3
P01037	Cystatin-SN	-2.3	-1.5	-0.9	-2.4	-1.8	0.2	0.1	-0.2
Q8TAX7	Mucin-7	-1.6	-1.7	-0.6	-2.7	-1.1	-0.6	0.1	0.2
P49913	Cathelicidin antimicrobial peptide	-7.5	-6.7	-5.8	-8.1	-6.5	-4.7	0.1	-0.2
P54108	Cysteine-rich secretory protein 3	-5.0	-5.2	-4.6	-5.7	-5.4	-3.3	0.1	-0.2
P23280	Carbonic anhydrase 6	-2.6	-1.5	-1.5	-3.2	-1.8	-0.1	0.1	-0.2
P09228	Cystatin-SA	-7.2	-8.1	-6.4	-8.2	-7.8	-5.5	0.0	-0.1
P52566	Rho GDP-dissociation inhibitor 2	-5.7	-6.3	-5.0	-6.4	-6.0	-4.4	0.0	-0.1
P15104	Glutamine synthetase	-8.1	-8.6	-7.1	-7.9	-9.4	-6.3	0.0	-0.1
Q96DA0	Zymogen granule protein 16 homolog B	-5.2	-5.1	-4.8	-4.6	-3.5	-2.8	1.2	-1.4

Bibliography

1. McKhann, G. M. et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's Dement. 7, (2011).

2. Duyckaerts, C., Delatour, B. & Potier, M. C. Classification and basic pathology of Alzheimer disease. Acta Neuropathol. 118, 5–36 (2009).

3. Castellani, R. J., Rolston, R. K. & Smith, M. A. Alzheimer disease. Disease-a-Month 56, 484–546 (2010).

4. Scheltens, P. et al. Alzheimer's disease. Lancet 388, 505–517 (2016).

5. Hebert, L. E., Weuve, J., Scherr, P. A. & Evans, D. A. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. Neurology 80, 1778–1783 (2013).

6. Nichols, E. et al. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. 18, 88–106 (2019).

7. Chan, K. Y. et al. Epidemiology of alzheimer's disease and other forms of dementia in China, 1990-2010: A systematic review and analysis. Lancet 381, 2016–2023 (2013).

8. Lopez, O. L. & Kuller, L. H. Epidemiology of aging and associated cognitive disorders: Prevalence and incidence of Alzheimer's disease and other dementias. Handbook of Clinical Neurology vol. 167 (Elsevier B.V., 2019).

9. Hippius, H. & Neundörfer, G. The discovery of Alzheimer's disease. Dialogues Clin. Neurosci. 5, 101–108 (2003).

10. Bateman, R. J. et al. Autosomal-dominant Alzheimer's disease: A review and proposal for the prevention of Alzheimer's disease. Alzheimer's Res. Ther. 2, 1–13 (2011).

11. Gomez, W., Morales, R., Maracaja-Coutinho, V., Parra, V. & Nassif, M. Down syndrome and alzheimer's disease: Common molecular traits beyond the amyloid precursor protein. Aging. 12, 1011–1033 (2020).

12. Lane, C. A., Hardy, J. & Schott, J. M. Alzheimer's disease. Eur. J. Neurol. 25, 59–70 (2018).

13. Verghese, P. B., Castellano, J. M. & Holtzman, D. M. Apolipoprotein E in Alzheimer's disease and other neurological disorders. The Lancet Neurology vol. 10 241–252 (2011).

14. Varadarajan, S., Yatin, S., Aksenova, M. & Butterfield, D. A. Review: Alzheimer's amyloid β-peptide-associated free radical oxidative stress and neurotoxicity. J. Struct. Biol. 130, 184–208 (2000).

15. Hozumi, I. et al. Patterns of levels of biological metals in CSF differ among neurodegenerative diseases. J. Neurol. Sci. 303, 95–99 (2011).

16. Rahman, M. A. et al. Emerging risk of environmental factors: insight mechanisms of Alzheimer's diseases. Environ. Sci. Pollut. Res. 27, 44659–44672 (2020).

17. Mawanda, F. & Wallace, R. Can infections cause Alzheimer's disease? Epidemiol. Rev. 35, 161–180 (2013).

18. Hill, J. M. et al. Pathogenic microbes, the microbiome, and Alzheimer's disease (AD). Front. Aging Neurosci. 6, 1–5 (2014).

19. Long, J. M. & Holtzman, D. M. Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. Cell 179, 312–339 (2019).

20. Sakono, M. & Zako, T. Amyloid oligomers: Formation and toxicity of Aβ oligomers. FEBS J. 277, 1348–1358 (2010).

21. Mori, H., Takio, K., Ogawara, M. & Selkoe, D. J. Mass spectrometry of purified amyloid β protein in Alzheimer's disease. J. Biol. Chem. 267, 17082–17086 (1992).

22. Michno, W., Blennow, K., Zetterberg, H. & Brinkmalm, G. Refining the amyloid β peptide and oligomer fingerprint ambiguities in Alzheimer's disease: Mass spectrometric molecular characterization in brain, cerebrospinal fluid, blood, and plasma. J. Neurochem. 159, 234–257 (2021).

23. Shoji, M. et al. Production of the Alzheimer amyloid β protein by normal proteolytic processing. Science. 258, 126–129 (1992).

