Università degli Studi di Cagliari

Philosophy Doctor

in

Life, Environmental and Drug Sciences
Biomedical Curriculum

XXX Cycle

Salivary biomarkers in Multiple Sclerosis and Autoimmune Hepatitis explored by an integrate top-down and bottom-up platform

BIO/10

Presented by: Doctor Barbara Liori
PhD coordinator: Professor Enzo Tramontano
Tutor: Doctor Alessandra Olianas

Final exam academic year 2016 – 2017

Thesis presented in the examination session February to March 2018

CONTENTS

ABSTRACT

Multiple Sclerosis and Autoimmune Hepatitis are serious diseases whose diagnosis is extremely difficult. To date, no causes have been found in the manifestation of both pathologies, but several factors that lead to their progression such as infectious agents, environment and ethnicity and genetic predisposition. In this study, the salivary proteome and peptidome of affected individuals from these pathologies has been explored by mass spectrometry, through a top-down and bottom-up platform integrated and compared with groups of healthy controls, with the aim to assess whether qualitative and quantitative changes in proteins and salivary peptides could be associated with immune defenses distinctive imbalance of any illness and in order to have suggestions of specific potential salivary biomarkers.

The comparative analysis of salivary proteome in Multiple Sclerosis patients with respect to controls allowed the identification and the structural characterization of new proteoforms of salivary cystatins never detected before in saliva. Moreover, this study highlighted quantitative alterations at the level of different peptides and proteins of specific glands secretion as well as some proteins non-specifically detectable in oral cavity.

The proteoforms detected and characterized in saliva for the first time during this study were cystatin A Thr_{96}→Met and its acetylated derivative; cystatin B N-terminally acetylated and CMC at Cys_3; N-terminally truncated cystatin D with the N-terminal Q converted to pyro-E and lacking the first 5 amino acid residues (pGlu-cystatin D Cys_{26}→R Des_{1-5}); N-terminally truncated forms of cystatin SN and SN P_{11}→L lacking the first 4 amino acids (cystatin SN Des_{1-4} and cystatin SN P_{11}→L Des_{1-4}) and the first 7 amino acids (cystatin SN Des_{1-7} and cystatin SN P_{11}→L Des_{1-7}); N-terminally truncated cystatin SA lacking the first 7 amino acids (cystatin SA Des_{1-7}); oxidized derivatives of cystatins SN and S1 at W_{23} and W_{107}.

The quantitative analysis on Multiple Sclerosis subject performed on 102 salivary peptides/proteins, showed a high number of statistically variated proteins belonging to cystatins family. Among these, cystatin A Thr_{96}→Met, cystatin SN Des_{1-4} and SN and P_{11}→L; oxidized derivatives of cystatins SN and S1 were also found with altered level in Multiple Sclerosis with respect to controls group. Moreover, a higher number of protein statistically variated were found among those not specifically secreted from salivary glands such as S100A7, S100A8-SNO, antileukoproteinase and ASVD.

This preliminary comparative analysis of salivary proteome in Autoimmune Hepatitis patients with respect to controls allowed to determine that the levels of proteins and peptides secreted by salivary glands, such as S-type cystatins, histatins and statherins and their naturally occurring proteoforms deriving from post-translational modifications were higher in autoimmune subjects respect to control. Further studies will have to be carried out to better explain the high overexpression of proteins involved in the protection of the oral cavity in subjects affected by autoimmune hepatitis. A possible cause could be the simultaneous presence of autoimmune diseases involving the oral cavity in half of patients with Autoimmune Hepatitis recruited for this study. The concomitance of these pathologies could either have damage the oral mucosa more or modified the natural bacterial flora present in the oral cavity or both, generating an over-expression by the protein classes involved in the protection of the oral cavity and explaining its high presence.

1. Introduction

1.1 Proteomics

Omic technologies like genomics (analysis of genes), transcriptomics (analysis of mRNA) proteomics (analysis of proteins), and metabolomics (analysis of metabolites) represent strategies to monitor a variety of molecular and organismal processes and these tecniques have been widely applied to characterize complex biochemical systems and to study pathophysiological processes.

Proteomics is a large-scale study of proteins (Anderson & Anderson, 1998) and refers to a set of proteins produced in an organism, system, or biological context. It is not constant since it differs from cell to cell and changes over time. Includes a high dynamic range of protein expression and degradation, and the complexity is increased by a plethora of post-translational modifications (PTMs), and sequence variations wich make such analyses challenging (Gregorich & Ge, 2014).

The revolution in proteomics and systems biology is driven by the development of powerful mass-spectrometry-based methods (Aebersold & Mann, 2016) that are both fast and sensitive and has become the method of choice for rapidly identify proteins and determine details of their primary structures (Aebersold & Mann, 2003). Information derived from these technologies is generally more specific compared to classical biochemical methods, e.g. immunological approaches.

Today mass-spectrometry-based methods investigations have been supplemented by the development of many experimental platforms, that may be classified mainly in top-down and bottom-up (Cui et al., 2011), each one characterized by advantages and drawbacks.

Bottom-up strategy is a common method to identify proteins and characterize their amino acid sequences by proteolytic digestion of proteins, by using specific enzymes tipically trypsin, prior to analysis by mass spectrometry. This approach requires a pre-purification step to purify the selected protein from the complex mixture which can be done by different methods such as gel electrophoresis, or by gel-free-approaches like liquid chromatography. Alternatively, the the protein mixture is digested directly without any pre-purification step and further analyzed by liquid chromatography coupled to mass spectrometry: this different technique is called “shotgun” proteomics.

The main drawbacks of bottom-up/shotgun strategies is that some fragments, generated by trypsin digestion, can be too small and their structural and PTMs information can be

lost during mass spectrometry analysis; however they have the advantage to allow greater protein sequence coverage than top-down approach.

Top-down approach focuses on the detection and characterization of the intact proteins and does not use proteolytic digestion: in this case the main advantage consists in obtaining more information about PTMs (Hubbard & Cohen, 1993; Mann & Jensen, 2003; Uy & Wold, 1977), which are important in the comprehension of protein biological functions (Bogdanov & Smith, 2005). Top-down platforms are intrinsically limited by the sample treatments necessary for coupling with mass spectrometry (typically treatment with formic acid or trifluoroacetic acid), which inevitably excludes proteins that are insoluble in acidic solution. Moreover, intact high-molecular weight proteins and heterogeneous glycosylated proteins are not accessible, in their naturally occurring forms, even to the best high-level MS apparatus.

Moreover, the integration of bottom-up and top-down approaches can be applied to the analysis of large protein fragments but also avoids redundant peptide sequences. Describing and understanding the complete and quantitative proteome as well as its structure, function and dynamics is a central and fundamental challenge of biology.

In a different point of view, proteomic platforms can be classified in qualitative and quantitative. Qualitative platforms define the complete set of proteins present in a sample, PTMs comprised. Quantitative platforms can be further classified into two groups: relative quantification, which compares the amounts of selected proteins or whole proteomes between samples and allows a quantitative ratio or relative change, and absolute quantification, which provides information about the absolute amount or the concentration of a protein within a sample. A powerful quantitative proteomic method to study relative peptide abundances between two or more biological samples can be obtained by label-free LC-MS profiling, in which the samples retain their native isotope composition (Nikolov, Schmidt, & Urlaub, 2012). Moreover, label- free methods permit direct comparison of multiple samples across multiple conditions (Wang et al., 2003), allowing complex experimental designs. These characteristics make label-free mass spectrometry methods appropriate for large scale biomarker studies.

Technology platforms incorporating mass spectrometry for proteomic biomarker discovery include both pattern-based methods that produce MS-derived protein pattern via surface-enhanced laser desorption-ionization (SELDI) (Petricoin et al., 2002), matrix-assisted laser desorption–ionization (MALDI) (Villanueva et al., 2004) or electrospray (VerBerkmoes et al., 2002) and identity-based methods that yield lists of sequence

identified peptides from liquid chromatography (LC)-MS/MS analysis of proteolytically digested proteins (Tirumalai et al., 2003; Wang et al., 2003).

All mass spectrometers, regardless of type, ionization mode or performance characteristics, produce mass spectra, which plot the mass-to-charge *ratio* of the ions observed (*x* axis) versus detected ion abundance (*y* axis). Subsequent interpretation of spectra, using parameters such as isotope distribution and/or accurate mass and amino acid sequence information (in tandem MS (MS/MS) experiments), may allow portions of the spectrum to be labeled with a protein or peptide identity. In this regard mass spectrometry offers powerfull advantage for the biomarker discovery, characterization and evaluation, because of its capacity of globally examining the protein expression profiles under given conditions.

The term biomarker generally refers to a measurable indicator of some biological condition.

There is a notable difference between biomarker discovery and biomarker validation; the discovery of a biomarker is challenging and expensive (requiring expensive equipment, trained personnel, and precious biological specimens) and the clinical validation component can be even more challenging, with coordinating the implementation of a technology across many geographic locations and recruiting many patients to test biomarker robustness (Crutchfield et al., 2016). The discovery of biomarkers use mass spectrometry as the principal technique (Rifai et al., 2006) to analyse proteins in the human proteome and to produce a list of candidate protein biomarkers each of which can be differentially expressed in disease and control samples. After the discovery of the potential biomarker, it's necessary to verify if it is possible use it such as a "true" biomarker: candidates identified using methods and biological materials optimized for discovery are translated to methods and materials suitable for verification. The next phase of biomarker development is validation and clinical assay development, which typically requires measurement of thousands of patient samples with single-digit measurement coefficient of variation values. A platform change is again required, because mass spectrometry is currently not able to achieve the combination of high throughput with high measurement accuracy and precision, so that it is not yet routinely accepted for such tests by the FDA (Kingsmore, 2006). Thus, this mandates toward the development of suitable antibodies for each biomarker candidate quantification, and in fact immunotechniques, such as radioimmunoassay (RIA) or ELISA offer higher level of

sensitivity compared with more sophisticated nonimmuno-based technologies, such as mass spectrometry (Vitzthum, Behrens, Anderson, & Shaw, 2005).

For these reasons the opportunity for the clinical application of mass spectrometry will likely not be "protein based", but rather utilize the investigation of PTMs, the presence or concentration of small molecule metabolites, or profiling metabolic flux (Crutchfield, Thomas, Sokoll, & Chan, 2016).

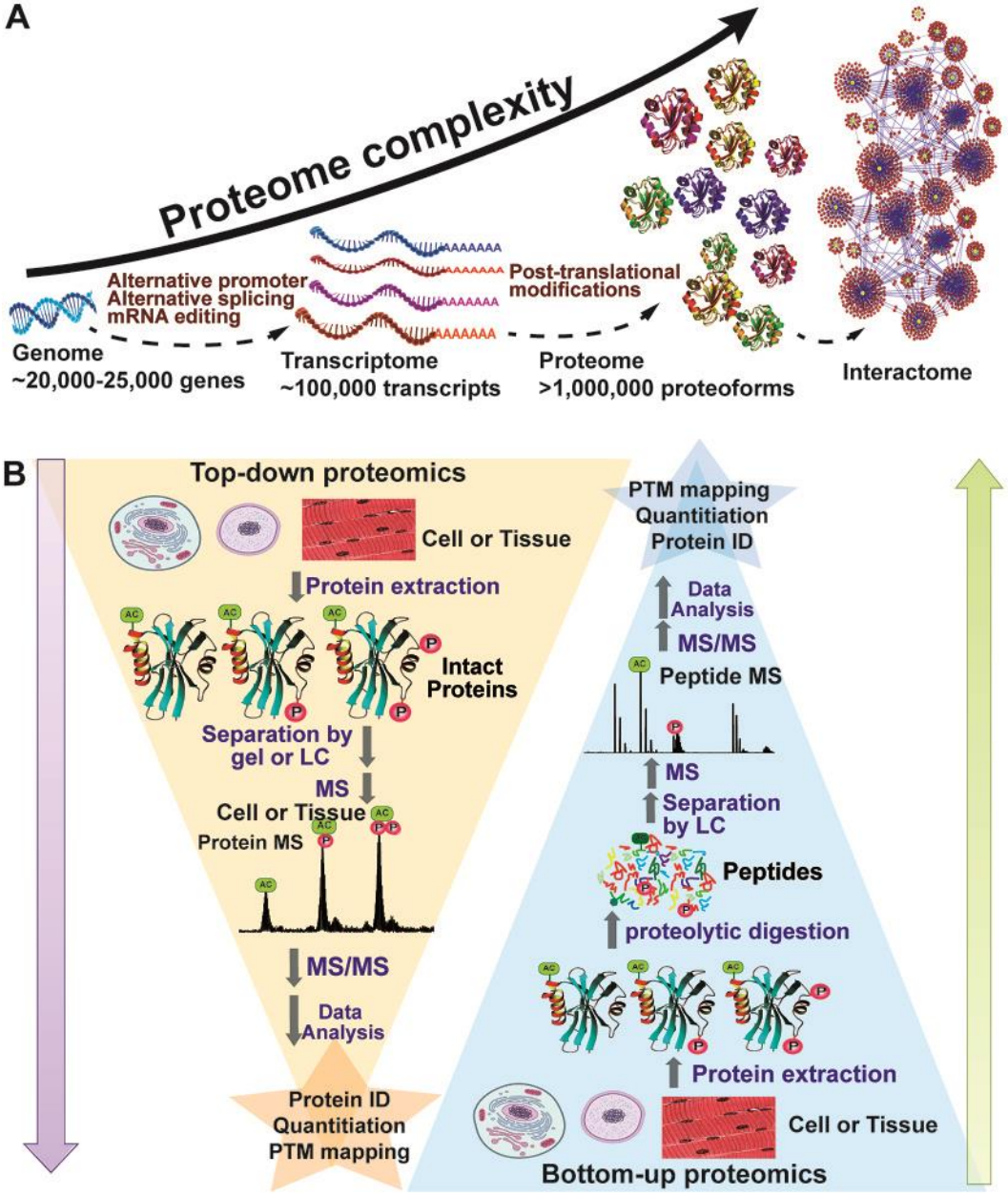


Figure 1. A. Exponential increase of proteome complexity owing to alternative splicing of mRNAs and post-translational modifications of proteins. **B.** Schematic depicting the basic workflow of top-down proteomics compared with that of bottom-up proteomics (Cai, Tucholski, Gregorich, & Ge, 2016).

1.2 Human saliva

Saliva is a clear slightly acid acqueous liquid (pH 6-7), hypotonic fluid if compared to plasma, composed by significant amounts of organic components, such as amylase, lysozyme, lipase, acid phosphatase, lacto peroxidase, superoxide dismutase, various peptide hormones, glycosaminoglycans, lipids and inorganic ions such as sodium, chloride, potassium, calcium, magnesium, bicarbonate, phosphate (Chiappin, Antonelli, Gatti, & De Palo, 2007). Saliva is produced by several types of salivary glands. Each type of salivary gland secretes saliva with characteristic composition and properties. Approximately 90% of total salivary volume results from the activity of 3 pairs of major salivary glands: parotid, submandibular, and sublingual glands, with the bulk of the remainder from minor salivary glands located at various oral mucosal sites (Greabu et al., 2009). The secretions from these different glands can be affected by different form of *stimuli*, time of day, diet, age, gender, a variety of disease states, and several pharmacological agents. The term "saliva" specifically refers to the salivary gland secretion, while "whole saliva" is used to indicate the complex solution deriving also from the contribution of gingival crevicular fluid, mucosal transudation, bronchial and nasal secretions, serum and blood derivatives from oral wounds, desquamated epithelial cells, non-adherent bacteria and bacteria products, and food residues (Humphrey & Williamson, 2001).

The complex mix of salivary constituents provides an effective set of systems to lubricate and to protect the soft and hard tissues and indeed qualitative and quantitative variation of saliva can effect human health in different ways. Specific functions of saliva are summarised in table 1.

Over the past decades many studies, have been devoted to identify several different salivary components and to characterize various classes of proteins and peptides. Most of them are specific of the oral cavity, while others are not part of the normal salivary constituents, and can reach saliva by several ways: intracellularly (through passive transfer, by diffusion) and extracellularly (ultrafiltration). The major salivary protein families that together constitute ≥ 95% of the salivary proteins content are proline rich proteins (acidic, basic and glycosylated), amylase, mucins, cystatins, histatins and statherin that have undergone a complex series of molecular processes, which ultimately define their structures.

Table 1. Saliva functions and components.

Functions	Components
Lubrication	➢ Mucins ➢ Proline-rich glycoproteins ➢ Water
Antimicrobial activity	➢ Lysozyme ➢ Lactoferrin ➢ Lactoperoxides ➢ Mucins ➢ Cystatins ➢ Histatins ➢ Immunoglobulins ➢ Proline-rich glycoproteins
Maintaining mucosa integrity	➢ Mucins ➢ Electrolytes ➢ Water
Cleansing	➢ Water
Buffer capacity and remineralisation	➢ Bicarbonate ➢ Phosphate ➢ Calcium ➢ Fluoride ➢ Statherin ➢ Proline-rich anionic proteins
Preparing food for swallowing	➢ Water ➢ Mucins
Digestion	➢ Amylase ➢ Lipase ➢ Ribonucleases ➢ Proteases ➢ Water ➢ Mucins
Taste	➢ Water ➢ Gustin
Phonation	➢ Water ➢ Mucins

1.2.1 Proline-rich proteins (PRPs)

Proline rich proteins represent the major fraction of salivary proteins, and they can be divided into acidic, basic and glycosylated proteins (Bennick, 1982). The genes encoding for the acidic proline-rich proteins (aPRP) are *PRH1* and *PRH2*, while basic proline-rich proteins (bPRP) are encoded by *PRB1* to *PRB4* and *PBII(SMR3B)* localized on chromosome 12p13 (Isemura, 2000; Maeda, 1985; Maeda et al., 1985).

Acidic PRP

Acidic PRPs are secreted by parotid (70%), submandibular and sublingual glands (30%). They are encoded by *PHR1* (expressing PIF-s, Db-s and Pa) and by *PRH2* that codifies for PRP1 and PRP2 proteins. Glutamic and aspartic acid residues located in the first 30 amino acids of the N-terminus confer the acidic character at these proteins. In adults the five acidic PRP isoforms are mainly phosphorylated at Ser_8 and Ser_{22} and display a pyro-glutamic residue at the N-terminus, although minor amounts of mono-phosphorylated, non-phosphorylated and three-phosphorylated (also on Ser_{17}) proteoforms have been detected in whole saliva (Inzitari et al., 2005). Interestingly, in new-borns the percentages of non-phosphorylated and mono-phosphorylated acidic PRPs are much higher than in adults, suggesting that phosphokinase expression or activity is not yet fully developed at a young age (Inzitari et al., 2007).

PRP1, PRP2, and PIF-s proteins (all composed of 150-residues acidic) are partially cleaved post-translationally after Arg_{106}, generating PRP3, PRP4, and PIF-f, respectively, and a C-terminal 44 residue fragment (PC peptide) that is identical in all proteins (Bennick, 1982). Db-s (171 residues in length) differs from PRP1, PRP2 and PIF-s for the insertion of a 21 amino acid residues repeat after position Gly_{83} (Helmerost & Oppenheim, 2007). It is cleaved, like PRP1, PRP2, and PIF-s, after Arg_{106}, but shifting cleavage site in Arg_{127}, generate Db-f and the same PC peptide. In the acidic Pa protein, the Arg_{103} residue is substituted by a Cys_{103} residue, abolishing the protease recognition site and it is usually present in human saliva as a dimer. An enzyme capable of acidic PRPs processing has been purified from human sublingual glands, and has been tentatively classified as a metal- and thiol- dependent protease (Cai & Bennick, 2004).