24. Dawkins, E. & Small, D. H. Insights into the physiological function of the β-amyloid precursor protein: Beyond Alzheimer's disease. J. Neurochem. 129, 756–769 (2014).

25. Brothers, H. M., Gosztyla, M. L. & Robinson, S. R. The physiological roles of amyloid-β peptide hint at new ways to treat Alzheimer's disease. Front. Aging Neurosci. 10, 1–16 (2018).

26. Saido, T. & Leissring, M. A. Proteolytic degradation of amyloid β-protein. Cold Spring Harb. Perspect. Med. 2, (2012).

27. Saito, T., Takaki, Y., Iwata, N., Trojanowski, J. & Saido, T. C. Alzheimer's disease, neuropeptides, neuropeptidase, and amyloid-beta peptide metabolism. Science of aging knowledge environment vol. 2003 PE1 (2003).

28. Hernandez, P., Lee, G., Sjoberg, M. & MacCioni, R. B. Tau phosphorylation by cdk5 and Fyn in response to amyloid peptide Aβ25-35: Involvement of lipid rafts. J. Alzheimer's Dis. 16, 149–156 (2009).

29. Zhang, H. et al. Review interaction between aβ and tau in the pathogenesis of alzheimer's disease. Int. J. Biol. Sci. 17, 2181–2192 (2021).

30. De Strooper, B. & Karran, E. The Cellular Phase of Alzheimer's Disease. Cell vol. 164 603–615 (2016).

31. Hughes, T. F., Snitz, B. E. & Ganguli, M. Should mild cognitive impairment be subtyped? Curr. Opin. Psychiatry 24, 237–242 (2011).

32. Aisen, P. S. et al. On the path to 2025: Understanding the Alzheimer's disease continuum. Alzheimer's Research and Therapy 9, 60 (2017).

33. Olsson, B. et al. CSF and blood biomarkers for the diagnosis of Alzheimer's disease: a systematic review and meta-analysis. Lancet Neurol. 15, 673–684 (2016).

34. Ashton, N. J., Ide, M., Zetterberg, H. & Blennow, K. Salivary Biomarkers for Alzheimer's Disease and Related Disorders. Neurol. Ther. 8, 83–94 (2019).

35. Schöll, M. et al. Biomarkers for tau pathology. Molecular and Cellular Neuroscience 97, 18–33 (2019).

36. Hampel, H. et al. Total and phosphorylated tau protein as biological markers of Alzheimer's disease. Experimental Gerontology 45, 30–40 (2010).

37. Zetterberg, H. Blood-based biomarkers for Alzheimer's disease—An update. Journal of Neuroscience Methods 319, 2–6 (2019).

38. Qu, Y. et al. Blood biomarkers for the diagnosis of amnestic mild cognitive impairment and Alzheimer's disease: A systematic review and meta-analysis. Neuroscience and Biobehavioral Reviews 128, 479–486 (2021).

39. Htike, T. T., Mishra, S., Kumar, S., Padmanabhan, P. & Gulyás, B. Peripheral Biomarkers for Early Detection of Alzheimer's and Parkinson's Diseases. Molecular Neurobiology 56, 2256–2277 (2019).

40. Gleerup, H. S., Hasselbalch, S. G. & Simonsen, A. H. Biomarkers for Alzheimer's disease in saliva: A systematic review. Disease Markers 2019, 4761054 (2019).

41. Livingston, G. et al. The Lancet Commissions Dementia prevention, intervention, and care. Lancet 390, 2673–2734 (2017).

42. Fillit, H. & Green, A. Aducanumab and the FDA — where are we now? Nature Reviews Neurology 17, 129–130 (2021).

43. Knopman, D. S., Jones, D. T. & Greicius, M. D. Failure to demonstrate efficacy of aducanumab: An analysis of the EMERGE and ENGAGE trials as reported by Biogen, December 2019. Alzheimer's Dement. 17, 696–701 (2021).

44. Yates, J. R. A century of mass spectrometry: From atoms to proteomes. Nature Methods 8, 633–637 (2011).

45. 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).

46. Schaffer, L. V. et al. Identification and Quantification of Proteoforms by Mass Spectrometry. Proteomics 19, e1800361 (2019).

47. Tipton, J. D. et al. Analysis of intact protein isoforms by mass spectrometry. Journal of Biological Chemistry 286, 25451–8 (2011).

48. Cupp-Sutton, K. A. & Wu, S. High-throughput quantitative top-down proteomics. Molecular Omics 16, 91–99 (2020).

49. Nakayasu, E. S. et al. Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation. Nat. Protoc. 16, 3737–3760 (2021).

50. Uchida, H. & Ovitt, C. E. Novel impacts of saliva with regard to oral health. J. Prosthet. Dent. S0022-3913(21)00273-0 (2021)

51. Pedersen, A. M. L., Sørensen, C. E., Proctor, G. B., Carpenter, G. H. & Ekström, J. Salivary secretion in health and disease. J. Oral Rehabil. 45, 730–746 (2018).

52. Cabiddu, G. et al. Proteomic characterization of the mucosal pellicle formed in vitro on a cellular model of oral epithelium. J. Proteomics 222, 103797 (2020).

53. Castagnola, M. et al. Salivary biomarkers and proteomics: Future diagnostic and clinical utilities. Acta Otorhinolaryngol. Ital. 37, 94–101 (2017).