Acidic PRPs (entire and cleavage) are effective inhibitors of calcium phosphate crystal growth, but not of primary calcium phosphate precipitation at physiological concentrations. Their calcium-binding affinity and inhibitory activities reside in the highly charged N-terminal 30 amino acids containing the two phosphoserine residues

(Bennick, Cannon, & Madapallimattam, 1981; Hay et al., 1987). Data have provided ample evidence for the importance of protein phosphorylation and the simultaneous presence of both phosphoserines for the biological functions of acidic PRPs pertaining to mineral homeostasis (Hay et al., 1987). Phosphopeptides derived from acidic PRPs are more effective in the inhibition of calcium phosphate precipitation than the intact PRPs. It seems that the structural features and the concentrations of phosphopeptides that are actually present in whole saliva are adequate to fulfil, in part, the functions carry out by both petides and entire proteins PRP (Madapallimattam & Bennick, 1990; Minaguchi, Madapallimattam, & Bennick, 1988).

Basic proline-rich proteins

Basic PRPs constitute more than 30% of the proteins secreted by parotid glands, and are encoded by four genes *PRB1* to *PRB4* and *PBII (SMR3B).* At least four alleles named small (*S*), medium (*M*), large (*L*) and very large (*VL*) are present at *PRB1* and *PRB3* loci, and three *S, M, L* at *PRB2* and *PRB4* loci in the western population (Azen, Minaguchi, Latreille, & Kim, 1990; Lyons, Stein, & Smithies, 1988). These alleles encode for pre-proproteins which, after peptide-signal removal, undergo extensive and complete proteolytic cleavages before ghiandolar secretion, thus only fragments of the proproteins can be detected in saliva. Proteins and peptides deriving from *PRB1* gene are: II-2 peptide (from *S, M, L* alleles), P-E peptides and IB-6 protein (from *S* allele), Ps-1 protein (from *M* allele) and Ps-2 protein (from *L* allele). From *PRB2*, IB-1, P-J, P-H, P-F peptides and IB-8a protein (from *L* allele) have been characterized while *PRB3* and *PRB4* genes give rise to glycosylated proteins and *PRP4* gene also to P-D peptide (from *S, M, L* alleles). Moreover, P-J, P-F and IB-8a can be further cleaved during granule maturation (Azen et al., 1996; Azen, Latreille, & Niece, 1993; Cabras et al., 2009, 2012a; Castagnola et al., 2012; Chan & Bennick, 2001; Lyons et al., 1988; Manconi et al., 2016a; Messana et al., 2015, 2004, 2008a; Stubbs et al., 1998). Given the number of protein sequences obtained from cDNA or genomic DNA large-scale studies, several other potential bPRP species should be detected in human saliva (Manconi et al., 2016a). It should be outlined that the deep knowledge on the multiple bPRP species detectable in saliva, including their natural variants, has been possible thanks to the application of top-down proteomics and peptidomics platforms, for their ability to investigate complex protein mixtures in their naturally occurring forms (Messana et al., 2008b). Some protein masses still pending for a definitive characterization were tentatively attributed to bPRPs family on the basis of their chromatographic properties and the absence of absorption at 270–280 nm

(Castagnola et al., 2012). MS-based glycoproteomic approach has permitted the identification for the N- and O-linked profiling of glycosylation occupancy at site-specific level of PRP3M (Manconi et al., 2016b) and IB-8a glycoproteins.

IB-8a protein, due to a single nucleotide polymorphism Ser_{100}→Pro, exists in two species (Azen et al., 1996). The IB-8a proteoform carrying a Pro_{100} is named IB-8a Con1− and it is not glycosylated on Asn_{98}. On the other hand, IB-8a carrying a Ser_{100} is named Con1+ and it is glycosylated on Asn_{98}. The complete characterization of the non-glycosylated IB-8a Con1− and six different glycoforms *N*-glycosylated at Asn_{98} of IB-8a Con1+ have been characterized in adult human saliva by HPLC–ESI–MS (Cabras et al., 2012a).

Characterization of glycoprotein species is a difficult task, due to their high heterogeneity deriving from the combination of multiple glycosylation sites and different oligosaccharide structures (Manconi et al., 2016b) and their functions in the oral cavity are not well-established. Glycosylation of the bPRPs endows these molecules with unique lubricating properties (Hatton, Loomis, Levine, & Tabak, 1985), common to all highly glycosylated proteins in saliva (Levine et al., 1987). Lubrication helps to protect hard and soft oral tissues against abrasive forces during mastication and facilitates speech. gPRP binds to a variety of bacteria, particularly *F. nucleatum*, but also to various streptococci, including *S. mitis* and *S. sanguis* (Bergey et al., 1986; Gillece-Castro et al., 1991; Levine et al., 1987; Nagata et al., 1983; Ruhl, Sandberg, & Cisar, 2004). Moreover, they bind and precipitate tannins, a dietary constituent with side-effects that are potentially toxic if not neutralized by salivary proteins (Hagerman & Butler, 1981; Lu & Bennick, 1998). The biological functions of basic PRPs, and the purpose for the generation of multiple highly homologous peptides, are otherwise poorly understood. It may be of interest that a subset of small basic PRP peptides in parotid secretion differs between caries-susceptible and caries resistant individuals. This suggests that the proteolytic pattern of basic PRPs could be of diagnostic value and potentially provide markers for caries susceptibility (Ayad et al., 2000).

1.2.2 Cystatins

Cystatins are a superfamily of evolutionary related proteins whose main function in saliva is to provide protection of the oral cavity by inhibiting cysteine proteases (Dickinson, 2002). The mammalian superfamily includes type 1 cystatins (cystatins A and B) and type 2 cystatins (C, D, E, F, S, SN and SA). The genes encoding for the type 1 cystatins are *CSTA* and *CSTB*, while type 2 cystatins are encoded by a multigene family of eight

to nine members. Human cystatins A (also named stefin A or acid cysteine protease inhibitor or epidermal SH-protease inhibitor) and B (also named stefin B or CPI-B or Liver thiol proteinase inhibitor) are inhibitor of cathepsin L, cathepsin S and cathepsin H. They are both single chain proteins exhibiting 54% sequence identity with 98 amino acid residues, a molecular mass of about 11 kDa, lack of signal peptide, disulfide bonds and phosphorylations and are generally expressed intracellularly (Turk, Stoka, & Turk, 2008).

Cystatin A is expressed in the epidermis (Räsänen, Järvinen, & Rinne, 1978), in oral mucosa (Järvinen, Pernu, Rinne, Hopsu-Havu, & Altonen, 1983), in lymphocytes, in neutrophils (Davies & Barrett, 1984), in reticulum cells of lymphoid tissue (Rinne et al., 1983), in Hassall's corpuscles and in thymus medullary cells (Söderström et al., 1994).

Cystatin B is a multifunctional protein widely expressed in different cell types and tissues (Hopsu-Havu et al., 1984; Järvinen & Rinne, 1982; Suzuki et al., 2000). Besides its well-known inhibitory role against cysteine proteinases, additional and/or alternative functions are being discovered, such as lysosome-associated function connected to the molecular pathogenesis of progressive myoclonus epilepsy of the Unverricht-Lundborg type (Alakurtti et al., 2005), prevention of apoptosis (Kopitar-Jerala et al., 2005; Yang et al., 2010), reduction of oxidative stress (Lehtinen et al., 2009) and neuroprotective role by a chaperone-like activity binding with Aβ (Skerget et al., 2010). Recently, our group was able to establish that cystatin B is present in saliva mainly as S-modified derivatives, namely Cys_3 S-glutathionylation, S-cysteinylation and S-S 2-mer (Cabras et al., 2012b).

Type 2 cystatins comprise five major S-type (S, S1, S2, SA, SN) and two minor (C and D). They are composed of about 115-120 amino acid residues with a molecular mass ranging from 13.5 to 14.5 kDa and present signal peptide and disulfide bridges (Lupi et al., 2003). Among the family of salivary cystatins, only cystatin S is phosphorylated. In contrast to other salivary phosphoproteins, cystatin S shows an interesting extent of heterogeneous phosphorylation not observed in other salivary phosphoproteins. It is either non-phosphorylated, or phosphorylated at Ser_3 (cystatin S1), or diphosphorylated at Ser_1 and Ser_3 (cystatin S2) (Isemura et al., 1991; Ramasubbu et al., 1991). Removal of the phosphate groups by alkaline phosphatase treatment reduces the affinity of cystatins (called cysteine-containing phosphoproteins at the time) for hydroxyapatite (Shomers, Tabak, Levine, Mandel, & Hay, 1982). Even though S-type cystatins exhibit a very high sequence identity (about 88%), cystatins SN and SA display greater inhibition toward papain-like cysteine proteinases with respect to cystatin S (Dickinson, 2002).

Human cystatin C consists of a single polypeptide chain of 120 amino acid containing four conserved cysteine residues, which can form two disulfide bonds (Turk & Bode, 1991; Turk, Stoka, & Turk, 2008). This protein has a broad distribution being found in most body (Grubb, 2000), and under pathological conditions it may form amyloid deposit in brain arteries of young adults, leading to cerebral hemorrhage (Olafsson & Grubb, 2000).

Human cystatin D consists of a single polypeptide chain of 122 amino acid residues found in two natural forms with Cys/Arg at position 26 due to gene polymorphism (Balbín et al., 1993). Cystatin D exhibits antiproteinase activity against cathepsin S, H and L but, unlike the other cystatins, not for cathepsin B (Alvarez-Fernandez, Liang, Abrahamson, & Su, 2005; Balbín et al., 1993). Cystatin D is a multifunctional protein and shows several activities not related to the antiproteinase activity, such as inhibition of proliferation, migration and invasion of colon carcinoma cells (Alvarez-Díaz et al., 2009), regulation of antigen presenting cells activity (Nashida et al., 2013) and modulation of gene expression related to its previously unpredicted nuclear activity localization (Ferrer-Mayorga et al., 2015). It displays a narrow tissue distribution with respect to other cystatins because it has been originally found only in saliva (Balbín et al., 1993) and only more recently in colon cancer cells (Ferrer-Mayorga et al., 2015). Moreover, cystatin D may undergo internalization into antigen-presenting cells originated from parotid glands which are responsible for secretion of cystatin D in other body fluids than saliva (Nashida et al., 2013).

The broad spectrum of functions displayed by cystatins explains the reason why variation in their level and structure may affect human health in numerous ways (Grubb, 2000), and indeed they are recently being considered as possible biomarkers in various pathologies none only confined to the oral cavity (Kos et al., 2000; Martini et al., 2017).

Histatins

Histatins are low molecular weight peptides, whose name given by the Oppenheim group derives from the high number of histidine residues in their structure (Oppenheim et al., 1988), secreted both by major and minor salivary glands (Ahmad, Piludu, Oppenheim, Helmerhorst, & Hand, 2004). It is widely accepted that all the members of this family arise from two parent peptides, histatin 1 and histatin 3, with a very similar sequence and are encoded by *HIS1* and *HIS2* respectively, located on chromosome 4q13 (Oppenheim,

Salih, Siqueira, Zhang, & Helmerhorst, 2007). Despite the very high sequence similarity, these two peptides follow different PTM pathways.

Histatin 1 is phosphorylated at Ser_2 in the Ser_2-Asp_3-Glu_4 consensus sequence. In histatin 3, Glu_4 is substituted by Ala_4, abolishing the kinase recognition site and preventing the phosphorylation of Ser_2 (Oppenheim et al., 1988). Before secretion, histatin 3 is exposed to an extensive proteolytic cleavage, generating at first histatin 6 (His-3 Fr. 1/25), subsequently histatin 5 (His-3 Fr. 1/24) and then other fragments (Castagnola et al., 2004), but the enzymes responsible for this processing have not yet been identified. The high number of histatin fragment has stimulated a new nomenclature based on the name of the parent peptide (histatin 1 or histatin 3) and the position in the parent sequence of the first and last amino acid residues (Castagnola et al., 2004).

Histatins exhibit multiple antifungal and antibacterial activities that may or may not be affected by proteolysis. The antifungal activity of histatin 5 exceeds that of histatin 3, pointing toward a potential biological advantage for the generation of this fragment (Helmerhorst, Venuleo, Beri, & Oppenheim, 2005; Xu, Levitz, Diamond, & Oppenheim, 1991). The fungicidal domain, consisting of residues 12-25 in histatin 3, is present in most of the longer naturally occurring histatin fragments (Troxler, Offner, Xu, Vanderspek, & Oppenheim, 1990), suggesting that post-translational proteolysis would not necessarily reduce the antimicrobial properties associated with the unprocessed parent protein. Functional comparisons between native histatin 1 and recombinant expressed histatin 1, lacking phosphate, have shown that the phosphate group has no significant impact on the antifungal properties of histatin 1 (Driscoll et al., 1995). This observation is consistent with the fact that the middle region, rather than the N-terminus, of histatins contains the fungicidal domain (Lamkin & Oppenheim, 1993). On the contrary, non-phosphorylated recombinant histatin 1 exhibits reduced affinity for hydroxyapatite, as compared with native histatin 1, thus indicating that the phosphate group specifically, or its negative charge, is a determining factor in governing the functional interaction of histatin 1 with tooth enamel mineral (Driscoll et al., 1995).

1.2.4 Statherin

Statherin is a small tyrosine-rich phospho-peptide (almost ¼ of the statherin sequence) of 43 amino acids secreted by parotid and submandibular glands, and it was the first fully sequenced salivary phosphoprotein (Schlesinger & Hay, 1977). It's codified by the *STATH* gene located on chromosome 4q13.3 (Sabatini, Carlock, Johnson, & Azen, 1987)

and maybe di-phosphorylated on Ser_2 and Ser_3, but also present such as mono- and non-phosphorylated isoforms, and can be observed in low quantities as a cyclo-statherin (Cabras et al., 2006; Messana et al., 2008a). The cyclo-structure Gln_{37} derives from an intra-molecular bridge between Lys_6 and Gln_{37} generated by the action of oral transglutaminase 2 on statherin (Cabras et al., 2006). In adult human saliva are always detectable *N*- and C-terminal truncated statherin proteoforms (Inzitari et al., 2006).
Has been demonstrated that statherin (in particular the cyclized form) play a key role in the oral calcium homeostasis, having high affinity for the hydroxyapatite, in the teeth mineralization and in the formation of the enamel pellicle (Cabras et al., 2006; Schlesinger & Hay, 1977; Schwartz, Hay, & Schluckebier, 1992). Statherin is the most effective inhibitor of calcium phosphate precipitation among other salivary phosphoproteins. The capacity to prevent crystal growth is localized in the N-terminal six amino acid residues containing both phosphoserines and also aspartic and glutamic residues (Moreno et al., 1979; Raj et al., 1992).

P-B peptide is the product of *PROL3* gene, localized on chromosome 4q13.3, very close to the statherin gene, suggesting a functional relationship with this protein. Its structure was included in the bPRPs family, differently from classical bPRPs, P-B peptide is not a fragment of a bigger pro-protein. P-B peptide is secreted both from parotid and Sm/Sl glands (Messana et al. 2008a) and it displays three Tyr residues in the sequence. While the statherin role is known, none specific function for P-B peptide has been proposed to date (Messana, 2008a; Isemura, 2000; Inzitari, 2006).

1.2.5 α-defensins

α-defensins also named human neutrophil peptides, are basic peptides rich in tyrosine and cysteine residues, the latter forming three disulphide bonds between residues 1 and 6, 2 and 4, and 3 and 5, resulting in peptides forming a triple-stranded β-sheet structure with a β-hairpin loop containing cationic charged molecules. *DEFA1*, *DEFA3* and *DEFA4* genes, located in chromosome 8p23.1, encode for α-defensin 1, 3 and 4 respectively, while α-defensin 2 derives from a proteolytic cleavage of the *N*-amino-terminal residue of α-defensin 1 or 3 (Valore & Ganz, 1992).
Defensins are not secreted from salivary gland and their presence in saliva is justified because defensins represent the major components detected in the gingival crevicular fluid and among them α-defensins 1, 2 and 3 are in major concentration, whereas α-

defensin 4 with minor amounts (Pisano et al., 2005). α-defensins 1 to 4 are expressed in neutrophils in which they play a role in the oxygen-independent killing of phagocytosed microorganisms linked to their antimicrobial activity and are involved in the regulation of the cell volume, cytokine production (Chaly et al., 2000; Lehrer & Lu, 2012), chemotaxis and inhibition of natural killer cells (Goebel, 2000). Moreover the α-defensin 4, also called corticostatin, exhibits pro-inflammatory effects through its anti-corticotropin property, which inhibits the production of cortisol (Singh et al., 1988).

1.2.6 β-thymosins

β-thymosins are ubiquitous peptides having interesting intra- and extra-cellular functions, whose name derives from their first characterization from thymus extracts (Hannappel, 2007; Kleint, Goldstein, & White, 1965). Thymosins β4, β4 oxidized (encoded by *TMSB4X* gene clustered on chromosome Xp22.2) and β10 (encoded by *TMSB10* located on chromosome 2p11.2) have been detected in whole saliva even if they mainly derive from gingival crevicular fluid (Badamchian et al. 2007; Inzitari et al. 2009; Castagnola, et al. 2011a). β-thymosins are involved in the prevention of actin filament polymerization, induction of metalloproteinases, chemotaxis, angiogenesis, inhibition of inflammation and bone marrow stem cell proliferation. They have been also associated to cancer and metastasis formation (Huff et al. 2001; Hannappel 2007; Hannappel 2010).

1.2.7 S100 family

S100 family are low molecular weight acidic proteins with two distinct calcium ion binding domains. Their name derived from the observation that the first identified S100 proteins were obtained from the soluble bovine brain fraction upon fractionation with 100% saturated ammonium sulphate (Moore, 1965).

25 proteins belonging to the S100 family have been identified and distinguished in three subfamily named S100A, S100B and S100P (Marenholz et al., 2004) encoded by genes located on chromosome 1q21 (Sedaghat & Notopoulos, 2008). This clustered organization gave rise to the systematic nomenclature of S100 proteins: in particular polypeptides encoded by genes located within the cluster on chromosome 1 were assigned as S100A proteins with numbers A1–A16, reflecting the position of the gene in the cluster (Marenholz, Lovering, & Heizmann 2006).

They have no intrinsic catalytic activity but, after calcium binding, structural modifications allow them to bind and modulate the action of other proteins. These

proteins appear to be rather young, as they are only present in vertebrates (Shang, Cheng, & Zhou, 2008). S100 proteins, are constitutively expressed in neutrophils, myeloid cells, platelets, osteoclasts and chondrocytes but can be induced and overexpressed in several cell types (macrophages, monocytes, keratinocytes, fibroblasts) in acute and chronic inflammatory and oxidative stress conditions (Edgeworth, 1991; Vogl, 1999; Eckert, 2004; Carlsson, 2005; Sedaghat, 2008; Lim, 2009; Goyette, 2011). It has been demonstrated their involvement in a wide range of intracellular and extracellular functions: regulation of calcium homeostasis, cytoskeletal rearrangement, contraction and motility, cell growth and differentiation, membrane organization, arachidonic acid transport, chemotaxis, apoptosis, promotion of wound repair, protection against microbial proliferation, control of ROS formation, inflammation and protein phosphorylation and secretion (Ravasi, 2004; Santamaria-Kisiel, 2006; Lim, 2009; Sedaghat, 2008; Thorey, 2001; Donato, 2003). Their activity can be altered and regulated through formation of homodimers and heterodimers and by numerous PTMs: phosphorylation, methylation, acetylation and oxidation that can change their ability to bind ions (Ca^{2+}, Zn^{2+} and Cu^{2+}) or target proteins (Lim, 2009; Andrassy, 2006; Zimmer, 2003).

S100A7, S100A8, S100A9, S100A11 and S100A12 were already detected in human saliva (Castagnola et al. 2011a).

Salivary S100A7 (psoriasin) was detected in two isoforms of which the variant D_{27} is most abundant. Both S100A7 variants were N-terminal acetylated following the loss of the initial methionine.

Four isoforms of S100A9 (calgranulin B) were firstly detected in human granulocytes (Strupat et al., 2000), and characterized in human saliva (Castagnola et al., 2011a); two isoforms, defined as long-types, were found to be acetylated following loss of the N-terminal methionine residue and differed from each other in phosphorylation of the Thr_{112}. The other two isoforms, defined as short-types, cleaved of five N-terminal amino acid residues (MTCKM) and differed in the phosphorylation of the same residue of the long-types.