54. Messana, I. et al. Trafficking and postsecretory events responsible for the formation of secreted

human salivary peptides. Mol. Cell. Proteomics 7, 911–926 (2008).

55. Castagnola, M. et al. Top-down platform for deciphering the human salivary proteome. Journal of Maternal-Fetal and Neonatal Medicine 25, 27–43 (2012).

56. Boroumand, M. et al. Saliva, a bodily fluid with recognized and potential diagnostic applications. J. Sep. Sci. 44, 3677–3690 (2021).

57. Amado, F. M. L., Ferreira, R. P. & Vitorino, R. One decade of salivary proteomics: Current approaches and outstanding challenges. Clin. Biochem. 46, 506–517 (2013).

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

59. Pfaffe, T., Cooper-White, J., Beyerlein, P., Kostner, K. & Punyadeera, C. Diagnostic potential of saliva: Current state and future applications. Clin. Chem. 57, 675–687 (2011).

60. Farah, R. et al. Salivary biomarkers for the diagnosis and monitoring of neurological diseases. Biomedical Journal 41, 63–87 (2018).

61. Vivacqua, G. et al. Salivary alpha-synuclein in the diagnosis of Parkinson's disease and Progressive Supranuclear Palsy. Park. Relat. Disord. 63, 143–148 (2019).

62. Corey-Bloom, J. et al. Salivary levels of total huntingtin are elevated in Huntington's disease patients. Sci. Rep. 8, 1–9 (2018).

63. Corey-Bloom, J. et al. Levels of interleukin-6 in saliva, but not plasma, correlate with clinical metrics in huntington's disease patients and healthy control subjects. Int. J. Mol. Sci. 21, 1–12 (2020).

64. Obayashi, K. et al. Salivary chromogranin A: Useful and quantitative biochemical marker of affective state in patients with amyotrophic lateral sclerosis. Intern. Med. 47, 1875–1879 (2008).

65. Ngounou Wetie, A. G. et al. A Pilot Proteomic Analysis of Salivary Biomarkers in Autism Spectrum Disorder. Autism Res. 8, 338–350 (2015).

66. 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).

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

68. El, K. et al. Saliva diagnostics-Current views and directions Impact statement. Exp. Biol. Med. 242, 459–472 (2017).

69. Spielmann, N. & Wong, D. Saliva: Diagnostics and therapeutic perspectives. Oral Dis. 17, 345–354 (2011).

70. Shen, L. et al. Proteomics analysis of blood serums from Alzheimer's disease patients using iTRAQ labeling technology. J. Alzheimer's Dis. 56, 361–378 (2017).

71. Chang, K. A., Kim, H. J. & Suh, Y. H. The role of S100a9 in the pathogenesis of Alzheimer's disease: The therapeutic effects of S100a9 knockdown or knockout. Neurodegener. Dis. 10, 27–29 (2012).

72. Watt, A. D. et al. Peripheral α-defensins 1 and 2 are elevated in Alzheimer's disease. J. Alzheimer's Dis. 44, 1131–1143 (2015).

73. Zhang, G. H., Murthy, K. D., Binti Pare, R. & Qian, Y. H. Protective effect of Tβ4 on central nervous system tissues and its developmental prospects. Eur. J. Inflamm. 18, 1–11 (2020).

74. Mathews, P. M. & Levy, E. Cystatin C in aging and in Alzheimer's disease. Ageing Research Reviews 32, 38–50 (2016).

75. Lehtinen, M. K. et al. Cystatin B deficiency sensitizes neurons to oxidative stress in progressive myoclonus epilepsy, EPM1. J. Neurosci. 29, 5910–5915 (2009).

76. Ii, K., Ito, H., Kominami, E. & Hirano, A. Abnormal distribution of cathepsin proteinases and endogenous inhibitors (cystatins) in the hippocampus of patients with Alzheimer's disease, parkinsonism-dementia complex on Guam, and senile dementia and in the aged. Virchows Archiv A Pathological Anatomy and Histopathology 423, 185–94 (1993).

77. 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).

78. Contini, C. et al. RP-HPLC-ESI-IT Mass Spectrometry Reveals Significant Variations of the Human Salivary Protein Profile Associated with Predominantly Antibody Deficiencies. J. Clin. Immunol. 40, 329–339 (2020).

79. Serrao, S. et al. Top-Down Proteomics of Human Saliva Discloses Significant Variations of the Protein Profile in Patients with Mastocytosis. J. Proteome Res. 19, 3238–3253 (2020).

80. 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, (1998).

81. Deutsch, E. W. et al. The ProteomeXchange consortium in 2020: Enabling "big data" approaches in proteomics. Nucleic Acids Res. 48, 1145–1152 (2020).

82. Ghallab, N. A. Diagnostic potential and future directions of biomarkers in gingival crevicular fluid and saliva of periodontal diseases: Review of the current evidence. Archives of Oral Biology 87, 115–124 (2018).

83. Iavarone, F. et al. Top down proteomic analysis of gingival crevicular fluid in deciduous, exfoliating and permanent teeth in children. J. Proteomics 226, 10389 (2020).

84. Grassl, N. et al. Ultra-deep and quantitative saliva proteome reveals dynamics of the oral microbiome. Genome Med. 8, 44 (2016).

85. Gazzar, M. El. Immunobiology of S100A8 and S100A9 Proteins and Their Role in Acute Inflammation and Sepsis. Int. J. Immunol. Immunother. 2, 2 (2015).

86. Cristóvaõ, J. S. & Gomes, C. M. S100 Proteins in Alzheimer's Disease. Frontiers in Neuroscience 13, 463 (2019).

87. Lodeiro, M. et al. Aggregation of the inflammatory s100a8 precedes a? plaque formation in transgenic app mice: Positive feedback for s100a8 and a? productions. Journals Gerontol. - Ser. A Biol. Sci. Med. Sci. 72, 319–328 (2017).

88. Kummer, M. P. et al. Mrp14 deficiency ameliorates amyloid β burden by increasing microglial phagocytosis and modulation of amyloid precursor protein processing. J. Neurosci. 32, 17824–17829 (2012).

89. Shepherd, C. E. et al. Inflammatory S100A9 and S100A12 proteins in Alzheimer's disease. Neurobiol. Aging 27, 1554–1563 (2006).

90. Ha, T. Y. et al. S100a9 knockdown decreases the memory impairment and the neuropathology in Tg2576 mice, AD animal model. PLoS One 5, e8840 (2010).

91. Smith, M. A., Harris, P. L. R., Sayre, L. M. & Perry, G. Iron accumulation in Alzheimer disease is a source of redox-generated free radicals. Proc. Natl. Acad. Sci. U. S. A. 94, 9866–9868 (1997).

92. Duncan, A. J. & Heales, S. J. R. Nitric oxide and neurological disorders. Molecular Aspects of Medicine 26, 67–96 (2005).

93. Wang, W. Y., Tan, M. S., Yu, J. T. & Tan, L. Role of pro-inflammatory cytokines released from microglia in Alzheimer's disease. Annals of Translational Medicine 3, 136 (2015).

94. Tarafdar, A. & Pula, G. The role of NADPH oxidases and oxidative stress in neurodegenerative disorders. International Journal of Molecular Sciences 19, 3824 (2018).

95. Benedyk, M. et al. HaCaT keratinocytes overexpressing the S100 proteins S100A8 and S100A9 show increased NADPH oxidase and NF-κB activities. J. Invest. Dermatol. 127, 2001–2011 (2007).

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

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

98. Alonso, R. et al. Fungal infection in patients with Alzheimer's disease. J. Alzheimer's Dis. 41, 301–311 (2014).

99. Lövheim, H. et al. Herpes simplex infection and the risk of Alzheimer's disease: A nested case-control study. Alzheimer's Dement. 11, 587–592 (2015).

100. Welling, M. M., Nabuurs, R. J. A. & Van Der Weerd, L. Potential role of antimicrobial peptides in the early onset of Alzheimer's disease. Alzheimer's Dement. 11, 51–57 (2015).

101. Williams, W. M. et al. Antimicrobial peptide β-defensin-1 expression is upregulated in Alzheimer's brain. J. Neuroinflammation 10, 127 (2013).

102. Soscia, S. J. et al. The Alzheimer's disease-associated amyloid β-protein is an antimicrobial peptide. PLoS One 5, e9505 (2010).

103. Oppenheim, J. J. & Yang, D. Alarmins: Chemotactic activators of immune responses. Current Opinion in Immunology 17, 359–365 (2005).

104. Carion, T. W. et al. Antimicrobial effects of thymosin beta-4 and ciprofloxacin adjunctive therapy in pseudomonas aeruginosa induced keratitis. Int. J. Mol. Sci. 21, 1–11 (2020).

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

106. Chopp, M. & Zhang, Z. G. Thymosin β4 as a restorative/regenerative therapy for neurological injury and neurodegenerative diseases. Expert Opin. Biol. Ther. 15, 9–12 (2015).

107. Magister, Š. & Kos, J. Cystatins in immune system. Journal of Cancer 4, 45–56 (2013).

108. Bernstein, H. G. et al. Cystatin a-like immunoreactivity is widely distributed in human brain and accumulates in neuritic plaques of alzheimer disease subjects. Brain Res. Bull. 33, 477–481 (1994).

109. Soond, S. M., Kozhevnikova, M. V., Zamyatnin, A. A. & Townsend, P. A. Cysteine cathepsin protease inhibition: An update on its diagnostic, prognostic and therapeutic potential in cancer. Pharmaceuticals 12, 87 (2019).

110. Wu, Z. et al. Cathepsin B plays a critical role in inducing Alzheimer's disease-like phenotypes following chronic systemic exposure to lipopolysaccharide from Porphyromonas gingivalis in mice. Brain. Behav. Immun. 65, 350–361 (2017).

111. Hook, G., Hook, V. & Kindy, M. The cysteine protease inhibitor, E64d, reduces brain amyloid-β and improves memory deficits in alzheimer's disease animal models by inhibiting cathepsin B, but not BACE1, β-secretase activity. J. Alzheimer's Dis. 26, 387–408 (2011).

112. Butinar, M. et al. Stefin B deficiency reduces tumor growth via sensitization of tumor cells to oxidative stress in a breast cancer model. Br. Dent. J. 217, 3392–3400 (2014).

113. Sun, T., Turk, V., Turk, B. & Kopitar-Jerala, N. Increased expression of stefin B in the nucleus of T98G astrocytoma Cells Delays caspase activation. Frontiers in Molecular Neuroscience 5, 93 (2012).