S100A11 was found to be acetylated at the N-terminal residue following methionine loss (Castagnola et al., 2011a).

Derivatives of S100A8 and S100A9 with different degree of oxidation are recently characterized by Cabras et al. (Cabras et al., 2015) through top-down proteomic approach on the intact proteins and peptides present in the acidic supernatant of whole saliva as

well as a bottom-up approach on the tryptic digests of salivary enriched fractions. S100A8 oxidation involved Met_1 and Met_{78}, Trp_{54}, and Cys_{42}. Three proteoforms of S100A8 showed the Cys_{42} residue oxidized to sulfonic acid (S100A8-SO_3H). The first showed a further oxidation at Trp_{54} (S100A8-$SO_3H/W_{54}ox$), the other two forms were isobaric derivatives of S100A8-SO_3H. One form was also oxidized at Trp_{54} and Met_{78} (S100A8-$SO_3H/W_{54}ox/M_{78}ox$), the other was dioxidized at Trp_{54} (S100A8-$SO_3H/W_{54}diox$). All these proteoforms have been named hyper-oxidized S100A8 (Cabras et al., 2015). It was also demonstrated the presence *in vivo* of a glutathionylation of Cys_{42} in S100A8 (S100A8-SSG) and Cys_{42} of S100A8 was also found linked to a S100A9 originated a disulphide bond with Cys_3 of S100A9(L) (S100A8/A9-SSdimer) (Cabras et al., 2015).

1.2.8 Antileukoproteinase

Antileukoproteinase, also known as human secretory leukocyte protease inhibitor (SLPI), is an 11.7 kDa cationic protein and a member of the innate immunity-associated proteins. It is a non-glycosylated, acid-stable, cysteine-rich, 107-amino acid, single-chain polypeptide (Thompson & Ohlsson 1986). The human *SLPI* gene is localized on chromosome 20q12-13.2 (Kikuchi et al., 1998) and the protein is produced by neutrophils, macrophages, β-cells of pancreatic islets, epithelial cells investing the renal tubules, acinar cells of parotid and submandibular glands, acinar cells of submucosal glands, and epithelial cells lining mucous membranes of respiratory and alimentary tracts (Abe et al., 1991; Fahey & Wira 2002; Farquhar et al., 2002; Jin et al. 1997).

SLPI was first isolated from secretions of patients with chronic obstructive pulmonary disease and cystic fibrosis and was thereby considered a major anti-elastase inhibitor (Tegner, 1978). Futher, SLPI was isolated from parotid saliva (Thompson & Ohlsson 1986), in a variety secretions such as whole saliva, seminal fluid, cervical mucus, synovial fluid, breast milk, tears, and cerebral spinal fluid, as in secretions from the nose and bronchi (Castagnola et al. 2011; Farquhar et al., 2002). Moreover, it is also found in neurons and astrocytes in the ischemic brain tissue (Wang et al., 2003).

The main function of SLPI is to protect local tissue against the detrimental consequences of inflammation. It protects the tissues by inhibiting the proteases, such as cathepsin G, elastase, and trypsin from neutrophils; chymotrypsin and trypsin from pancreatic acinar cells; and chymase and tryptase from mast cells (Gipson, T. S., 1999; He et al., 2003).

It also has bactericidal and antifungal properties.

1.2.9 pIgR fragments

Polymeric ImmunoGlobulin Receptor (pIgR), a type I transmembrane glycoprotein transports polymeric IgA across mucosal epithelial cells, playing the main role in the adaptive immune response on mucosal surfaces (Asano et al., 2011; Kaetzel, 2005). It is upregulated by pro-inflammatory cytokines, hormones and microbial factors, through a signaling pathway involving toll-like receptors 3 and 4 (Kaetzel, 2005). A proteolytic cleavage occurring in the glycosylated extracellular portion of pIgR generates the secretory component (19-603 residues), which has been detected also in human saliva (Ramachandran et al., 2006). The cleavage occurs by action of unknown proteases, probably released by activated neutrophils (Kaetzel, 2005), and the highly conserved sequence 602-613 (PRLFAEEKAVAD) is believed to be the cleavage signal (Asano et al., 2011). The AVAD peptide originates by a cleavage occurring in this region at the level of K_{609}, and the ASVD peptide derives from AVAD by the trypsin-like cleavage at R_{622}. The cleavage releasing the C terminal glycine from both fragments could be made by several proteases, including cathepsins and matrix metallopeptidases. AVAD and ASVD peptides do not derive from the secretory component, and have a sequence partially overlapped to the transmembrane portion (639-661) of pIgR. Thus, they should originate by degradation of pIgR after its release from disrupted cell membranes. High levels of pIgR have been associated to the invasion and metastasis of the hepatocellular carcinoma (Ai et al., 2011), while a down regulation of pIgR in intestinal mucosa of animal models subsequent to acute liver necrosis has been observed (Fu et al., 2012).

1.3 Human saliva in proteomics research

Human saliva is a complex fluid that can be collected easily and by a non-invasive method.

There are different methods to collect whole saliva, such as:

- saliva can be dripped from the lower lip into a graduate tube with a funnel;
- saliva can be allowed to accumulate in the floor of the mouth and the sabject split every 30 sec;
- saliva can be continuously aspirated from the floor of the mouth into a tube by an aspirator;

- saliva can be collected by a preweighed swab, cotton roll, or gauze sponge placed in the mouth at the orificies of the major glands and is removed for reweighing at the end of the collection period (Navazesh, 1993).

In a comparative study of these methods, it was found that the suction and swab methods introduced some degree of stimulation and variability and thus are not recommended for unstimulated whole saliva collection (Navazesh & Christensen, 1982).

Saliva is already used routinely by clinical laboratories for detection of secretory IgA antibodies, determination of salivary cortisol, hormones and for genetic purposes. However, recent reports suggest that in the near future human saliva will be a relevant diagnostic fluid for clinical diagnosis and prognosis (Lee & Wong, 2009) being used to monitor and diagnose not only diseases confined to the oral cavity but also systemic pathologies. Recent proteomic platforms have analysed the human salivary proteome, characterising about 3000 differentially expressed proteins and peptides, many of them of microbiological origin (Grassl et al., 2016). By integrating top-down and bottom-up approaches, it has been possible to obtain the characterization of the salivary proteome in different physiological states, such as age, diet, circadian variations (Cabras et al., 2014; Castagnola et al., 2012; Castagnola et al., 2011). These studies allowed to obtain a "reference" protein pattern for the proteomic study of saliva in different pathological conditions, such as autoimmune disorders and genetic diseases.

Because a proteomic strategy able to characterise the whole saliva proteome does not exist (Messana et al., 2013) and many studies on the same disease have been carried out with different instruments and experimental plans, it is necessary to implement top-down and bottom-up approaches, in order to take advantage of the two strategies and to minimize their limitations (Cabras et al., 2014).

Several excellent reviews have recently been published outlining the possibility to use saliva as a diagnostic fluid (Wang et al., 2015; Schafer et al., 2014; Cuevas-Córdoba & Santiago-García, 2014) evidencing significant quali/quantitative difference in some class of peptides/proteins.

Salivary proteome in subjects with human oral carcinoma has showed increased levels of transferrin (Jou et al., 2010), many truncated form of cystatin SN (Shintani, Hamakawa, Ueyama, Hatori, & Toyoshima, 2010) and 46 peptides/proteins with significantly different levels when compared to healty subjects (Hu & Wong, 2007).

Moreover, studies performed on whole saliva from subjects affected by head and neck squamous cell carcinoma have identified several potential tumor markers using SDS-

PAGE-MALDI TOF/TOF-MS (Jarai et al., 2012). Xiao and colleagues evidenced potetial salivary biomarkers in lung cancer by 2-D-DIGE combined with mass-spectrometry technique (Xiao et al., 2012).

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

In SAPHO patients, salivary proteome shows an increase levels and frequency of S100A12 protein. The high expression of this pro-inflammatory protein is probably related to the inflammatory response and to the altered neutrophil responses to functional *stimuli* that characterize SAPHO syndrome suggesting a possible application as a salivary biomarker (Sanna, et al., 2015). The analysis of the salivary proteome of autistic spectrum disorders patients demonstrated hypo-phosphorylation of salivary peptides, suggesting potential asynchronies in the phosphorylation of other secretory proteins, which could be relevant in central nervous system development either during embryonic development or in early infancy (Castagnola et al., 2008).

Down's syndrome, is a frequent genetic disorder in humans, have increased risk of health problems associated with this condition. Research conducted on the salivary proteome has highlighted as opposed to controls, in Down syndrome subjects the concentration of the major salivary proteins of gland origin did not increase with age; as a consequence concentration of acidic proline rich proteins and S cystatins were found significantly reduced in older Down syndrome subjects with respect to matched controls; levels of the antimicrobial α-defensins 1 and 2 and histatins 3 and 5 were significantly increased in whole saliva of older Down syndrome subjects with respect to controls; S100A7, S100A8, and S100A12 levels were significantly increased in whole saliva of Down syndrome subjects in comparison with controls. The increased level of S100A7 and S100A12 may be of particular interest as a biomarker of early onset Alzheimer's disease, which is frequently associated with Down syndrome (Cabras et al., 2013).

The salivary proteome of Wilson's disease patients reflected oxidative stress and inflammatory conditions characteristic of the pathology, highlighting differences that could be useful clues of disease exacerbation (Cabras et al., 2015).

Analysis of both whole saliva and parotid saliva by top-down proteomics platform has been applied to study the effects of pilocarpine treatment on salivary proteins and peptides in patients with Sjögren's syndrome (Peluso et al., 2007).

Giusti and colleagues demonstrated that sclerosis affect the salivary proteome and showed that the chaperon GRP78/BiP increased in saliva of rheumatoid arthritis patients (Giusti et al., 2007), suggesting its potential role as rheumatoid arthritis biomarker (Giusti et al., 2010). Several studies evidenced significant modifications of the peptide fraction in patients with Type 1 diabetes, probably due to increased activity of oral proteases (Hirtz et al. 2006; Caseiro et al., 2013).

The greatest challenge of salivary diagnostics is to identify disease diagnostic markers and successfully translate these research efforts from the laboratory into the clinic. These proteomics approaches are a powerful tools that are utilize for various studies involving qualitative and quantitative protein expression, isoform analysis, post-translational modification analysis, and biomarker discovery (Chugh et al., 2010).

1.4 Multiple sclerosis

Multiple Sclerosis (MS) is a chronic disease that attacks the central nervous system; it affects the brain, spinal cord and optic nerves. It constitutes the leading cause of non-traumatic disability in young adults and it is more common in women than in men with a *ratio* of 3:1. Patients usually experience a first neurological episode known as a clinically isolated syndrome (CIS). This event evolves either into a RR course (85%) or a primary progressive (PP) course (15%). RR patients will evolve into a secondary progressive (SP) course after a period that could vary between 10 and 20 years. While demyelination and inflammation are considered as initial and prominent mechanisms in relapsing-remitting (RR) MS, neurodegeneration is more present in progressive phases of MS, and probably constitutes the main cause of permanent disability accumulation (Mahad et al., 2015).

The instrument for measuring and evaluating the progression of the disease and the effect of treatments in patients with MS is the Expanded Disability Status Scale (EDSS) that was developed in 1950s by Dr. Kurtzke (Kurtzke, 1975).

The EDSS is accepted in clinical trials for evaluating the disability level through two factors: walking ability and scores related to eight functional systems that are pyramid,

cerebellar, encephalic trunk, sensitive, sphinctic, visual, cerebral and others that are variable affected by disease (Lavery et al., 2014).

The global distribution of MS (Figure 2) can be generalised as increasing with distance north or south of the equator, but that summary conceals many places with disproportionately high or low frequencies (Kurtzke, 1975; 1993). In Europe, the countries with the greatest spread of the disease are Denmark (227 cases per 100,000 inhabitants), Sweden (189) Hungary (176), United Kingdom (164). In Eastern Europe, France, Spain and Portugal, data on the prevalence of MS are lower than in other countries.

Always in the European overview, Italy is in an intermediate position with 113 cases per 100,000 inhabitants. There are about 68,000 MS patients in Italy, for a total of about 1800 new cases every year (Browne et al., 2013). Sardinia is one of the regions at the highest risk for MS in the adult population and in the pediatric MS. The risk of MS is estimated to be significantly higher than those reported elsewhere, among the highest values worldwide with a prevalence of 33.3 cases per 100,000 when disease onset occurred within the range of 0-18 years (Dell'Avvento et al., 2016).

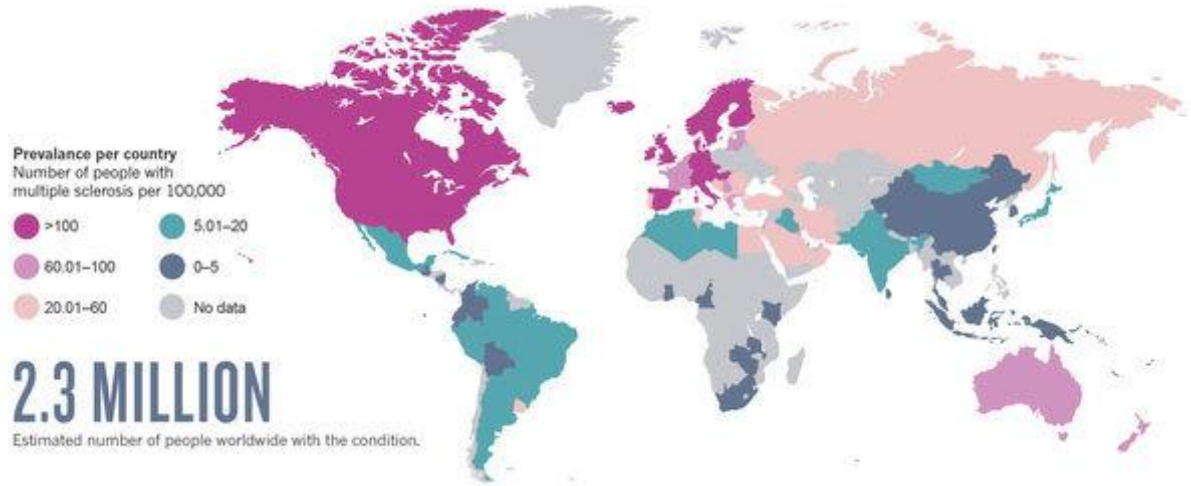


Figure 2. Global distribution of MS. The highest incidence in northern latitudes reflects the likely association with reduced exposure to sunlight and vitamin D deficiency (Schmidt, 2016).

To date no causes have been found in the manifestation of this disease, but there are several factors that affect its development such as: exposure to infectious agents (viruses and bacteria) especially during the first years of life, the environment and ethnicity and

the genetic predisposition. It would be the combination of several factors that trigger the autoimmune mechanism at the base of the onset of symptoms (multifactorial origin).

At the basis of myelin loss there is an alteration in the immune system's response which, under normal conditions, has the task of defending the organism from external agents, mainly viruses and bacteria. In particular, on the basis of a population of 3 million, infection with Epstein-Barr virus as a young adult increases the risk of subsequently developing MS (Ascherio et al., 2005). These data lend support to the so-called hygiene hypothesis whereby individuals not exposed to infections early in life, because of a clean environment, make aberrant responses to infections when encountering these challenges as young adults. Lang and colleagues (Lang et al., 2002) describe a basis for molecular mimicry between Epstein-Barr virus and a self-protein, so that an immune response to the virus inadvertently cross-reacts with myelin and induces demyelination; four DRB1* restricted T-cell receptor peptide contacts are identical for myelin basic protein (which is one of the constituents of myelin) and Epstein-Barr virus. Studies investigating pathological changes suggest that a high proportion of B cells, accumulating in lesions of chronic MS, are infected by Epstein-Barr virus (Serafini et al., 2007).

Moreover, some reaserchers have suggested environmental risk factors in MS aetiology such as low sunlight, vitamin D deficiency, diet, geomagnetism, air pollutants, radioactive rocks, cigarettes and toxins (Hernan et al., 2004).

MS is not an infectious disease and is not transmitted from individual to individual. Similarly, genetic predisposition does not mean that MS is hereditary or transmitted from parents to children with their own chromosomes and MS has a familial recurrence rate of about 20%.

Different studies focused on genetic factors determining familial clustering and individual susceptibility to MS. Individuals with MS who were adopted soon after birth and those having affected members of their adoptive family, have the same risk as does the general population and, therefore, a substantially lower frequency than that observed in the biological relatives of index cases (Ebers et al., 1995). The same is true for step-siblings of index cases (Dyment et al., 2006). The age-adjusted risk for half-siblings is lower than that for full siblings and with no difference in risk for half-siblings reared together or apart (Ebers et al., 1995). The recurrence risk is higher for the children of conjugal than single-affected parents (Ebers et al., 2000).

The association between MS and alleles of the MHC was identified in the early 1970s (Compston et al., 1976; Terasaki et al., 1976). These markers have been refined as DR15

and DQ6 and the corresponding genotypes DRB1*1501, DRB5*0101, DQA1*0102, and DQB2*0602 (Olerup et al., 1991). The association is strongest in northern Europeans but is seen in all populations apart from Sardinians and some other Mediterranean groups in whom MS is associated with DR4 (DRB1*0405–DQA1*0301–DQB1*0302) (Marrosu et al., 1992).

There is not currently a single test available to confirm the diagnosis of MS in a certain and unquestionable sure way. Diagnosis is formulated by three elements: patient symptoms, neurological examination and instrumental (magnetic resonance) and biological analysis (blood and cerebrospinal fluid).

Choosing the proper biological sample is a crucial step in identifying biomarker candidates that might be suitable for clinical use. The blood is the first and more obvious choice, given that it is the most used type of sample for diagnosis and follow-up in clinical practice due to its accessibility and minimally invasive collection procedure. However, several drawbacks must be considered. Albumin and Igs constitute approximately 75% of the total protein weight in blood plasma/serum, and 20 additional proteins make up most of the remaining weight. The other hundreds of low-abundance proteins account for only approximately 1% of the protein weight in plasma/serum (Jaros et al., 2013). The depletion of the most abundant proteins is the standard solution employed in proteomic studies, but we should always consider the possibility that the depletion of these proteins might also deplete other proteins that may be biomarker candidates (Koutroukides et al., 2011). Some authors consider the blood a viable sample to detect CNS alterations when a disruption of the blood–brain barrier (BBB) occurs (Dagley et al., 2013), but such disruptions do not always occur in MS patients. Furthermore, there are not many CNS proteins that are detectable in the blood. However, as an inflammatory disease, the basis of MS depends on cells and humoral factors that are produced in peripheral immune organs and released into the blood. Nevertheless, all studies concerning the identification of potential biomarkers based on immunological features have failed to identify specific markers. As a multifactorial disorder, MS likely involves a set of differentially expressed proteins and genes rather than one particular marker. These proteins, if analyzed in a multifactorial manner, might be useful in combination for prognostic, diagnostic, or patient stratification goals. The CSF is a clear fluid present in the subarachnoid space, which surrounds the CNS. CSF is present in the intracerebral space of the ventricular system and the spinal cord and flows in a unidirectional manner (Nilsson et al., 1992). The CSF is an ideal sample type for identifying modifications in the CNS and is therefore

an interesting source of biomarker candidates. Notably, CSF proteomic data must be carefully evaluated because an increased level of a specific protein may not be related directly to increased expression in the brain tissue but could be due to the degeneration of the CNS tissue, which can also be a biomarker. The major drawback of employing CSF as the biological fluid for clinical applications is the invasive nature of its collection. Furthermore, depletion of high-abundance proteins, such as albumin, IgG, transferrin, and transthyretin, should also be performed prior to the proteomic analysis of CSF. Because MS is a brain disorder, the CNS tissue would be the ideal tissue for discovering biomarkers. However, the impossibility of collecting such tissue from living patients prevents using CNS tissue markers. Despite this limitation, CNS tissue is still a rich source of information for better understanding and characterizing the pathobiology of MS. With this aim, CNS tissues from postmortem human samples and EAE models have been investigated by proteomic techniques and cross-compared with the CSF and serum results. Eventually, the CNS results may be extrapolated to the peripheral tissue in the search for biomarker candidates.

1.5 Autoimmune hepatitis

Autoimmune hepatitis (AIH) is a rare chronic liver disease of unknown ethiology characterized by the loss of immunological tolerance to autologous liver tissue. If not diagnosed and properly treated, the disease can progress to cirrhosis, liver failure and death.

The diagnosis of AIH is based on the combination of clinical and laboratory findings, such as the presence of circulating autoantibodies, hypergammaglobulinemia, increased serum transaminases, with typical histological abnormalities.

The prevalence of AIH differs according to ethnicity: it ranges from 16 to 18 cases per 100,000 inhabitants in Europe (http://www.easl.eu/), while it reaches 42.9 cases per 100,000 and 24.5 cases per 100,000 in Alaska natives and New Zealand, respectively. This variability may be attributed to differences in genetics, environmental factors and study population, and these heterogeneous factors make it challenging to understand the global epidemiology of AIH.

Autoimmune hepatitis has a strong female predominance with a ratio of 3:1 and it is often diagnosed in subjects aged between 40 and 50 years.

The disease is subclassified into two major types: type 1 autoimmune hepatitis (AIH-1) that represents about 90% of cases and type 2 autoimmune hepatitis (AIH-2) which accounts for the remaining 10% of cases. While AIH-1 is frequent in adults with an incidence of about 0.1-1.9/100,000 among Caucasian people and Northern Europe, AIH-2 is more common in children and young adults in Southern Europe, United States and Japan.

AIH-1 is characterized by the detection of antinuclear antibodies (ANA) and/or smooth muscle autoantibodies (SMA) and perinuclear anti-neutrophil cytoplasmic antibodies (p-ANCA) (Muratori et al., 2005). On the contrary, AIH-2 is characterized by the presence of specific autoantibodies, namely anti-liver/kidney microsomal antibody (anti-LKM type 1 or rarely anti-LKM type 3) and/or antibodies against liver cytosol type 1 antigen (anti-LC1) (Krawitt, 1998).

The causes of AIH are still unknown, although remarkable progress in the understanding of the disease pathogenesis has been made over the last years. The prevalent hypothesis suggests the development of AIH in genetically predisposed individuals after their exposure to triggering factors like microbes, viruses and xenobiotics. This results in the activation of an immune response by T lymphocytes which is responsible for the hepatic necrotic and inflammatory damage. Lymphocytes comprise around 25% of intrahepatic immune cells and include effector CD4 T cells, cytotoxic T cells, B cells, invariant iNK, and MAIT cells (Racanelli & Rehermann, 2006). In autoimmune hepatitis, CD4 T cells dominate in the early stage of the disease, particularly Th1 cells (Löhr et al., 1996). Subsequently, T helper cells recruit cytotoxic CD8 T cells leading to higher CD8 T cell frequencies and possibly more aggressive cytotoxic T cell mediated destruction of hepatocytes as the disease progresses (Taubert et al., 2014)

The strongest genetic association has been observed with the genes of the major histocompatibility complex (MCH), in particular, those located within the short arm of chromosome 6, encoding the HLA class II DRB1 alleles (Czaja, 2002). Several studies indicate an association with polymorphisms in genes located outside of the major histocompatibility complex (MHC), like the cytotoxic T lymphocyte antigen-4 (Agarwal et al., 2000), the gene promoter of tumor necrosis factor-alpha (TNF-α) (Cookson et al., 1999) and Fas (Agarwal et al., 2007). Recently, the activating KIR (Killer cell immunoglobulin-like receptors) gene KIR2DS1 has been found highly expressed in patients with type 1 autoimmune hepatitis, suggesting its potential involvement in the pathogenesis of the disease (Littera et al., 2016).

The importance of intestinal microbiota in disease prevention and its role in immune tolerance has been recently highlighted. In particular a study of 24 patients with AIH evidenced an increased permeability of the intestinal barrier, bacterial translocation and microbial alterations with respect to healthy controls and these alterations were correlated with the severity of the disease (Lin et al., 2015).

Several studies have been conducted in order to identify potential biomarkers in AIH. Li et al. have identified candidate biomarkers in AIH patient's serum by a targeted iTRAQ (isobaric tags for relative and absolute quantification) identification; in this case biomarker verification was performed by 2-DE analysis on serum proteins from a mouse model of AIH induced by treatment with concanavalin A (ConA). Candidates were further validated in independent cohorts of ConA treated mice and AIH patients by ELISA (enzyme-linked immuno sorbent assay), suggesting that nine proteins were differentially expressed in AIH mice treated with con-A. Two of these, the third component of complement (C3) and alpha-2-macroglobulin (A2M) were also up-regulated in AIH patient's sera by a targeted iTRAQ identification (Li et al., 2013).

Tahiri et al. were able to demonstrate that liver arginase, CK 8/18, HSP 60, HSP 70, HSP 90, and VCP represent potential candidate targets for autoantibodies in AIH-1 by a bottom-up approach (Tahiri et al., 2008).

To date, there are no top-down and bottom-up studies that use saliva as predictive fluid in AIH. The possible use of salivary proteins variation at the diagnostic level in this pathology could be very promising, being less invasive than liver biopsy, which is the only certainty for disease recognition today.

1.6 Objectives of the study

It is probably surprising for most people to learn that saliva has been used in diagnostics for more than two thousand years. Ancient doctors of traditional Chinese medicine have concluded that saliva and blood are "brothers" in the body and they come from the same origin. It's believed that changes in saliva are indicative of the wellness of the patient. Saliva offers some distinctive advantages: smaller sample aliquots, the possibility of a dynamic study, greater sensitivity, non-invasive, stress free and easy collection procedure, a good cooperation with patients, the possibility to collection somewhere and anywhere, no special equipment and not a trained technician are needed for collection, correlation with levels in blood, potentially valuable for children and older adults, more accurate than blood for detection of many oral and systemic diseases, may provide a cost-

effective approach for the screening of large populations, could eliminate the potential risk of contracting infectious disease for both a technician and the patient. Advances in the use of saliva as a diagnostic fluid have been affected by current technological developments: enzyme-linked fluorescence technique, Western blot assays, polymerase chain reaction (PCR).

Comparing what has been reported in the previous studies, the main objective of this study was to investigate whether the autoimmune response observed in MS and in AIH could be associated with qualitative and quantitative variations of salivary proteins and peptides in patients compared to control subjects for suggestions of a potential biomarker of these conditions.

2. Materials and methods

2.1 Materials

Chemicals and reagents, all of LC–MS grade, were purchased from Merck (Darmstadt, Germany), Waters Corporation (MA, USA), Thermo Fischer Scientific (IL, USA), Bio-Rad (Hercules, CA, USA), GE Healthcare (LC, United Kingdom), Santa Cruz Biotechnology (TX, USA) and Sigma Aldrich (St. Louis, MO, USA).

2.2 Samples

2.2.1 Study subjects

The study protocol and written consent form were approved by the Ethical Committee of the University Hospital of Cagliari, and has therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. All rules were respected and written consent forms were obtained by the donors.

2.2.2 Sample collection

Unstimulated whole saliva (from 0.2 to 1 mL) was collected with a soft plastic aspirator at the basis of the tongue between 8 a.m. and 13 p.m., when salivary secretion is at a maximum (Dawes, 1972). Samples were collected at least 30 min after any food or beverage had been consumed and teeth had been cleaned. After collection, salivary samples were kept in an ice bath and immediately mixed 1:1 v/v *ratio* with a 0.2% solution of 2,2,2-trifluoroacetic acid (TFA) containing 50 μM of leucine enkephalin as internal standard (Sztáray, Memboeuf, Drahos, & Vékey, 2011). The solution was centrifuged at 13400 rpm for 10 min at 4°C to separate the precipitate from the acidic soluble fraction that was immediately analyzed by RP-HPLC-ESI-MS (33 μL, corresponding to 16.5 μL of saliva) or stored at −80 °C until analysis. Also precipitates were stored at -80°C.

2.3 Experimental methods

2.3.1 RP-HPLC-low resolution ESI-MS analysis

Peptides and proteins search were made by low-resolution reversed phase (RP)-HPLC-ESI-MS analysis of the acid soluble fraction of whole saliva samples. The measurements were carried out by a Surveyor HPLC system connected to a LCQ Advantage mass spectrometer (Thermo Fisher Scientific, San Jose, CA). The mass spectrometer was equipped with an ESI source. The chromatographic column was a Vydac (Hesperia, CA) C8 column with 5 μm particle diameter (150 x 2.1 mm). The following solutions were

used: (eluent A) 0.056% (v/v) aqueous TFA, and (eluent B) 0.05% (v/v) TFA in acetonitrile-water 80/20. The gradient applied for the analysis of saliva was linear from 0 to 55% of B in 40 min, and from 55% to 100% of B in 10 min, at a flow rate of 0.10 ml/min toward the ESI source. During the first 5 min of separation, the eluate was diverted to waste to avoid instrument damage because of the high salt concentration. Mass spectra were collected every 3 ms in the *m/z* range 300-2000 in positive ion mode. The MS spray voltage was 5.0 kV, and the capillary temperature was 260 °C. MS resolution was 6000.

2.3.2 Top-down proteomics experiments

Intact proteins quantification by low resolution RP-HPLC-ESI-MS

Deconvolution of averaged ESI-MS spectra was performed by MagTran 1.0 software (Zhang & Marshall, 1998). Average experimental mass values (Mav) were compared with average theoretical values using PeptideMass program available on the Swiss-Prot data bank (http://us.expasy.org/tools/proteomics). The relative abundance of the salivary proteins was determined by measuring the area of RP-HPLC-ESI-MS eXtracted Ion Current (XIC) peaks, considered when the S/N *ratio* was at least 5. This value is linearly proportional to the peptide concentration and it can be used to monitor relative abundances, under constant analytical conditions (Levin, Schwarz, Wang, Leweke, & Bahn, 2007). In determining the XIC peak area the right choice of *m/z* values for the detection of the protein of interest is relevant to avoid *m/z* of ESI potentially overlapping spectra belonging to other proteins that elute very close in crowded areas of chromatographic elution. The window for all these values was in a range of ± 0.5 *m/z*. In order to minimize errors associated with sample dilution, the XIC peak value of each protein/peptide analyzed was correct with respect to the XIC peak value of the leucine encephalin internal standard. The estimated percentage error of the XIC procedure was <8%. XIC peaks were considered when the signal to noise *ratio* was at least 5. Tables 2 reports proteins and peptides investigated in the present study, the Swiss-prot codes, the elution times, the experimental and theoretical average (low-resolution) mass values and the multiply-charged ions utilized to selectively XIC peaks used to quantify proteins/peptides and their derivatives.

Table 2. Proteins and peptides investigated, Swiss-Prot code, elution time (RT), experimental (exp.) and theoretical (th.) average mass values (Mav) and multiply charge ions used for XIC quantification are reported.

Proteins/Peptides	Swiss-Prot Code	RT (min.)	Mav (exp.)	Mav (th.)	Layout *m/z* (Charge)
aPRPs					
PRP-1 0P	P02810	23.2	15355 ± 2	15354–15355	$1280.54_{(+12)}$ $1182.11_{(+13)}$ $1024.63_{(+15)}$ $960.65_{(+16)}$ $904.20_{(+17)}$
PRP-1 1P	P02810	22.9	15435 ± 2	15434–15435	$1287.20_{(+12)}$ $1188.26_{(+13)}$ $1029.96_{(+15)}$ $965.65_{(+16)}$ $908.91_{(+17)}$
PRP-1 2P	P02810	22.2	15515 ± 2	15514–15515	$1293.87_{(+12)}$ $1194.42_{(+13)}$ $1035.29_{(+15)}$ $970.65_{(+16)}$ $913.61_{(+17)}$
PRP-1 3P	P02810	21.6	15595 ± 2	15594–15595	$1418.67_{(+11)}$ $1300.53_{(+12)}$ $1200.57_{(+13)}$ $1040.63_{(+15)}$ $975.65_{(+16)}$
PRP-3 0P	P02810	23.8	11002 ± 1	11001–11002	$1376.21_{(+8)}$ $1101.17_{(+10)}$ $917.81_{(+12)}$ $786.84_{(+14)}$
PRP-3 1P	P02810	23.4	11082 ± 1	11081–11082	$1584.09_{(+7)}$ $1386.20_{(+8)}$ $1008.42_{(+11)}$ $924.47_{(+12)}$ $853.44_{(+13)}$
PRP-3 2P	P02810	22.8	11162 ± 1	11161–11162	$1595.51_{(+7)}$ $1396.20_{(+8)}$ $1015.69_{(+11)}$ $931.14_{(+12)}$ $859.59_{(+13)}$
PRP-3 2P Des R_{106}	P02810	22.8	11004 ± 1	11005–11006	$1573.20_{(+7)}$ $1223.83_{(+9)}$ $1001.49_{(+11)}$ $847.57_{(+13)}$
PRP-3 3P	P02810	21.6	15595 ± 2	15594–15595	$1418.67_{(+11)}$ $1300.53_{(+12)}$ $1200.57_{(+13)}$ $1040.63_{(+15)}$ $975.65_{(+16)}$
P-C	P02810	15.0	4371 ± 1	4370.8	$1457.93_{(+3)}$ $1093.70_{(+4)}$
P-C Des Q_{44}	P02810	14.9	4242.6 ± 0.5	4242.6	$1415.22_{(+3)}$ $1061.67_{(+4)}$
P-C Fr. 1-14	P02810	8.7	1471.7 ± 0.1	1471.6	$1471.71_{(+1)}$ $736.36_{(+2)}$
P-C Fr. 1-25	P02810	11.8	2521.8 ± 0.2	2521.8	$1261.90_{(+2)}$ $841.60_{(+3)}$
P-C Fr. 5-25	P02810	11.3	2083.3 ± 0.2	2083.3	$1042.65_{(+2)}$ $695.44_{(+3)}$
P-C Fr. 15-44	P02810	12.6	2917.2 ± 0.2	2917.2	$1459.63_{(+2)}$ $973.42_{(+3)}$
P-C Fr. 26-35	P02810	8.2	990.5 ± 0.1	990.1	$990.51_{(+1)}$ $495.76_{(+2)}$
P-C Fr. 26-44	P02810	9.5	1867.0 ± 0.1	1867.0	$933.96_{(+2)}$ $622.98_{(+3)}$
P-C Fr. 36-44	P02810	6.9	895.4 ± 0.1	894.9	$895.43_{(+1)}$ $448.22_{(+2)}$
Pa 2-mer 4P	P02810	23.6	30921 ± 3	30920.5	$1628.40_{(+19)}$ $1547.03_{(+20)}$ $1473.41_{(+21)}$ $1406.49_{(+22)}$ $1345.38_{(+23)}$ $1237.83_{(+25)}$ $1146.21_{(+27)}$

Proteins/Peptides	Swiss-Prot Code	RT (min.)	Mav (exp.)	Mav (th.)	Layout *m/z* (Charge)
Db-s 3P	P02810	22.7	17713 ± 2	17712.6	$1772.27_{(+10)}$ $1611.24_{(+11)}$ $1477.06_{(+12)}$ $1363.51_{(+13)}$ $1266.19_{(14)}$ $1042.92_{(+17)}$
Db-s 2P	P02810	22.9	17633 ± 2	17632.6	$1603.97_{(+11)}$ $1470.39_{(+12)}$ $1357.36_{(+13)}$ $1260.48_{(+14)}$ $1176.51_{(+15)}$ $1038.22_{(+17)}$
Db-s 1P	P02810	23.4	17553 ± 2	17552.6	$1756.27_{(+10)}$ $1463.73_{(+12)}$ $1351.21_{(13)}$ $1254.77_{(+14)}$ $1171.18_{(+15)}$ $1098.05_{(+16)}$
Db-s 0P	P02810	23.8	17473 ± 2	17472.6	$1748.27_{(+10)}$ $1457.06_{(+12)}$ $1345.06_{(+13)}$ $1249.05_{(+14)}$ $1165.85_{(+15)}$ $1093.05_{(+16)}$
Db-f 3P	P02810	23.0	13360 ± 2	13359.8	$1670.99_{(+8)}$ $1336.99_{(+10)}$ $1215.54_{(+11)}$ $1114.33_{(+12)}$ $1028.69_{(+13)}$ $955.28_{(+14)}$
Db-f 2P	P02810	23.3	13280 ± 2	13279.8	$1660.99_{(+8)}$ $1328.99_{(+10)}$ $1208.27_{(+11)}$ $1107.66_{(+12)}$ $1022.53_{(+13)}$ $949.57_{(+14)}$
Db-f 1P	P02810	23.9	13200 ± 2	13199.9	$1886.70_{(+7)}$ $1650.99_{(+8)}$ $1467.66_{(+9)}$ $1320.99_{(+10)}$ $943.85_{(+14)}$ $881.00_{(+15)}$
Db-f 0P	P02810	24.1	13120 ± 2	13119.9	$1875.28_{(+7)}$ $1640.99_{(+8)}$ $1458.77_{(+9)}$ $1313.00_{(+10)}$ $938.14_{(+14)}$ $875.67_{(+15)}$
bPRPs					
P-J	P02811/2	14.5	5943.6 ± 0.5	5943.5	$1486.90_{(+4)}$ $1189.72_{(+5)}$ $991.60_{(+6)}$ $850.09_{(+7)}$ $743.95_{(+8)}$
IB-1	P02812	19.4	9593.3 ± 2	9593.4	$1919.68_{(+5)}$ $1599.90_{(+6)}$ $1371.49_{(+7)}$ $1066.94_{(+9)}$ $960.35_{(+10)}$ $873.13_{(+11)}$
P-F	P02812	14.7	5842.5 ± 1	5842.5	$1461.63_{(+4)}$ $1169.51_{(+5)}$ $974.76_{(+6)}$ $835.65_{(+7)}$ $731.32_{(+8)}$
P-H	P02811/2	15.2	5590.1 ± 1	5590.1	$1398.53_{(+4)}$ $1119.03_{(+5)}$ $932.69_{(+6)}$ $799.59_{(+7)}$ $699.77_{(+8)}$
Type 1 cystatins					
Cystatin A	P01040	31.8	11006 ± 2	11006.5	$1835.42_{(+6)}$ $1573.36_{(+7)}$ $1376.82_{(+8)}$ $1223.95_{(+9)}$ $1101.66_{(+10)}$ $1001.60_{(+11)}$