114. Torres, P., Castro, M., Reyes, M. & Torres, V. A. Histatins, wound healing, and cell migration. Oral Diseases 24, 1150–1160 (2018).

115. Sugiyama, K. Anti-lipopolysaccharide activity of histatins, peptides from human saliva. Experientia 49, 1095–1097 (1993).

116. Wang, G. Human antimicrobial peptides and proteins. Pharmaceuticals 7, 545–594 (2014).

117. Melino, S., Santone, C., Di Nardo, P. & Sarkar, B. Histatins: Salivaryx peptides with copper(II)- and zinc(II)-binding motifs Perspectives for biomedical applications. FEBS J. 281, 657–672 (2014).

118. Huang, X. et al. Zinc-induced Alzheimer's Aβ1-40 aggregation is mediated by conformational factors. J. Biol. Chem. 272, 26464–26470 (1997).

119. Kochańska, B., Kędzia, A., Kamysz, W., Maćkiewicz, Z. & Kupryszewski, G. The effect of statherin and its shortened analogues on anaerobic bacteria isolated from the oral cavity. Acta Microbiol. Pol. 49, 243–251 (2000).

120. Amano, A., Kataoka, K., Raj, P. A., Genco, R. J. & Shizukuishi, S. Binding sites of salivary statherin for Porphyromonas gingivalis recombinant fimbrillin. Infect. Immun. 64, 4249–4254 (1996).

121. Ryder, M. I. Porphyromonas gingivalis and Alzheimer disease: Recent findings and potential therapies. J. Periodontol. 91, S45–S49 (2020).

122. Poole, S., Singhrao, S. K., Kesavalu, L., Curtis, M. A. & Crean, S. J. Determining the presence of periodontopathic virulence factors in short-term postmortem Alzheimer's disease brain tissue. J. Alzheimer's Dis. 36, 665–677 (2013).

123. Dominy, S. S. et al. Porphyromonas gingivalis in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. Sci. Adv. 5, (2019).

124. Mancini, M., Grappasonni, I., Scuri, S. & Amenta, F. Oral Health in Alzheimers Disease: A Review. Curr. Alzheimer Res. 7, 368–373 (2010).

125. Ship, J. A., DeCarli, C., Friedland, R. P. & Baum, B. J. Diminished submandibular salivary flow in dementia of the Alzheimer Type. Journals Gerontol. 45, M61–6 (1990).

126. Deary, I. J. et al. Age-associated cognitive decline. Br. Med. Bull. 92, 135–152 (2009).

127. Campisi, J. et al. From discoveries in ageing research to therapeutics for healthy ageing. Nature 571, 183–192 (2019).

128. Hoffman, J. M., Lyu, Y., Pletcher, S. D. & Promislow, D. E. L. Proteomics and metabolomics in ageing research: From biomarkers to systems biology. Essays Biochem. 61, 379–388 (2017).

129. Rivero-Segura, N. A., Bello-Chavolla, O. Y., Barrera-Vázquez, O. S., Gutierrez-Robledo, L. M. & Gomez-Verjan, J. C. Promising biomarkers of human aging: In search of a multi-omics panel to understand the aging process from a multidimensional perspective. Ageing Res. Rev. 64, 101164 (2020).

130. Johnson, A. A., Shokhirev, M. N., Wyss-Coray, T. & Lehallier, B. Systematic review and analysis of human proteomics aging studies unveils a novel proteomic aging clock and identifies key processes that change with age. Ageing Res. Rev. 60, 101070 (2020).

131. Santos, A. L. & Lindner, A. B. Protein Posttranslational Modifications: Roles in Aging and Age-Related Disease. Oxid. Med. Cell. Longev. 2017, 1–19 (2017).

132. Cabras, T. et al. Age-dependent modifications of the human salivary secretory protein complex. J. Proteome Res. 8, 4126–4134 (2009).

133. Castagnola, M. et al. The surprising composition of the salivary proteome of preterm human newborn. Mol. Cell. Proteomics 10, M110.003467 (2011).

134. Morzel, M. et al. Saliva electrophoretic protein profiles in infants: Changes with age and impact of teeth eruption and diet transition. Arch. Oral Biol. 56, 634–642 (2011).

135. Manconi, B. et al. Modifications of the acidic soluble salivary proteome in human children from birth to the age of 48months investigated by a top-down HPLC-ESI-MS platform. J. Proteomics 91, 536–543 (2013).

136. Messana, I. et al. Chrono-proteomics of human saliva: Variations of the salivary proteome during human development. J. Proteome Res. 14, 1666–1677 (2015).

137. Johnson, D. A., Yeh, C. K. & Dodds, M. W. J. Effect of donor age on the concentrations of histatins in human parotid and submandibular/sublingual saliva. Arch. Oral Biol. 45, 731–740 (2000).

138. Denny, P. C. et al. Age-related Changes in Mucins from Human Whole Saliva. J. Dent. Res. 70, 1320–1327 (1991).

139. Chang, W. I., Chang, J. Y., Kim, Y. Y., Lee, G. & Kho, H. S. MUC1 expression in the oral mucosal epithelial cells of the elderly. Arch. Oral Biol. 56, 885–890 (2011).

140. Tanida, T. et al. Influence of aging on candidal growth and adhesion regulatory agents in saliva. J. Oral Pathol. Med. 30, 328–335 (2001).