Proteins/Peptides	Swiss-Prot Code	RT (min.)	Mav (exp.)	Mav (th.)	Layout *m/z* (Charge)
Cystatin A Acetyl	P01040	33.0	11049 ± 2	11048.5	$1842.43_{(+6)}$ $1579.37_{(+7)}$ $1382.07_{(+8)}$ $1228.62_{(+9)}$ $1105.86_{(+10)}$ $1005.42_{(+11)}$
Cystatin A (T_{96}→M)	P01040	29.5	11036 ± 2	11036.7	$1840.45_{(+6)}$ $1577.68_{(+7)}$ $1380.59_{(+8)}$ $1227.31_{(+9)}$ $1104.68_{(+10)}$
Cystatin B Acetyl	P04080	33.0	11182 ± 2	11181.6	$1864.61_{(+6)}$ $1598.38_{(+7)}$ $1398.71_{(+8)}$ $1243.41_{(+9)}$ $1119.17_{(+10)}$ $1017.52_{(+11)}$
Cystatin B Acetyl S-CMC	P04080	30.5	11239 ± 2	11239.8	$1874.30_{(+6)}$ $1606.69_{(+7)}$ $1405.98_{(+8)}$ $1249.87_{(+9)}$ $1124.98_{(+10)}$
Cystatin B SSG	P04080	32.8	11487 ± 2	11486.9	$1915.50_{(+6)}$ $1642.00_{(+7)}$ $1436.87_{(+8)}$ $1277.33_{(+9)}$ $1149.70_{(+10)}$ $1045.27_{(+11)}$
Cystatin B S-cysteinyl	P04080	32.9	11301 ± 2	11300.8	$1884.47_{(+6)}$ $1615.40_{(+7)}$ $1413.60_{(+8)}$ $1256.65_{(+9)}$ $1131.08_{(+10)}$ $1028.35_{(+11)}$
Cystatin B S-S dimer	P01034	34.3	22361 ± 3	22361.2	$1864.44_{(+12)}$ $1721.10_{(+13)}$ $1598.24_{(+14)}$ $1491.76_{(+15)}$ $1398.59_{(+16)}$ $1316.37_{(+17)}$ $1243.30_{(+18)}$ $1177.92_{(+19)}$ $1119.07_{(+20)}$ $1065.83_{(+21)}$ $1017.43_{(+22)}$ $973.24_{(+23)}$
Type 2 cystatins					
Cystatin S	P01036	35.3	14185 ± 2	14184.7	$1577.12_{(+9)}$ $1419.51_{(+10)}$ $1290.55_{(+11)}$ $1183.09_{(+12)}$ $1092.16_{(+13)}$
Cystatin S1	P01036	35.3	14265 ± 2	14264.7	$1585.97_{(+9)}$ $1427.48_{(+10)}$ $1297.80_{(+11)}$ $1189.73_{(+12)}$ $1098.29_{(+13)}$
Cystatin S1 mono-ox	P01036	35.2	14281 ± 2	14280.7	$1587.75_{(+9)}$ $1429.08_{(+10)}$ $1299.25_{(+11)}$ $1191.07_{(+12)}$ $1099.52_{(+13)}$
Cystatin S1 di-ox	P01036	35.2	14297 ± 2	14296.7	$1589.53_{(+9)}$ $1430.68_{(+10)}$ $1300.71_{(+11)}$ $1192.40_{(+12)}$ $1100.75_{(+13)}$
Cystatin S2	P01036	35.3	14345 ± 2	14344.7	$1594.86_{(+9)}$ $1435.48_{(+10)}$ $1305.07_{(+11)}$ $1196.40_{(+12)}$ $1104.44_{(+13)}$
Cystatin S2 mono-ox	P01036	35.2	14361 ± 2	14360.7	$1596.64_{(+9)}$ $1437.08_{(+10)}$ $1306.52_{(+11)}$ $1197.73_{(+12)}$ $1105.68_{(+13)}$

Proteins/Peptides	Swiss-Prot Code	RT (min.)	Mav (exp.)	Mav (th.)	Layout *m/z* (Charge)
Cystatin S2 di-ox	P01036	35.2	14377 ± 2	14376.7	$1598.42_{(+9)}$ $1438.68_{(+10)}$ $1307.98_{(+11)}$ $1199.06_{(+12)}$ $1106.91_{(+13)}$
Cystatin SN	P01037	34.6	14312 ± 2	14312.0	$1591.23_{(+9)}$ $1432.21_{(+10)}$ $1302.10_{(+11)}$ $1193.68_{(+12)}$ $1101.93_{(+13)}$
Cystatin SN mono-ox	P01037	33.9	14328 ± 2	14328.0	$1593.01_{(+9)}$ $1433.81_{(+10)}$ $1303.56_{(+11)}$ $1195.01_{(+12)}$ $1103.16_{(+13)}$
Cystatin SN di-ox	P01037	33.7	14344 ± 2	14344.0	$1594.79_{(+9)}$ $1435.41_{(+10)}$ $1305.01_{(+11)}$ $1196.34_{(+12)}$ $1104.39_{(+13)}$
Cystatin SN Des_{1-4}	P01037	33.4	13813 ± 2	13813.6	$1535.79_{(+9)}$ $1382.31_{(+10)}$ $1256.73_{(+11)}$ $1152.09_{(+12)}$ $1063.55_{(+13)}$
Cystatin SN Des_{1-7}	P01037	33.2	13440 ± 2	13440.2	$1494.37_{(+9)}$ $1345.03_{(+10)}$ $1222.85_{(+11)}$ $1121.03_{(+12)}$ $1034.87_{(+13)}$
Cystatin SN (P_{11}→L)	P01037	34.7	14328 ± 2	14328.2	$1593.03_{(+9)}$ $1433.83_{(+10)}$ $1303.57_{(+11)}$ $1195.02_{(+12)}$ $1103.18_{(+13)}$
Cystatin SN Des_{1-4} (P_{11}→L)	P01037	34.3	13830 ± 2	13829.6	$1537.67_{(+9)}$ $1384.01_{(+10)}$ $1258.28_{(+11)}$ $1153.51_{(+12)}$ $1064.85_{(+13)}$
Cystatin SN Des_{1-7} (P_{11}→L)	P01037	29.6	13456 ± 2	13456.7	$1496.19_{(+9)}$ $1346.67_{(+10)}$ $1224.34_{(+11)}$ $1122.39_{(+12)}$ $1036.13_{(+13)}$
Cystatin SA	P09228	36.8	14346 ± 2	14346.0	$1595.01_{(+9)}$ $1435.61_{(+10)}$ $1305.19_{(+11)}$ $1196.51_{(+12)}$ $1104.55_{(+13)}$
Cystatin SA mono-ox	P09228	36.6	14362 ± 2	14362.0	$1596.79_{(+9)}$ $1437.21_{(+10)}$ $1306.65_{(+11)}$ $1197.84_{(+12)}$ $1105.78_{(+13)}$
Cystatin SA Des_{1-7}	P09228	32.5	13474 ± 2	13474.3	$1498.14_{(+9)}$ $1348.43_{(+10)}$ $1225.94_{(+11)}$ $1123.86_{(+12)}$ $1037.48_{(+13)}$
Cystatin C	P01034	35.1	13343 ± 2	13343.1	$1483.57_{(+9)}$ $1335.32_{(+10)}$ $1214.02_{(+11)}$ $1112.93_{(+12)}$ $1027.40_{(+13)}$
Cystatin C mono-ox	P01034	38.4	13360 ± 2	13359.1	$1670.89_{(+8)}$ $1485.34_{(+9)}$ $1336.91_{(+10)}$ $1215.46_{(+11)}$ $1114.26_{(+12)}$ $1028.62_{(+13)}$ $955.22_{(+14)}$

Proteins/Peptides	Swiss-Prot Code	RT (min.)	Mav (exp.)	Mav (th.)	Layout *m/z* (Charge)
Cystatin D C_{26}→R	P28325	31.0	13908 ± 2	13907.7	$1739.46_{(+8)}$ $1546.30_{(+9)}$ $1391.77_{(+10)}$ $1265.34_{(+11)}$ $1159.98_{(+12)}$ $1070.82_{(+13)}$
Cystatin D C_{26}→R Des_{1-4}	P28325	31.0	13605 ± 2	13605.4	$1701.68_{(+8)}$ $1512.71_{(+9)}$ $1361.54_{(+10)}$ $1237.85_{(+11)}$ $1134.78_{(+12)}$ $1047.57_{(+13)}$
pGlu-cystatin DC_{26}→R Des_{1-5}	P28325	31.0	13517 ± 2	13517.3	$1690.66_{(+8)}$ $1502.92_{(+9)}$ $1352.73_{(+10)}$ $1229.85_{(+11)}$ $1127.44_{(+12)}$ $1040.79_{(+13)}$
Cystatin D C_{26}→R Des_{1-8}	P28325	31.0	13163 ± 2	13163.0	$1646.38_{(+8)}$ $1463.56_{(+9)}$ $1317.30_{(+10)}$ $1197.64_{(+11)}$ $1097.92_{(+12)}$ $1013.54_{(+13)}$
Histatins					
Histatin 1 0P	P15515	22.0	4848.2 ± 0.5	4848.2	$1617.06_{(+3)}$ $1213.05_{(+4)}$
Histatin 1	P15515	21.9	4928.2 ± 0.5	4928.2	$1643.72_{(+3)}$ $1233.05_{(+4)}$ $986.64_{(+5)}$ $822.37_{(+6)}$ $705.03_{(+7)}$ $617.03_{(+8)}$
Histatin 3	P15516	17.7	4062.4 ± 0.5	4062.4	$1355.14_{(+3)}$ $1016.61_{(+4)}$
Histatin 5	P15516	14.6	3036.3 ± 0.3	3036.3	$1013.12_{(+3)}$ $760.09_{(+4)}$
Histatin 6	P15516	14.3	3192.5 ± 0.3	3192.5	$1065.18_{(+3)}$ $799.14_{(+4)}$
Histatin 3 Fr. 2-6	P15516	5.5	597.3 ±0.05	597.7	$598.34_{(+1)}$
Histatin 3 Fr. 1-11	P15516	7.8	1334.7 ± 0.1	1335.5	$1335.67_{(+1)}$ $668.34_{(+2)}$ $445.89_{(+3)}$
Histatin 3 Fr. 1-12	P15516	8.5	1490.8 ± 0.1	1491.6	$1491.77_{(+1)}$ $746.39_{(+2)}$ $497.93_{(+3)}$
Histatin 3 Fr. 1-13	P15516	8.4	1618.9 ± 0.1	1619.8	$1619.86_{(+1)}$ $810.43_{(+2)}$ $540.63_{(+3)}$
Histatin 3 Fr. 28-32	P15516	12.5	686.3 ±0.05	686.7	$687.30_{(+1)}$
Statherins and PB					
Statherin 0P	P02808	28.6	5219.8 ± 0.5	5219.8	$1740.93_{(+3)}$ $1305.95_{(+4)}$ $1044.96_{(+5)}$
Statherin 1P	P02808	28.9	5299.7 ± 0.5	5299.7	$1767.59_{(+3)}$ $1325.94_{(+4)}$ $1060.96_{(+5)}$
Statherin	P02808	29.2	5379.7 ± 0.5	5379.7	$1794.25_{(+3)}$ $1345.94_{(+4)}$ $1076.95_{(+5)}$
Statherin Des_{1-9}	P02808	28.5	4127.6 ± 0.5	4127.6	$1376.87_{(+3)}$ $1032.90_{(+4)}$
Statherin Des_{1-10}	P02808	28.0	3971.4 ± 0.5	3971.4	$1986.71_{(+2)}$ $1324.81_{(+3)}$
Statherin Des_{1-13}	P02808	27.5	3645.0 ± 0.5	3645.0	$1823.51_{(+2)}$ $1216.01_{(+3)}$
Statherin SV1	P02808	27.8	5232.5 ± 0.5	5232.5	$1745.19_{(+3)}$ $1309.14_{(+4)}$ $1047.52_{(+5)}$
Statherin Des T_{42} F_{43}	P02808	27.9	5131.4 ± 0.5	5131.4	$1711.49_{(+3)}$ $1283.87_{(+4)}$ $1027.30_{(+5)}$

Proteins/Peptides	Swiss-Prot Code	RT (min.)	Mav (exp.)	Mav (th.)	Layout *m/z* (Charge)
Statherin Des D_1	P02808	28.7	5264.6 ± 0.5	5264.6	$1755.88_{(+3)}$ $1317.17_{(+4)}$ $1053.93_{(+5)}$
P-B	P02814	30.0	5792.7 ± 0.5	5792.7	$1931.92_{(+3)}$ $1449.19_{(+4)}$ $1159.55_{(+5)}$
P-B Des_{1-4}	P02814	30.0	5371.3 ± 0.5	5371.3	$1791.43_{(+3)}$ $1343.83_{(+4)}$ $1075.26_{(+5)}$
P-B Des_{1-5}	P02814	30.3	5215.1 ± 0.5	5215.1	$1739.37_{(+3)}$ $1304.78_{(+4)}$ $1044.03_{(+5)}$
P-B Des_{1-7}	P02814	30.1	5060.9 ± 0.5	5060.9	$1687.98_{(+3)}$ $1266.24_{(+4)}$ $1013.19_{(+5)}$
P-B Des_{1-12}	P02814	27.5	4549.3 ± 0.5	4549.3	$1517.46_{(+3)}$ $1138.34_{(+4)}$
α-defensins					
α-defensin 1	P59665	23.5	3442.0 ± 0.4	3442.0	$1722.03_{(+2)}$ $1148.36_{(+3)}$ $861.52_{(+4)}$
α-defensin 2	P59665	23.5	3371.0 ± 0.4	3371.0	$1686.49_{(+2)}$ $1124.66_{(+3)}$ $843.75_{(+4)}$
α-defensin 3	P59666	23.5	3486.1 ± 0.4	3486.1	$1744.03_{(+2)}$ $1163.03_{(+3)}$ $872.52_{(+4)}$
α-defensin 4	P12838	27.2	3707.8 ± 0.5	3709.4	$1855.71_{(+2)}$ $1237.48_{(+3)}$ $928.36_{(+4)}$
Thymosins					
Thymosin β4	P62328	18.5	4963.5 ± 1	4963.5	$1655.51_{(+3)}$ $1241.88_{(+4)}$ $993.71_{(+5)}$
Thymosin β4 sulfox	P62328	18.3	4979.5 ± 1	4979.5	$1660.84_{(+3)}$ $1245.88_{(+4)}$ $996.91_{(+5)}$
Thymosin β10	P63313	20.8	4936.5 ± 1	4936.5	$1646.52_{(+3)}$ $1235.14_{(+4)}$ $988.31_{(+5)}$
S100A					
S100A7 (D_{27})	P31151	37.0	11368 ± 2	11367.8	$1421.98_{(+8)}$ $1264.10_{(+9)}$ $1137.79_{(+10)}$ $1034.44_{(+11)}$
S100A8	P05109	40.4	10834 ± 2	10834.5	$1355.26_{(+8)}$ $1204.79_{(+9)}$ $1084.41_{(+10)}$ $985.92_{(+11)}$
S100A8-SSG	P05109	38.2	11140 ± 2	11139.8	$1393.48_{(+8)}$ $1238.76_{(+9)}$ $1114.98_{(+10)}$ $1013.71_{(+11)}$
S100A8-SNO	P05109	40.7	10863 ± 2	10863.5	$1358.94_{(+8)}$ $1208.06_{(+9)}$ $1087.35_{(+10)}$ $988.59_{(+11)}$
S100A8/A9-SS dimer		41.7	23986 ± 3	23985.0	$1600.00_{(+15)}$ $1500.06_{(+16)}$ $1411.88_{(+17)}$ $1333.50_{(+18)}$ $1263.37_{(+19)}$ $1200.25_{(+20)}$ $1143.14_{(+21)}$ $1091.23_{(+22)}$ $1043.83_{(+23)}$ $1000.38_{(+24)}$ $960.40_{(+25)}$ $923.50_{(+26)}$

Proteins/Peptides	Swiss-Prot Code	RT (min.)	Mav (exp.)	Mav (th.)	Layout *m/z* (Charge)
Hyper-oxidized S100A8	P05109	39.3	10915 ± 2	10914.6	$1365.33_{(+8)}$ $1213.73_{(+9)}$ $1092.46_{(+10)}$ $993.24_{(+11)}$
S100A8-SO3H/W54ox	P05109	40.4	10898 ± 2	10898.6	$1363.33_{(+8)}$ $1211.96_{(+9)}$ $1090.86_{(+10)}$ $991.78_{(+11)}$
S100A8-SO_2H	P05109	39.8	10866 ± 2	10866.5	$1359.31_{(+8)}$ $1208.39_{(+9)}$ $1087.65_{(+10)}$ $988.86_{(+11)}$
S100A9 short	P06702	42.2	12689 ± 2	12689.2	$1410.92_{(+9)}$ $1269.93_{(+10)}$ $1154.57_{(+11)}$ $1058.44_{(+12)}$ $977.10_{(+13)}$
S100A9 short mono-ox	P06702	42.0	12705 ± 2	12705.2	$1412.70_{(+9)}$ $1271.53_{(+10)}$ $1156.03_{(+11)}$ $1059.78_{(+12)}$ $978.33_{(+13)}$
S100A9 short P	P06702	42.2	12769 ± 2	12769.2	$1419.81_{(+9)}$ $1277.93_{(+10)}$ $1161.84_{(+11)}$ $1065.11_{(+12)}$ $983.25_{(+13)}$
S100A9 short P mono-ox	P06702	42.0	12785 ± 2	12785.2	$1421.59_{(+9)}$ $1279.53_{(+10)}$ $1163.30_{(+11)}$ $1066.44_{(+12)}$ $984.48_{(+13)}$
S100A9 long	P06702	41.9	13153 ± 2	13152.8	$1316.29_{(+10)}$ $1196.72_{(+11)}$ $1097.08_{(+12)}$ $1012.76_{(+13)}$ $940.50_{(+14)}$
S100A9 long P	P06702	41.9	13233 ± 2	13232.8	$1324.29_{(+10)}$ $1203.99_{(+11)}$ $1103.74_{(+12)}$ $1018.92_{(+13)}$ $946.21_{(+14)}$
S100A9 long SSG	P06702	41.5	13458 ± 2	13458.1	$1346.82_{(+10)}$ $1224.48_{(+11)}$ $1122.52_{(+12)}$ $1036.25_{(+13)}$ $962.30_{(+14)}$
S100A9 long SSG P	P06702	41.5	13538 ± 2	13538.1	$1354.82_{(+10)}$ $1231.75_{(+11)}$ $1129.18_{(+12)}$ $1042.40_{(+13)}$ $968.02_{(+14)}$
S100A9 long cyst	P06702	41.6	13272 ± 2	13272.0	$1328.21_{(+10)}$ $1207.55_{(+11)}$ $1107.01_{(+12)}$ $1021.93_{(+13)}$ $949.01_{(+14)}$
S100A9 long cyst P	P06702	41.6	13352 ± 2	13352.0	$1336.20_{(+10)}$ $1214.82_{(+11)}$ $1113.67_{(+12)}$ $1028.08_{(+13)}$ $954.72_{(+14)}$
S100A12	P80511	40.0	10444 ± 2	10443.8	$1492.99_{(+7)}$ $1306.49_{(+8)}$ $1161.43_{(+9)}$ $1045.39_{(+10)}$ $950.45_{(+11)}$
Antileukoproteinase					
Antileukoproteinase	P03973	26.2	11710 ± 2	11709.8	$1952.64_{(+6)}$ $1673.84_{(+7)}$ $1464.73_{(+8)}$ $1302.10_{(+9)}$
AVAD and ASVD					

Proteins/Peptides	Swiss-Prot Code	RT (min.)	Mav (exp.)	Mav (th.)	Layout *m/z* (Charge)
AVAD	P01833	25.2	3834.1 ± 0.3	3834.1	$1918.04_{(+2)}$ $1279.03_{(+3)}$ $959.52_{(+4)}$
ASVD	P01833	24.7	2490.7 ± 0.3	2490.7	$1246.35_{(+2)}$ $831.23_{(+3)}$

N.B. 0P: non-phosphorilation; 1P: mono-phosphorilation; 2P: di- phosphorilation; 3P: tri-phosphorilation; Ox: oxidation; SSG: S-glutathionilation; Cyst: S-cysteinilation; Acetyl: acetilation; CMC: S-(carboxymethyl)-cysteine residue; Hyper oxidized: proteoforms from S100A8-SO_3H, named S100A8-SO_3H/W_{54}ox/M_{78}ox e S100A8-SO_3H/W_{54}diox

2.4 Enriched fraction preparation and bottom-up experiments

Enriched fractions were obtained by preparative RP-HPLC (Dionex Ultimate 3000 instrument, Thermo Fisher Scientific, Sunnyvale, CA) of selected saliva samples. The chromatographic column was a reversed phase Vydac (Hesperia, CA) C8 column with 5 µm particle diameter (250 x 10 mm). The eluents used for preparative RP-HPLC were the same utilized for analytical HPLC low-resolution ESI-MS experiments. The step gradient was from 0 to 50% B in 30 minutes, from 50 to 65% B in 20 minutes, and from 65 to 100% B in 1 minute with a flow rate of 2.8 ml/min. Fractions corresponding to peaks eluting between 30 and 37 minutes were collected separately and lyophilized. Each fraction was solubilized in 100 µl of ultrapure H_2O, and 1/3 of the solution was acidified with 0.2% TFA (1:1 v/v *ratio*) to be checked by HPLC low-resolution ESI-MS. The remaining sample was lyophilized for HPLC high-resolution-ESI-MS/MS experiments. For the structural characterization of cystatin D proteoforms, an aliquot of the lyophilized sample was submitted to reduction of disulphide bonds at 100 °C for 5 min followed by an incubation at 50 °C for 15 min in 100 mM ammonium bicarbonate buffer pH 8.0 containing 10 mM dithiotreitol (DTT) in a final volume of 35 µl. The reaction sample was alkylated in the dark at 30 °C for 45 min using 55 mM iodoacetamide (IAM) and subsequently submitted to enzyme digestion. The trypsin digestion was performed by the kit "Trypsin Singles Proteomic Grade" (Sigma-Aldrich) according to the manufacturer's instructions. Digestion was stopped after 8 h by acidification with 0.1% FA, the sample was lyophilized and stored at -80° C until the analysis by HPLC high-resolution ESI-MS/MS.

2.5 RP-HPLC-high-resolution ESI-MS/MS experiments

The experiments were carried out by an Ultimate 3000 RSLC Nano System HPLC apparatus (Thermo Fisher Scientific, Sunnyvale, CA) coupled to a LTQ Orbitrap Elite apparatus (Thermo Fisher Scientific, Sunnyvale, CA). The columns were a Zorbax 300SB-C8 column (3.5 µm particle diameter; 1.0 x 150 mm) for the top-down analysis, and a Zorbax 300SB-C18 column (3.5 µm particle diameter; 1.0 x 150 mm) for the bottom-up. Eluents were: (eluent A) 0.1% (v/v) aqueous formic acid (FA) and (eluent B) 0.1% (v/v) FA in acetonitrile-water 80/20. For both top-down and bottom-up analyses the gradient was: 0-2 min 5% B, 2-40 min from 5% to 55% B (linear), 40-45 min from 70% to 99% B, at a flow rate of 50 µL/min. MS and MS/MS spectra were collected in positive mode with the resolution of 60000. The acquisition range was from 350 to 2000 *m/z* for the top-down and the bottom-up experiments. Tuning parameters: capillary temperature was 300 °C, and the source voltage 4.0 kV, S-Lens RF level 60% in both experiments. In data-dependent acquisition mode the five most abundant ions were selected and fragmented by using collision-induced dissociation (CID) or higher energy collision dissociation (HCD), with 35% normalized collision energy for 30 ms, isolation width of 5 *m/z*, activation q of 0.25. The inject volume was 20 µL. HPLC-ESI-MS and MS/MS data were generated by Xcalibur2.2 SP1.48 (Thermo Fisher Scientific) using default parameters of the Xtract program for the deconvolution. In top-down experiments protein sequences and sites of covalent modifications were validated by manual inspection of the experimental fragmentation spectra against the theoretical ones generated by MS-Product software available at the ProteinProspector website (http://prospector.ucsf.edu/prospector/mshome.htm). In bottom-up experiments, MS/MS data were analyzed by the Proteome Discoverer 1.4 program, based on SEQUEST HT cluster as a search engine (University of Washington, licensed to Thermo Electron Corporation, San Jose, CA) against the Swiss-Prot *Homo Sapiens* proteome (UniProtKB, Swiss-Prot, release 2017_02). The settings were: trypsin enzyme with a maximum of two missed cleavage sites, precursor mass search tolerance was 10 ppm and fragment mass tolerance 0.5 Da. Target FDR was: 0.01 (strict), 0.05 (relaxed). The following modifications were selected: in dynamic mode oxidation of methionine and carbamidomethylation of cysteine in static mode.

2.6 Statistical analysis

The software GraphPad Prism (version 6.0) was used for calculating means and standard deviations of protein XIC peak areas and for statistical analyses. A comparison between patients and controls was performed by using the following statistical tests depending on data distribution (normal or skewed), and variance (homogeneous or unequal): parametric t test (variance homogeneous); t test with Welch correction (normal distribution, variance unequal), and the nonparametric Mann-Whitney test (skewed distribution, variance unequal). Correlation between protein/peptide levels and clinical data was evaluated by using Pearson or Spearman test according to data distribution. Statistical analysis was considered significant when the p value was less than 0.05 (two-tailed).

2.7 Cystatins characterization data analysis SIFT

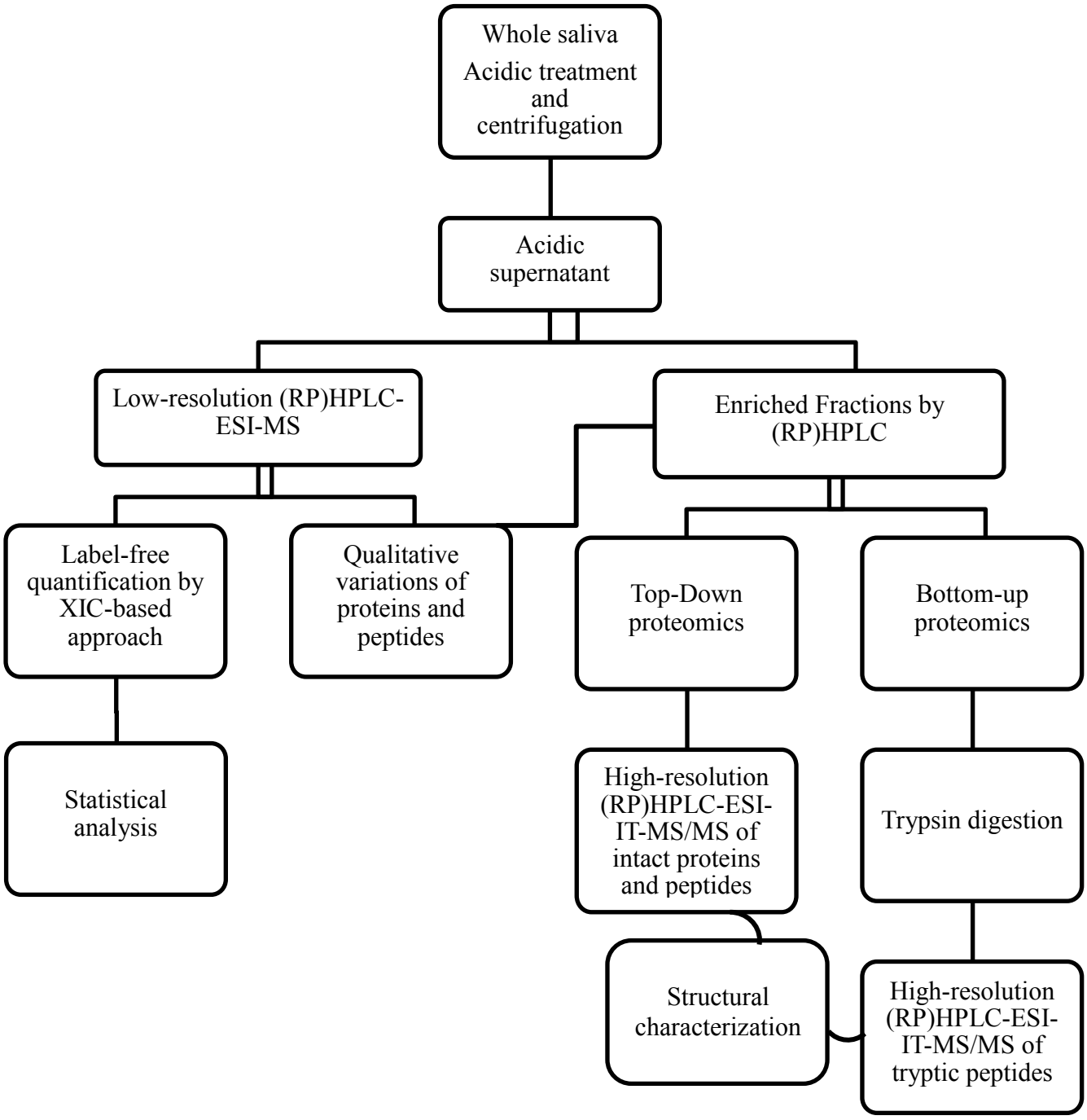
The potential impact of the amino acid substitution on cystatin A function was predicted by *SIFT Human Protein Prediction* software available at the SIFT website (http://sift.jcvi.org/). SIFT (Sorting Intolerant From Tolerant) algorithm is based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences, collected through PSI-BLAST (Ng & Henikoff, 2003). SIFT results with score <0.05 indicate amino acids deleterious on protein function.

2.8 Cystatins characterization data analysis BLAST

Amino acid sequence was compare by BLAST *Basic Local Alignment Search Tool* software available at the BLAST website (https://web.expasy.org/blast/). BLAST is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of proteins or the nucleotides of DNA sequences, with a library or database of sequences, and identify library sequences that resemble the query sequence above a certain threshold.

3. Results

This proteomics study allows evaluating the qualitative and quantitative differences in the soluble fraction of saliva of MS and AIH subjects with respect to a control groups matched by sex and age. Whole saliva, collected and mixed 1:1 v/v *ratio* with a solution of TFA, was centrifuged to separate the precipitate from the acidic soluble fraction and the latter analyzed by both RP-HPLC low and high resolution ESI-MS. Qualitative and quantitative analyses were performed to the acid supernatant according to the following scheme.



3.1 Salivary proteome in Multiple Sclerosis subjects

3.1.1 Patients Population

MS patients were recruited at the Multiple Sclerosis Center, Department of Medical Sciences and Public Health, University of Cagliari, Sardinia, Italy.

MS patients were 49 (39.6 ± 9.9 years old, males n = 24, females n = 25) that, based on the clinical course, were classified into relapsing-remitting (RR; n=38), primary progressive (PP; n=6) and secondary progressive (SP; n=5) according to McDonald criteria (Polman et al., 2011). 32 MS subjects were under therapy at the moment of saliva sampling (immunomodulatory drugs: Rebif, Avonex, Copaxone, Tecfidera, Betaferon; immunosuppressive drugs: Gilenya, Azathioprine, Aubagio; monoclonal antibodies: Lemtrada, Ocrelizumab, Tysabri) and 17 MS subjects were without any pharmacological treatment at the time of inclusion in the study (table 3).

The healthy control group comprised 54 subjects (41.0 ± 10.8 years old, males n = 23, females n = 31) with no history of neurological diseases for the comparative study.

Due the small number of subjects available for each group of MS with different clinical courses and presence of therapies at the time of inclusion in the study, we decided to analyze all the MS subject with respect to control group.

Table 3. Clinical features of MS group and demographic charateristics of MS and Control group. Duration refers to mean of the years from the diagnosis ± SD; EDSS (Expanded Disability Status Scale).

	Multiple Sclerosis			**Control**
	RR	**PP**	**SP**	
Number of subjects	37	6	5	54
Gender (M, F)	15, 22	5, 1	3, 2	23, 31
Age Mean ± SD	37.8 ± 9.2	53.2 ± 13.6	47.6 ± 8.5	41.0 ± 10.8
Duration Mean ± SD	8.9 ± 6.2	14.3 ± 7.1	20.4 ± 4.6	n/a
EDSS Mean ± SD	2.6 ± 1.6	5.9 ± 1.4	6.9 ± 0.7	n/a
Treatment Number of subjects	27[a]	3[b]	2[c]	n/a

3.1.2 Analysis of acidic soluble fraction of whole saliva by RP-HPLC-ESI-MS

The qualitative analysis of proteins/peptides from the acidic soluble fraction of whole saliva of MS patients and control subjects was performed by a integrate top-down and bottom-up proteomic approaches by RP-HPLC high resolution ESI-MS and ESI-MS/MS experiments.

On the contrary the quantitative analysis was performed by top-down RP-HPLC low resolution ESI-MS.

3.1.3 Qualitative analysis

The present study mainly focused on the already characterized salivary proteins (Messana et al., 2008a), however by investigating the salivary proteome of MS patients and healthy controls, several masses not previously characterized were detected in both groups, in the eluting range where salivary cystatins are typically observed. The characterization of these new proteoforms was performed by a top-down RP-HPLC high resolution ESI-MS and ESI-MS/MS proteomic approach on the intact proteins present in the acidic supernatant of WS as well as on salivary enriched fractions obtained by preparative RP-HPLC. They were proteoforms of cystatin A, cystatin B, cystatin D, cystatin SN, cystatin SA and oxidized forms of cystatin S1. Some of the above reported proteoforms were also characterized by a bottom-up approach.

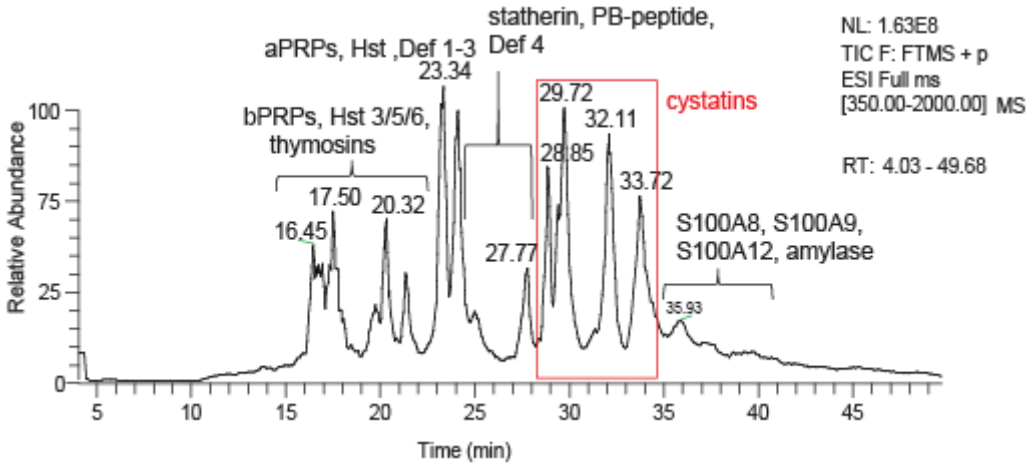


Figure 3. Typical HPLC high resolution ESI-MS profile of the acidic soluble fraction of saliva from a MS subject with the elution range of several salivary peptides/proteins.

The following sections report the structural characterization of the new proteoforms belonging to cystatin family describe in detail.

3.1.4 High-resolution top-down structural characterization of naturally occurring proteoforms of cystatins A.

In the peak eluting between 28.1-29.2 minutes of the HPLC high-resolution ESI-MS profile shown in panel A of figure 4, we identified novel proteoforms of cystatins A (experimental monoisotopic at $[M+H]^+$ 11000.7 *m/z*). In the first part of the peak (28.0-29.0 min) three proteins with experimental monoisotopic ions $[M+H]^+$ at 11030.6 ± 0.2 *m/z*, 11042.6 ± 0.2 *m/z* and 11072.7 ± 0.2 *m/z* were detected together cystatin A. High-resolution and deconvoluted mass spectra of these proteins are shown in panels B and C, respectively. The observed difference between the experimental mass value of cystatin A and the three co-eluting proteins could be justified by considering the presence of PTMs or some modified amino acidic residue not previously describe by the protein database. The mass difference of +42.06 Da between the mass of the cystatin A 11000.7 ± 0.2 *m/z* and the mass 11042.6 ± 0.2 *m/z* could be attributed to an acetylation at the N-terminus. Moreover, we have supposed that the mass 11030.6 ± 0.2 *m/z* and 11072.7 ± 0.2 *m/z* could be a natural variant of cystatin A with the substitution Thr_{96}→Met (already genetically described and available in the database http://www.uniprot.org/uniprot/P01040) and its acetylation form.

To characterize the proteins, an enriched fraction containing the proteoforms of cystatin A was collected by preparative RP-HPLC in the elution range of 29.2-32.0 min, and submitted to HPLC high-resolution ESI-MS/MS top-down analysis.

The protein with experimental monoisotopic ion $[M+H]^+$ at 11042.6 ± 0.2 *m/z* was identified as the proteoform of cystatin A acetylated at the N-terminus (theor. monoisotopic $[M+H]^+$ 11042.68 *m/z*) on the basis of high-resolution MS/MS spectra performed on the ions 921.73 ± 0.02 *m/z* ($[M+12H]^{+12}$) and 1128.64 ± 0.01 *m/z* ($[M+9H]^{+9}$). The protein with experimental monoisotopic ion $[M+H]^+$ at 11030.6 ± 0.2 *m/z* was identified as cystatin A Thr_{96}→Met (theoretical monoisotopic $[M+H]^+$ 11030.66 *m/z*) on the basis of high-resolution MS/MS spectra performed on the ion 920.73 ± 0.02 *m/z* ($[M+12H]^{+12}$) as shown by the high-resolution MS/MS annotated spectra reported in figure 5. The *y* and *b* ions series confirmed the presence of the methionine residue at position 96, in particular the fragment ion b_{96} [10808.60 *m/z*] provides definitive evidence for this modification. The protein with experimental monoisotopic ion $[M+H]^+$ at 11072.7 ± 0.2 *m/z* was tentatively identified as being the N-terminal acetylated form of cystatin A Thr_{96}→Met (theoretical monoisotopic ion $[M+H]^+$ 11072.67 *m/z*), but the fragmentation spectrum did not allow to confirm the structure.

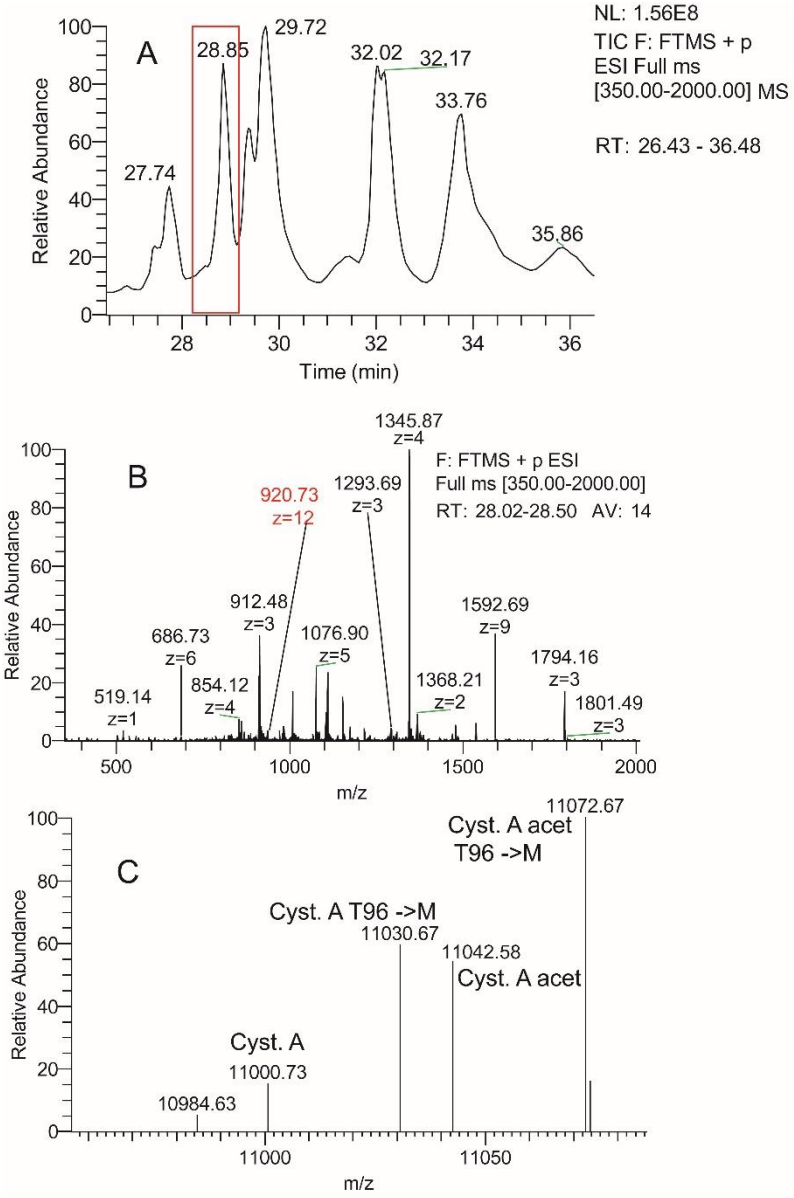


Figure 4. Characterization of cystatin A proteoforms. Enlargement, in the elution time 26.6-36.4 min (panel A), of the HPLC high-resolution ESI-MS profile shown in Figure 3, boxed peak corresponds to cystatins A. High-resolution mass spectra of cystatin A (panel B) proteoforms with the corresponding deconvoluted spectra (panel C). Red *m/z* values in panel B have been used for high-resolution MS/MS characterization of the proteoforms.