141. Salvolini, E. et al. Age-related modifications in human unstimulated whole saliva: A biochemical study. Aging Clin. Exp. Res. 12, 445–448 (2000).

142. Nagler, R. M. & Hershkovich, O. Relationships between age, drugs, oral sensorial complaints and salivary profile. Arch. Oral Biol. 50, 7–16 (2005).

143. Manconi, B. et al. Top-down HPLC-ESI–MS proteomic analysis of saliva of edentulous subjects evidenced high levels of cystatin A, cystatin B and SPRR3. Arch. Oral Biol. 77, 68–74 (2017).

144. François, M., Bull, C. F., Fenech, M. F. & Leifert, W. R. Current State of Saliva Biomarkers for Aging and Alzheimer's Disease. Curr. Alzheimer Res. 16, 56–66 (2018).

145. Xu, F., Laguna, L. & Sarkar, A. Aging-related changes in quantity and quality of saliva: Where do we stand in our understanding? J. Texture Stud. 50, 27–35 (2019).

146. Contini, C. et al. Top-Down Proteomics of Human Saliva Highlights Anti-inflammatory, Antioxidant, and Antimicrobial Defense Responses in Alzheimer Disease. Front. Neurosci. 15, 668852 (2021).

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

148. 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).

149. Yoav Benjamini & Yosef Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B 57, 289–300 (1995).

150. Kursa, M. B., Jankowski, A. & Rudnicki, W. R. Boruta - A system for feature selection. Fundam. Informaticae 101, 271–285 (2010).

151. Sun, S. et al. Analysis of age and gender associated n-glycoproteome in human whole saliva. Clin. Proteomics 11, 1–10 (2014).

152. Scott, J., Flower, E. A. & Burns, J. A quantitative study of histological changes in the human parotid gland occurring with adult age. J. Oral Pathol. Med. 16, 505–510 (1987).

153. Vissink, A., Spijkervet, F. K. L. & Van Nieuw Amerongen A. Aging and saliva : A review of the literature parotid gland submandibular gland. Spec. Care Dent. 16, 95–103 (1996).

154. Vissink, A. et al. Clinical management of salivary gland hypofunction and xerostomia in head-and-neck cancer patients: Successes and barriers. Int. J. Radiat. Oncol. Biol. Phys. 78, 983–991 (2010).

155. Gomes, L. H. et al. S100A8 and S100A9 - Oxidant scavengers in inflammation. Free Radic. Biol. Med. 58, 170–186 (2013).

156. Žerovnik, E. Putative alternative functions of human stefin B (cystatin B): binding to amyloid-beta, membranes, and copper. J. Mol. Recognit. 30, 1–7 (2017).

157. 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).

158. Kopitar-Jerala, N. The role of cystatins in cells of the immune system. FEBS Lett. 580, 6295–6301 (2006).

159. 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).

160. Žerovnik, E., Staniforth, R. A. & Turk, D. Amyloid fibril formation by human stefins: Structure, mechanism & putative functions. Biochimie 92, 1597–1607 (2010).

161. 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).

162. Ganz, T. Defensins: Antimicrobial peptides of innate immunity. Nat. Rev. Immunol. 3, 710–720 (2003).

163. Williams, W. M., Castellani, R. J., Weinberg, A., Perry, G. & Smith, M. A. Do β-defensins and other antimicrobial peptides play a role in neuroimmune function and neurodegeneration? The Scientific World Journal 2012, 905785 (2012).

164. Wang, S. et al. S100A8/A9 in inflammation. Front. Immunol. 9, 1298 (2018).

165. Kielland, A. et al. In vivo imaging of reactive oxygen and nitrogen species in inflammation using the luminescent probe L-012. Free Radic. Biol. Med. 47, 760–766 (2009).

166. Miranda, L. P. et al. Total chemical synthesis and chemotactic activity of human S100A12 (EN-RAGE). FEBS Lett. 488, 85–90 (2001).

167. Hofmann, M. A. et al. RAGE Mediates a Novel Proinflammatory Axis. Cell 97, 889–901 (1999).

168. 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).

169. Swindell, W. R. et al. Robust shifts in S100a9 expression with aging: A novel mechanism for chronic inflammation. Sci. Rep. 3, 1–13 (2013).

170. Beach, T. G., Monsell, S. E., Phillips, L. E. & Kukull, W. Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005-2010. J. Neuropathol. Exp. Neurol. 71, 266–273 (2012).

171. Kim, K. et al. Clinically accurate diagnosis of Alzheimer's disease via multiplexed sensing of core biomarkers in human plasma. Nat. Commun. 11, 1–9 (2020).

172. Liang, Q. et al. Metabolomics-based screening of salivary biomarkers for early diagnosis of Alzheimer's disease. RSC Adv. 5, 96074–96079 (2015).

173. Nagler, R. M. Salivary glands and the aging process: Mechanistic aspects, health-status and medicinal-efficacy monitoring. Biogerontology 5, 223–233 (2004).

174. Stoka, V., Turk, V. & Turk, B. Lysosomal cathepsins and their regulation in aging and neurodegeneration. Ageing Research Reviews 32, 22–37 (2016).

175. Cipollini, E. et al. Cystatin B and its EPM1 mutants are polymeric and aggregate prone in vivo. Biochim. Biophys. Acta - Mol. Cell Res. 1783, 312–322 (2008).