Figure 5. Cystatin $A_{T96\rightarrow M}$. Annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+12H]^{12+}$ 920.73 *m/z*.

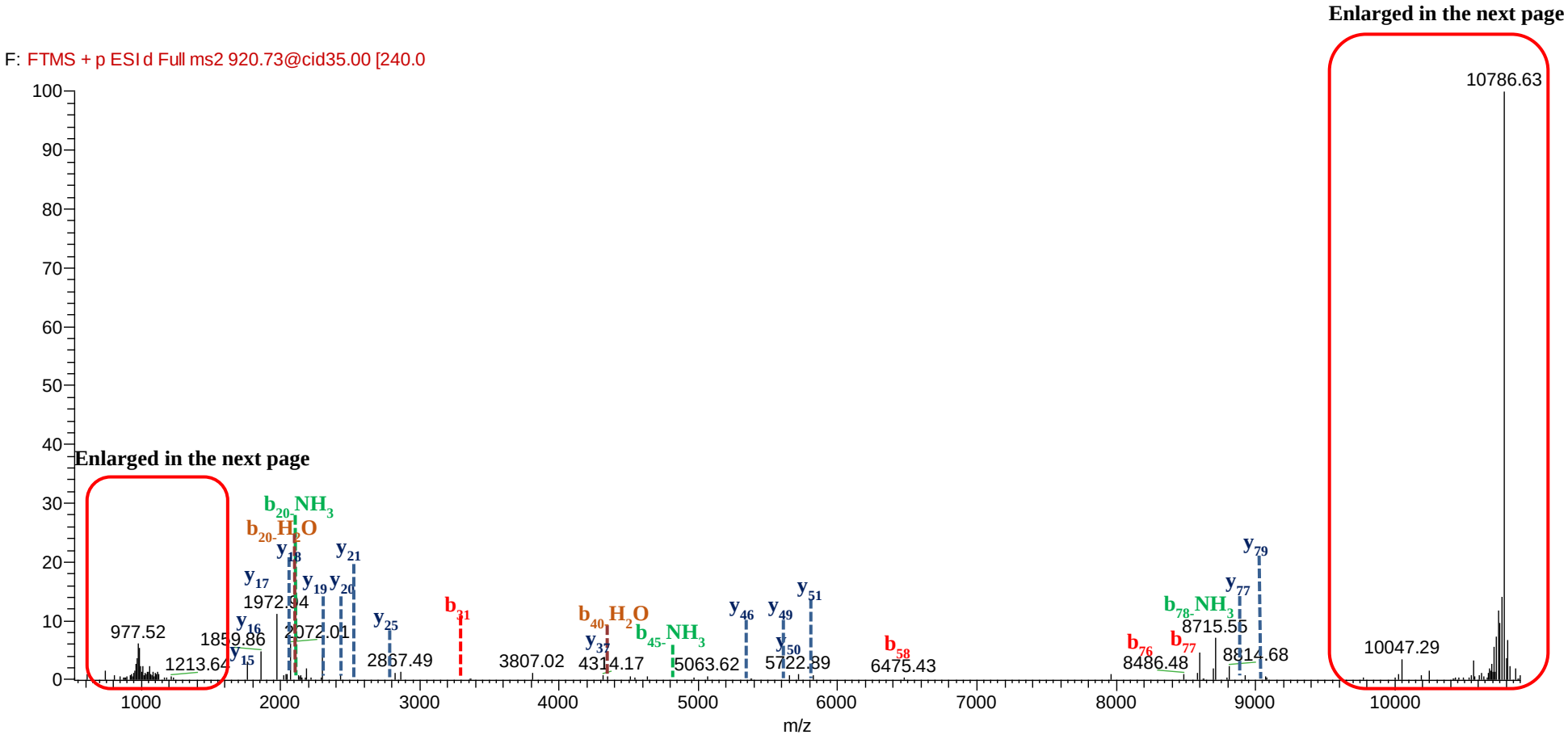


Figure 5a. Cystatin $A_{T96\rightarrow M}$. Enlargement in the mass range 500-1400 *m/z* of the annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+12H]^{12+}$ 920.73 *m/z*.

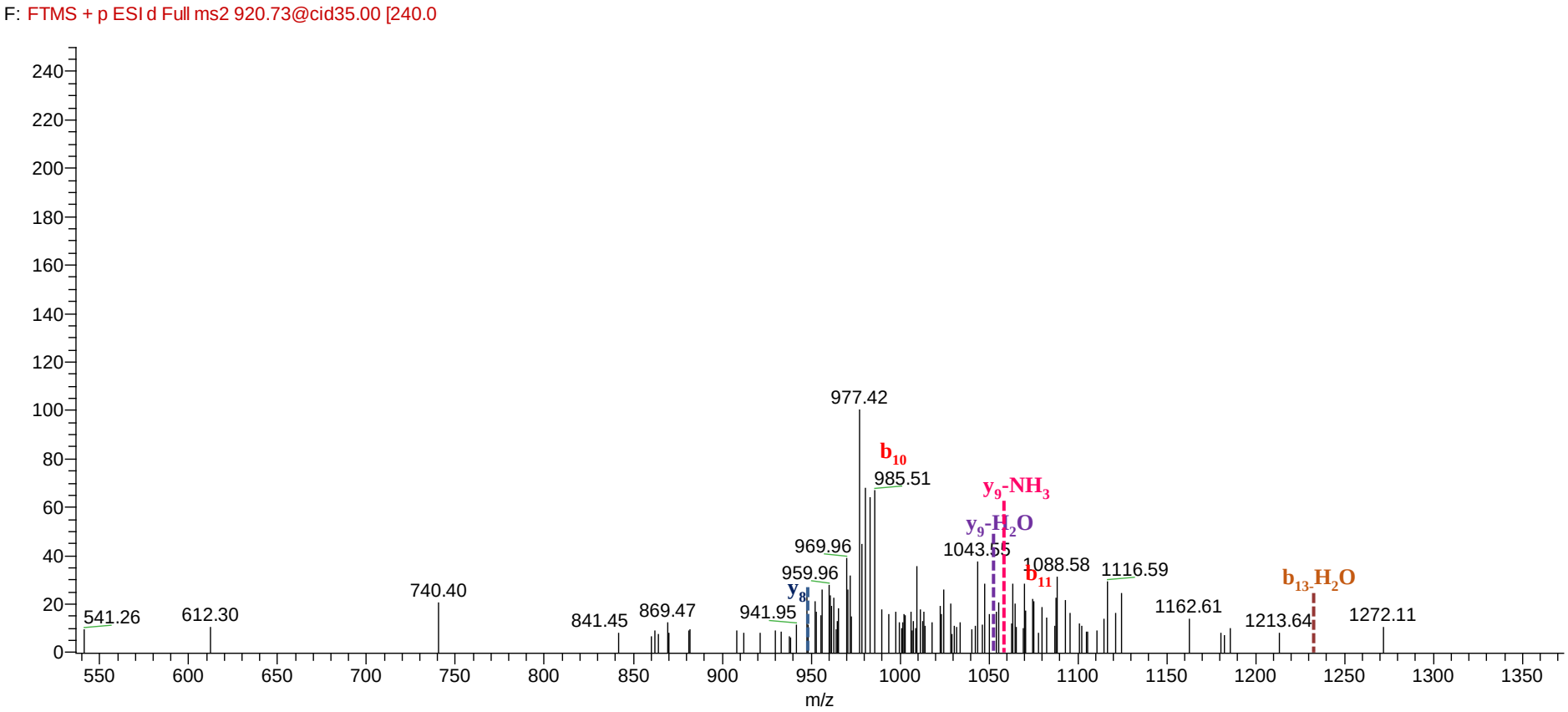
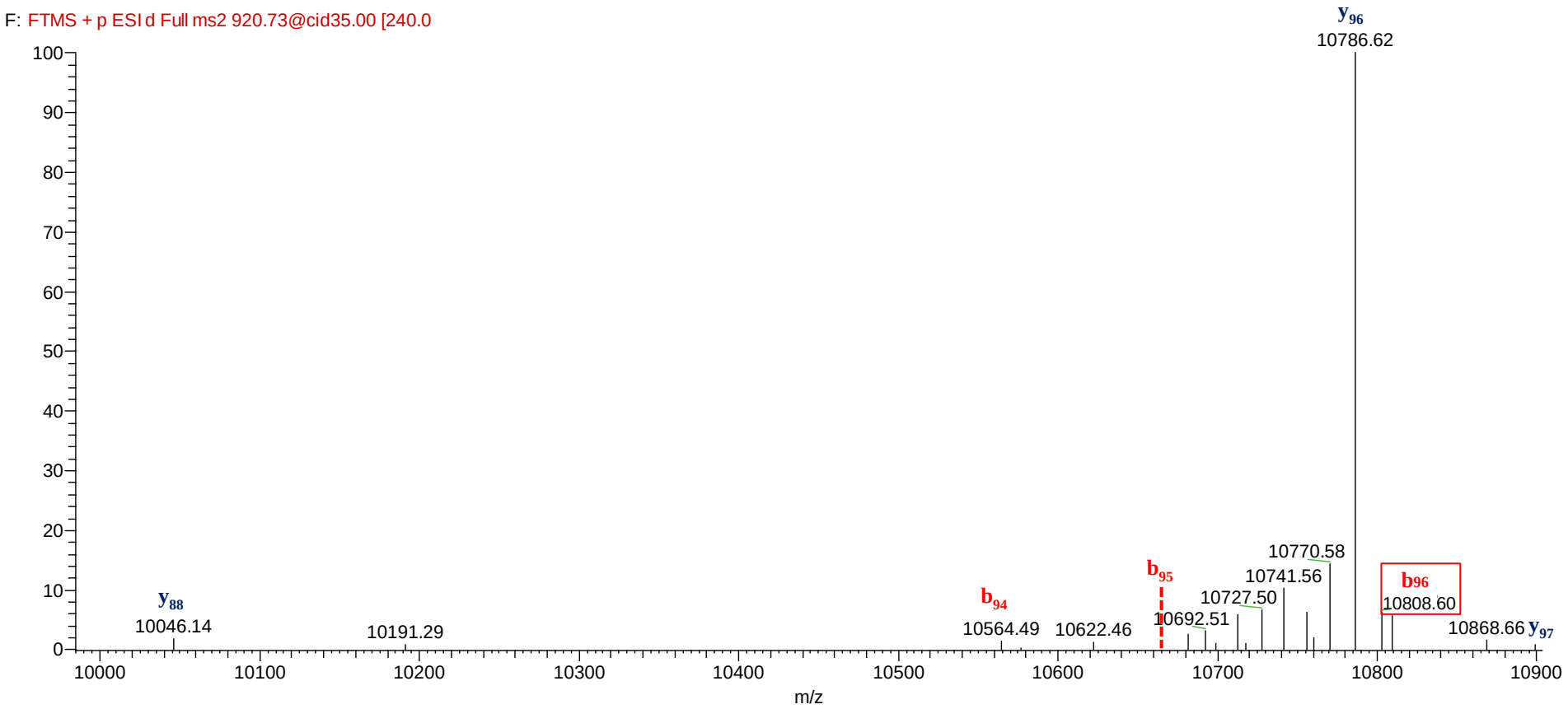


Figure 5b. Cystatin $A_{T96\rightarrow M}$. Enlargement in the mass range 10000-11000 *m/z* of the annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+12H]^{12+}$ 920.73 *m/z*.



3.1.5 High-resolution top-down structural characterization of naturally occurring proteoforms of cystatins B.

In the peak eluting between 28.1-29.2 minutes of the HPLC high-resolution ESI-MS profile shown in panel A of figure 7, a protein with monoisotopic ion $[M+H]^+$ at 11233.7 ± 0.2 *m/z* was detected together S-glutathionylated and S-cysteinylated cystatin B. Panel B reports the high-resolution mass spectra of cystatin B proteoforms with the corresponding deconvoluted spectra (panel C). To characterize the protein, an enriched fraction containing the protein co-eluting with cystatin B proteoforms was separately collected by preparative RP-HPLC in the elution range of 32.0-33.3 min, and submitted to HPLC high-resolution ESI-MS/MS top-down analysis. The monoisotopic ion $[M+H]^+$ at 11233.7 ± 0.2 *m/z* was characterized as cystatin B acetylated at the N-terminus, with Cys_3 converted to carboxymethyl cysteine (CMC) (theoretical monoisotopic ion $[M+H]^+$ at 11233.61 *m/z*) confirmed by top-down high-resolution MS/MS performed on the ion $[M+11H]^{+11}$ at 1022.61 ± 0.01 *m/z* as shown by the high-resolution MS/MS annotated spectra reported in figure 8. Particularly, the fragment ion y_{96} [10929.57 *m/z*] confirmed the Cys_3 modification.

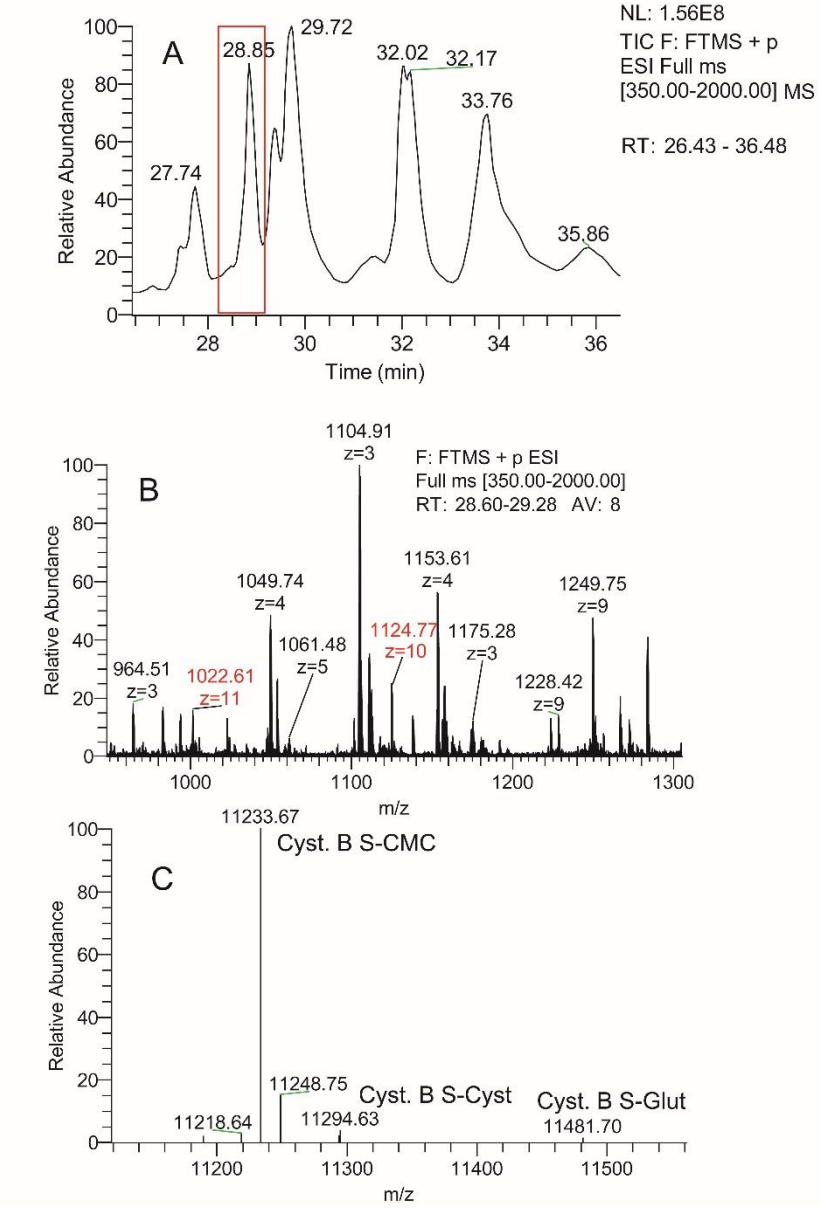
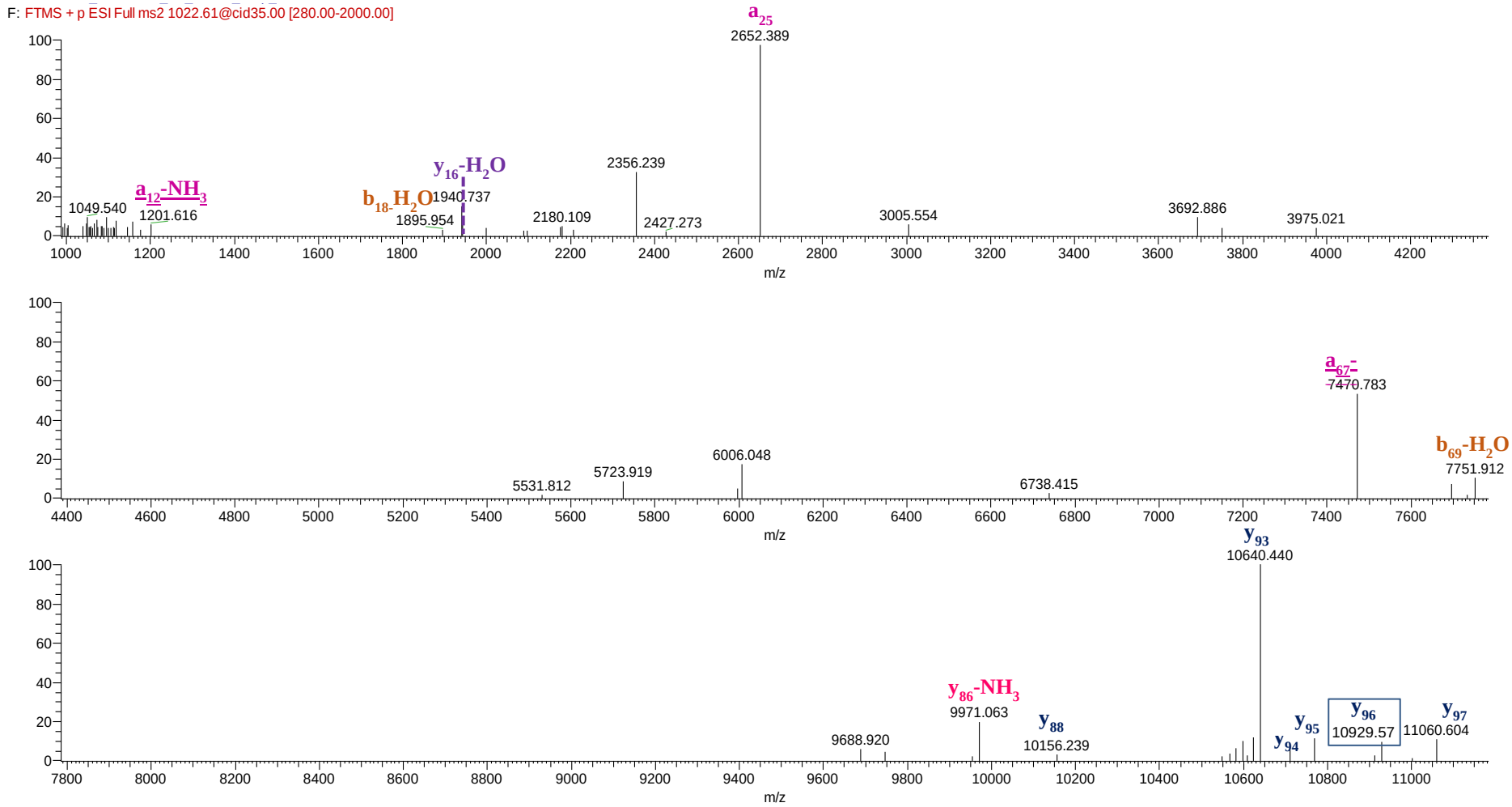


Figure 7. Characterization of cystatin B proteoform. Enlargement, in the elution time 26.6-36.4 min (panel A), of the HPLC high-resolution ESI-MS profile shown in Figure 3, boxed peak correspond to cystatins B. High-resolution mass spectra of cystatin B (panel B) proteoforms with the corresponding deconvoluted spectra (panel C). Red *m/z* values in panel B have been used for high-resolution MS/MS characterization of the proteoform.