176. Ulbrich, L. et al. Cystatin B and SOD1: Protein-protein interaction and possible relation to neurodegeneration. Cell. Mol. Neurobiol. 34, 205–213 (2014).

177. Di Giaimo, R. et al. New insights into the molecular basis of progressive myoclonus epilepsy: A multiprotein complex with cystatin B. Hum. Mol. Genet. 11, 2941–2950 (2002).

178. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685 (1970).

179. Tyanova, S. & Cox, J. Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research. in Methods in Molecular Biology 1711, 133–148 (2018).

180. Mellacheruvu, D. et al. The CRAPome: A contaminant repository for affinity purification-mass spectrometry data. Nat. Methods 10, 730–736 (2013).

181. Szklarczyk, D. et al. The STRING database in 2021: Customizable protein-protein networks,

and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 49, D605–D612 (2021).

182. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37, 1–13 (2009).

183. Raudvere, U. et al. G:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 47, W191–W198 (2019).

184. Guard, S. E., Ebmeier, C. C. & Old, W. M. Label-free immunoprecipitation mass spectrometry workflow for large-scale nuclear interactome profiling. J. Vis. Exp. 2019, 153 (2019).

185. Stubbs, M. T. et al. The refined 2.4 Å X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: A novel type of proteinase inhibitor interaction. EMBO J. 9, 1939–1947 (1990).

186. 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).

187. Jenko Kokalj, S. et al. Essential Role of Proline Isomerization in Stefin B Tetramer Formation. J. Mol. Biol. 366, 1569–1579 (2007).

188. Pol, E. & Björk, I. Role of the single cysteine residue, Cys 3, of human and bovine cystatin B (stefin B) in the inhibition of cysteine proteinases. Protein Sci. 10, 1729–1738 (2001).

189. Dos Remedios, C. G. et al. Actin binding proteins: Regulation of cytoskeletal microfilaments. Physiol. Rev. 83, 433–473 (2003).

190. Dalli, J. et al. Heterogeneity in neutrophil microparticles reveals distinct proteome and functional properties. Mol. Cell. Proteomics 12, 2205–2219 (2013).

191. Othman, A., Sekheri, M. & Filep, J. G. Roles of neutrophil granule proteins in orchestrating inflammation and immunity. FEBS J. (2021) doi:10.1111/febs.15803.

192. Stendahl, O. Actin Assembly and Regulation of Neutrophil Function : Effects of Cytochalasin B and Tetracaine on Chemotactic Peptide-Induced O Production and Degranulation. J. Leukoc. Biol. 244, 236–244 (1991).

193. Lacy, P. Mechanisms of degranulation in neutrophils. Allergy, Asthma Clin. Immunol. 2, 98–108 (2006).

194. Teahan, C. G., Totty, N. F. & Segal, A. W. Isolation and characterization of grancalcin, a novel 28 kDa EF-hand calcium-binding protein from human neutrophils. Biochem. J. 286, 549–554 (1992).

195. Xu, P., Roes, J., Segal, A. W. & Radulovic, M. The role of grancalcin in adhesion of neutrophils. 240, 116–121 (2006).

196. Yamaguchi, F. et al. S100 proteins modulate protein phosphatase 5 function: A link between Ca 2+ signal transduction and protein dephosphorylation. J. Biol. Chem. 287, 13787–13798 (2012).

197. Sun, H. Q., Yamamoto, M., Mejillano, M. & Yin, H. L. Gelsolin, a multifunctional actin regulatory protein. J. Biol. Chem. 274, 33179–33182 (1999).

198. Janji, B. et al. Phosphorylation on Ser5 increases the F-actin-binding activity of L-plastin and

promotes its targeting to sites of actin assembly in cells. J. Cell Sci. 119, 1947–1960 (2006).

199. 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).

200. Delclaux, C. et al. Role of Gelatinase B and Elastase in Human Polymorphonuclear Neutrophil Migration across Basement Membrane. Am. J. Respir. Cell Mol. Biol. 14, 288–295 (1996).

201. Barber, D. F., Faure, M. & Long, E. O. LFA-1 Contributes an Early Signal for NK Cell Cytotoxicity. J. Immunol. 173, 3653–3659 (2004).

202. DiScipio, R. G., Daffern, P. J., Schraufstätter, I. U. & Sriramarao, P. Human polymorphonuclear leukocytes adhere to complement factor H through an interaction that involves alphaMbeta2 (CD11b/CD18). J. Immunol. 160, 4057–66 (1998).

203. M. Hunt, J. & Tuder, R. Alpha 1 Anti-Trypsin: One Protein, Many Functions. Curr. Mol. Med. 12, 827–835 (2012).

204. de Koning, P. J. A. et al. Intracellular serine protease inhibitor SERPINB4 inhibits granzyme M-induced cell death. PLoS One 6, 1–9 (2011).

205. Fujimoto, N. et al. Extracellular matrix protein 1 inhibits the activity of matrix metalloproteinase 9 through high-affinity protein/protein interactions. Exp. Dermatol. 15, 300–307 (2006).

206. Hoyer, S., Oesterreich, K. & Wagner, O. Glucose metabolism as the site of the primary abnormality in early-onset dementia of Alzheimer type? J. Neurol. 253, 143–148 (1988).

207. Tajes, M. et al. Methylglyoxal reduces mitochondrial potential and activates Bax and caspase-3 in neurons: Implications for Alzheimer's disease. Neurosci. Lett. 580, 78–82 (2014).

208. Tajes, M. et al. Methylglyoxal produced by amyloid-β peptide-induced nitrotyrosination of triosephosphate isomerase triggers neuronal death in alzheimer's disease. J. Alzheimer's Dis. 41, 273–288 (2014).

209. Olah, J. et al. Triosephosphate isomerase deficiency : consequences of an inherited mutation at mRNA , protein and metabolic levels. Biochem. J. 683, 675–683 (2005).

210. Lennartsson, A., Pieters, K., Vidovic, K. & Gullberg, U. A murine antibacterial ortholog to human bactericidal/permeability-increasing protein (BPI) is expressed in testis, epididymis, and bone marrow. J. Leukoc. Biol. 77, 369–377 (2005).

211. Brown, R. B. & Hollingsworth, M. A. Mucin family of glycoproteins. Encycl. Biol. Chem. Second Ed. 2, 200–204 (2013).

212. Munroe, P. B. et al. Mutations in the gene encoding the human matrix Gla protein cause Keutel syndrome. Nat. Genet. 21, 142–144 (1999).

213. Freitas, A. et al. Bioinformatic analysis of the human brain extracellular matrix proteome in neurodegenerative disorders. Eur. J. Neurosci. 53, 4016–4033 (2021).

214. Rasmussentll, H. H., Flintll, T. & Krusell, T. A. Cloning , Expression , and Chromosome Mapping of Human Galectin-7. 270, 5823–5829 (1995).

215. Sewgobind, N. V., Albers, S. & Pieters, R. J. Functions and inhibition of galectin-7, an emerging target in cellular pathophysiology. Biomolecules 11, 1720 (2021).

216. Heinrich, C. et al. Copine-III interacts with ErbB2 and promotes tumor cell migration. Oncogene 29, 1598–1610 (2010).

217. Salih, D. A. et al. Genetic variability in response to amyloid beta deposition influences Alzheimer's disease risk. Brain Commun. 1, 1–13 (2019).

218. Kaur, G. & Levy, E. Cystatin C in Alzheimer's disease. Frontiers in Molecular Neuroscience 79 (2012) doi:10.3389/fnmol.2012.00079.

219. Baba, M., Ichinose, K., Tamai, M., Kawakami, A. & Ohyama, K. Similarity of autoimmune diseases based on the profile of immune complex antigens. Rheumatol. Int. 39, 323–325 (2019).

220. Aibara, N. et al. Proteomic approach to profiling immune complex antigens in cerebrospinal fluid samples from patients with central nervous system autoimmune diseases. Clin. Chim. Acta 484, 26–31 (2018).

221. Yamane, K. et al. Immune complexome analysis reveals the presence of immune complexes and identifies disease-specific immune complex antigens in saliva samples from patients with Sjögren's syndrome. Clin. Exp. Immunol. 204, 212–220 (2021).

222. Shmagel, K. V. & Chereshnev, V. A. Molecular bases of immune complex pathology. Biochem. 74, 469–479 (2009).

223. Kaname Ohyama, Yukitaka Ueki, A., Kawakami, Naoya Kishikawa, M. T., Makoto Osaki, Shimeru Kamihira, K. & Kuroda, N. and N. Immune Complexome Analysis of Serum and Its Application in Screening for Immune Complex Antigens in Rheumatoid Arthritis. Clin. Chem. 57, 905–909 (2011).

Acknowledgments

This thesis contains 220 pages of life, with double line spacing, to facilitate reading. It took 1185 days of work to be written, some happy and full of enthusiasm, others less, made up of fears and uncertainties. PhD is the highest qualification that can be achieved, even if it is more a pot of ingredients that can hardly be found in books. If we could prepare a PhD solution I would tell you to take a big cylinder and start with 100 mL of commitment, 100 mL of organization and 100 mL of patience, and then bring it up to volume with a lot of passion. In this way, in the blink of an eye, three years, a pandemic and an extension later it's time for acknowledgments.

I would like to thank all the people who contributed to my scientific education, starting with Prof. Massimo Castagnola and Prof. Irene Messana, pillars of our research group. Thanks to Barbara, a person full of life from whom I took much of my enthusiasm and curiosity. Thanks to Prof. Olianas, the world champion of generosity and support, in and out of work. Thanks to Prof. Turck for welcoming me into his research group in Munich, to Dr. Giuseppina Maccarrone for her guidance and to my colleagues Karin, Yan Yu, Jinqiu and Bozo. A sincere thanks to all of them, for their invaluable collaboration and for the unconditional affection they have reserved to me. Finally, the biggest thanks goes to my tutor, Prof. Cabras, for making me fall in love with biochemistry and proteomics, but above all for having believed in me all this time.

They say teamwork divides tasks and multiplies successes, so the first person I would like to mention among my colleagues is my personal stress management professor: Dr. Simone Serrao. I just want to thank him for always being by my side as a colleague and as a friend, for coffee breaks, for understanding and for all the times we gave each other strength.

A sincere thanks also goes to two special people who I met by chance, during an English course, long before starting the PhD who then became so precious to me: Mariano and Giulia.

And to another special person that brightened my life and multiplied my passion for this job: Neeta. Thank you for sharing this journey with me, for the support, the laughs, the listening and the after work together.