Figure 8. **Cystatin B-CMC**. Annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+11H]^{11+}$ 1022.61 *m/z*.



3.1.6 High-resolution top-down and bottom-up structural characterization of cystatin D proteoforms.

In the peak eluting between 31.0-32.0 minutes of the HPLC high-resolution ESI-MS profile shown in panel A of figure 10, we observed a protein with experimental $[M+H]^+$ monoisotopic ion 13509.7 ± 0.2 *m/z* whose high-resolution mass and deconvoluted spectra are reported in panels B and C respectively. The manual inspection of the high resolution MS/MS spectrum on deca *m/z* ion 1352.7 ± 0.02 allowed us to obtain a partial sequence of ten amino acid residues (FAISEYNKVI) that was used to query the bioinformatic software BLAST. BLAST program proposed the human cystatin D as the best attribution with a score of 35.8 and an E-value of 7^{-04}. The theoretical mass value of cystatin D reported in Swiss-Prot protein database (theoretical monoisotopic ion $[M+H]^+$ at 13850.7 ± 0.2 *m/z*) did not coincide however with the experimental mass of $[M+H]^+$ monoisotopic ion 13509.7 ± 0.2 *m/z*. The mass difference value of -341 Da could be justified by considering the presence of PTMs not described by the protein database.

The structural characterization of cystatin D was performed by an integrated top-down and bottom-up approach.

Top-down high-resolution MS/MS experiments performed on the ion 1352.67 ± 0.02 *m/z* ($[M+10H]^{+10}$) allowed evidencing a new proteoform of cystatin D $C_{26}\rightarrow R$ lacking the first 5 amino acids from the N-terminus (Des_{1-5}), with the N-terminal glutamine residue converted to pyro-glutamic acid and with two disulfide bonds (experimental Monoisotopic ion $[M+H]^+$ at 13509.2 ± 0.2 *m/z*) as shown by the high-resolution MS/MS annotated spectra reported in figure 12. The top-down approach yielded extensive sequence coverage of the N-terminus of cystatin D proteoform, allowing defining the truncation extents, but the core of the proteins located between the two disulfide bonds and R_{26} was not completely covered. For this reason, a fraction of this protein was collected by preparative RP-HPLC in the elution range of 36.0-37.0 min, submitted to trypsin digestion after reduction and alkylation with IAM, and peptides analysed by high-resolution MS/MS analysis. In particular, MS/MS sequencing performed on the ion 475.01 ± 0.02 *m/z* ($[M+4H]^{+4}$) allowed confirming the $C_{26}\rightarrow R$ substitution in the sequence of the tryptic peptide 9-26 (T_9LAGGIHATDLNDKSVQR$_{26}$) with monoisotopic $[M+H]^+$ at 1895.9 ± 0.02 *m/z*.

In the HPLC high-resolution ESI-MS spectrum profile of some saliva samples only, two more mass values of experimental monoisotopic ion $[M+H]^+$ at 13596.7 ± 0.2 *m/z* and $[M+H]^+$ at 13155.5 ± 0.2 *m/z* in the same elution time of cystatin D were detected. The

experimental monoisotopic ion $[M+H]^+$ at 13596.7 ± 0.2 *m/z* was tentatively attributed to the proteoform of cystatin D C_{26}→R lacking the first 4 amino acids from the N-terminus (Des_{1-4}) and carrying two disulfide bonds (theor. Monoisotopic ion $[M+H]^+$ at 13596.71 *m/z*), but the low intensity of the ions did not permit to confirm the attribution. High-resolution MS/MS experiments performed on the ion 1463.73 ± 0.02 *m/z* ($[M+9H]^{+9}$) allowed confirming the sequence of the proteoform of cystatin D C_{26}→R lacking the first 8 amino acids from the N-terminus (Des_{1-8}) with two disulfide bonds (theor. Monoisotopic ion $[M+H]^+$ at 13155.48*m/z*) as shown by the high-resolution MS/MS annotated spectra reported in figure 11.

It is noteworthy that we were not able to detect in the acidic soluble fraction of human saliva the mass corresponding both to the entire proteoforms of cystatin D C_{26}→R and to cystatin with C_{26}, probably either because this proteoform may be prone to precipitation under our experimental condition or because the C_{26} could carry unknown modifications.

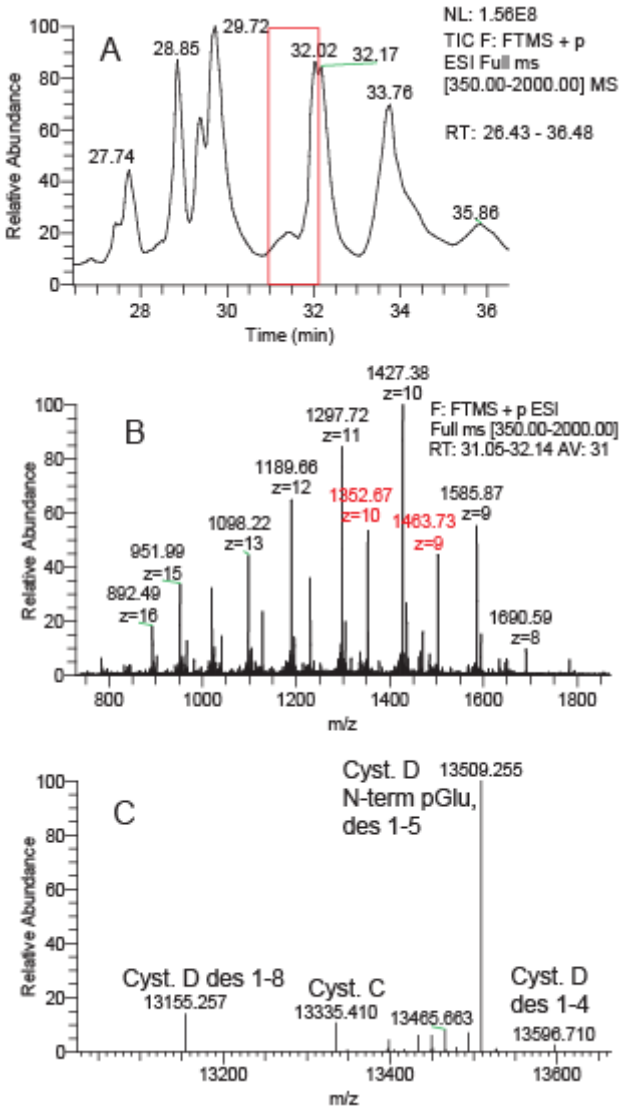


Figure 10. Characterization of cystatin D proteoforms. Enlargement, in the elution time 26.6-36.4 min (panel A), of the HPLC high-resolution ESI-MS profile shown in Figure 3, boxed peak corresponds to cystatins C and D. High-resolution mass spectra (panels B) and the corresponding deconvoluted spectra (panels C) of the cystatin C and cystatin D proteoforms respectively. Red *m/z* values in panel B have been used for high-resolution MS/MS characterization of the proteoforms.

Figure 11. Cystatin $D_{C46\rightarrow R}$ Des 1-8. Annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+9H]^{9+}$ 1463.73 *m/z*.

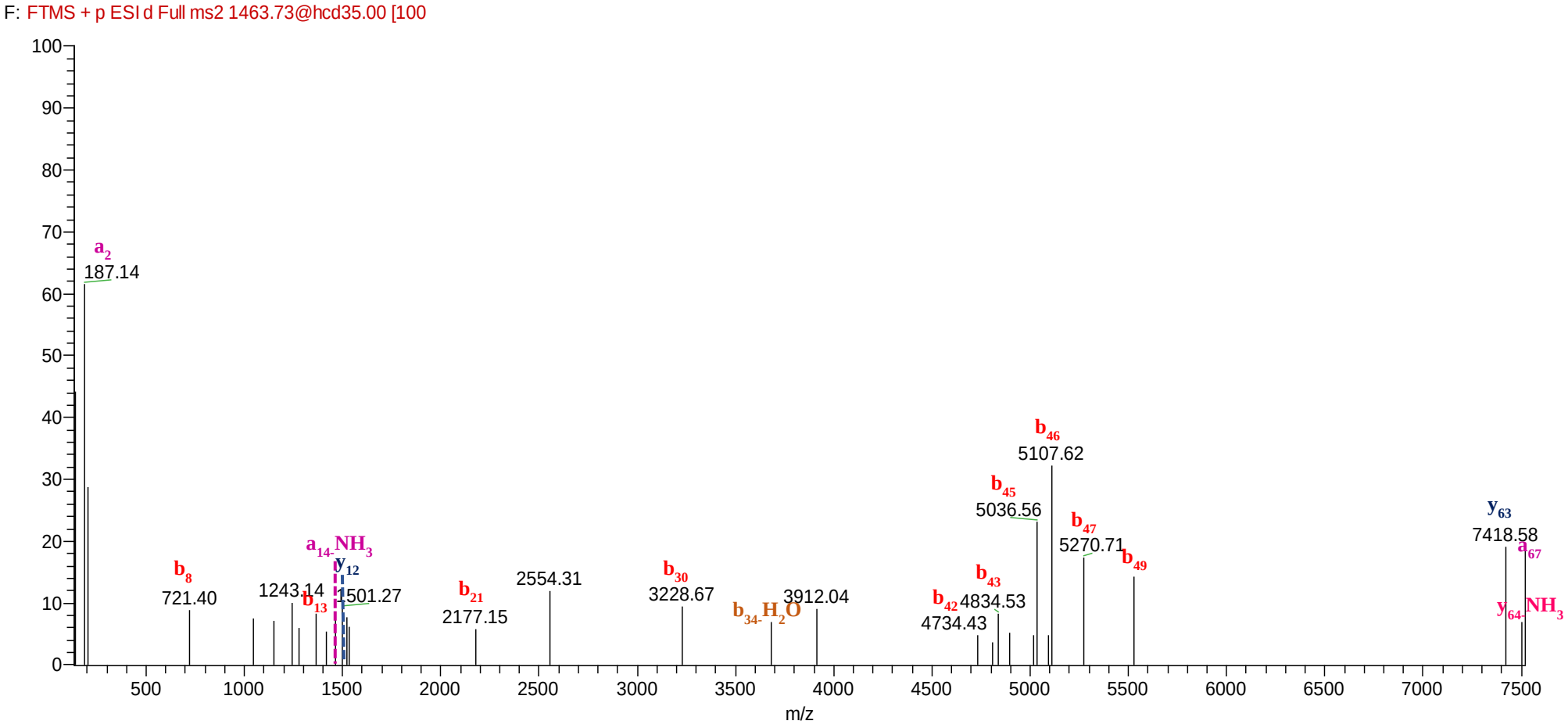
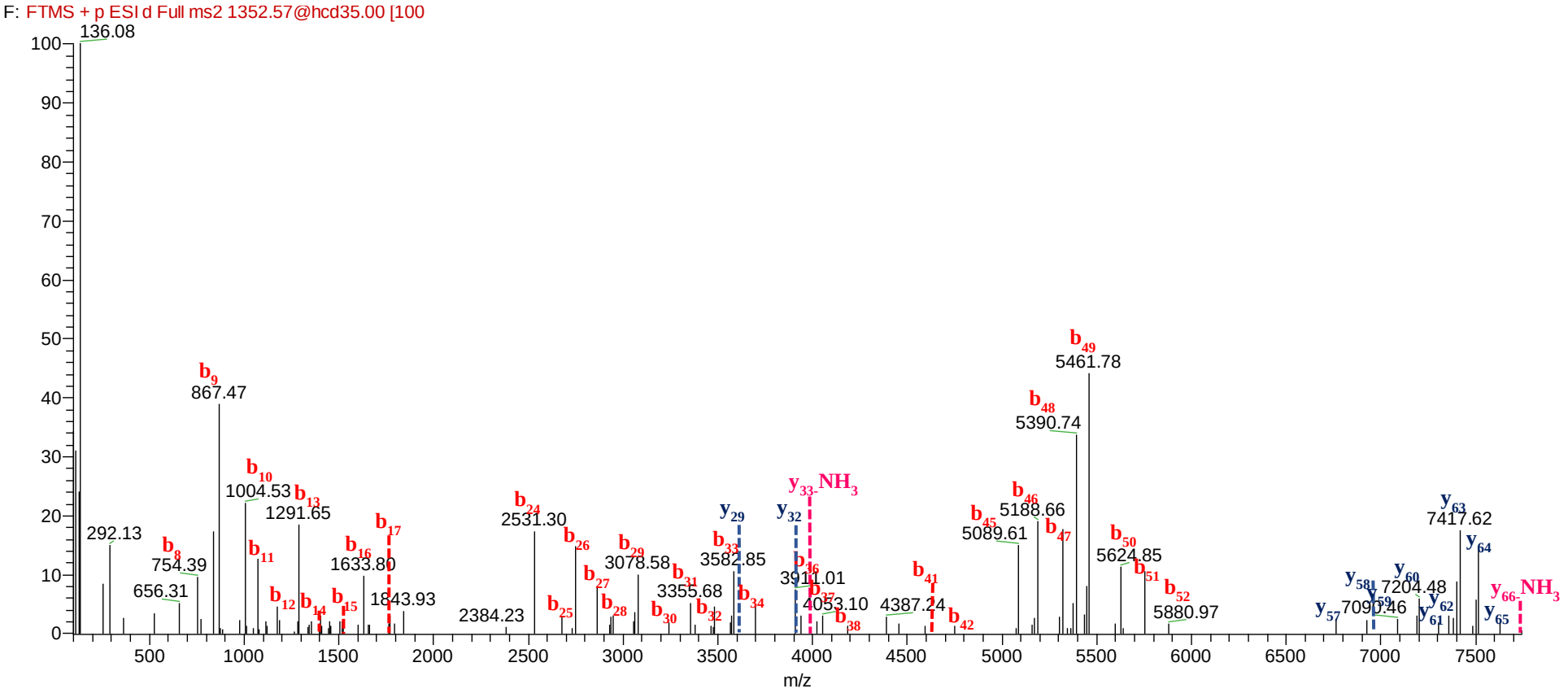


Figure 12. Cystatin $D_{C46\rightarrow R}$ Des 1-5, N-terminal pGlu. Annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+10H]^{10+}$ 1352.57 *m/z*.



3.1.7 High-resolution top-down structural characterization of cystatins SN proteoforms.

In the peak eluting between 31.5-32.5 minutes of the HPLC high-resolution ESI-MS profile shown in figure 14 panel A, we observed two proteins with experimental $[M+H]^+$ monoisotopic ion 13805.8 ± 0.2 *m/z* and $[M+H]^+$ 13432.7 ± 0.2 *m/z* co-eluting with cystatin SN (theoretical $[M+H]^+$ monoisotopic ion 14304.09 ± 0.2 *m/z*). The high-resolution mass and deconvoluted spectra of these proteins are reported in panel B and C respectively.

We have suppose that the mass difference of -498.26 and -871.37 Da between the mass of cystatin SN and the mass 13805.8 ± 0.2 *m/z* and 13432.7 ± 0.2 *m/z* respectively, could be attributed to a N-terminal cleavage of cystatin SN. The monoisotopic ion $[M+H]^+$ at 13805.8 ± 0.2 *m/z* was attributed to the proteoform of cystatin SN lacking the first 4 amino acids from the N-terminus (cystatin SN Des_{1-4}) and carrying two disulfide bonds (theor. monoisotopic $[M+H]^+$ 13805.84 *m/z*). The sequence was confirmed by top-down high-resolution MS/MS experiments performed on the ion 1256.72 ± 0.02 *m/z* ($[M+11H]^{+11}$) as shown by the high-resolution MS/MS annotated spectra reported in figure 15. In addition, the monoisotopic ion $[M+H]^+$ at 13432.7 ± 0.2 *m/z* was attributed to the proteoform of cystatin SN lacking the first 7 amino acids from the N-terminus (cystatin SN Des_{1-7}) and carrying two disulfide bonds (theor. monoisotopic $[M+H]^+$ 13432.72 *m/z*). High-resolution MS/MS experiments performed on the ion 1034.76 ± 0.02 *m/z* ($[M+13H]^{+13}$) allowed confirming the sequence (figure 16). Furthermore, two proteoforms with the mass differences of +16 and +32 Da with respect to cystatin SN suggested the presence of mono-oxidized and di-oxidized forms of this protein. In order to define the site of oxidation, the soluble fraction of saliva samples containing a high amount of the oxidized proteoforms, was submitted to top-down high-resolution MS/MS experiments using both HCD- and CID-based fragmentations. High-resolution HCD-based MS/MS spectrum of the $[M+11H]^{+11}$ ion 1103.17 ± 0.02 *m/z* allowed confirming a partial sequence of the mono-oxidized proteoform of cystatin SN (exp. monoisotopic $[M+H]^+$ 14320.1± 0.2 *m/z*), but failed in the identification of the amino acid involved in the oxidation, as shown by the high-resolution MS/MS annotated spectra reported in figure 17. More complete coverage was obtained by a selected ion monitoring (SIM) experiment carried out on 1194.85 ± 0.02 *m/z* ($[M+12H]^{+12}$) with CID fragmentation, as shown by the high-resolution MS/MS annotated spectra reported in figure 18. The y_{99} ion [11807.90 *m/z*] confirmed tryptophan (W_{23}) as oxidized residue.

During this study, we were able to detect and confirm the structure of the natural variant of cystatin SN with the substitution P_{11}→L (experimental monoisotopic ion $[M+H]^+$ 14320.1 ± 0.2 *m/z*) (already genetically described and available in the database http://www.uniprot.org/uniprot/P01037) and similarly to cystatin SN, characterized for the first time two of its N-terminally truncated proteoforms (experimental monoisotopic ions $[M+H]^+$ 13821.8 ± 0.2 *m/z* and 13448.7 ± 0.2 *m/z*). Panel D of figure 14 displays the high-resolution mass spectra and the corresponding deconvoluted spectra (panel E) of these proteoforms. The monoisotopic $[M+H]^+$ value 14320.1 ± 0.2 *m/z* was in agreement with the natural variant of cystatin SN P_{11}→L (theor. monoisotopic $[M+H]^+$ 14320.13 *m/z*) and the sequence was confirmed by top-down high-resolution MS/MS experiments performed on the ion 956.22 ± 0.01 *m/z* ($[M+15H]^{+15}$) as shown by the high-resolution MS/MS annotated spectra reported in figure 19. Among the *b* fragment ions, the diagnostic b_{11} ion [1367.74 *m/z*] confirmed the presence of L_{11}.

High-resolution MS/MS experiments performed on the ion 1064.61 ± 0.02 *m/z* ($[M+13H]^{+13}$) as shown by the high-resolution MS/MS annotated spectra reported in figure 20 allowed confirming the sequence of the proteoform of cystatin SN P_{11}→L lacking the first 4 amino acids from the N-terminus (cystatin SN P_{11}→L Des_{1-4}) and carrying two disulfide bonds (exp. monoisotopic $[M+H]^+$13821.8 ± 0.2 *m/z*), whereas the monoisotopic $[M+H]^+$ ion at 13448.7 ± 0.2 *m/z* was consistent with the proteoform of cystatin SN P_{11}→L lacking the first 7 amino acids from the N-terminus (cystatin SN P_{11}→L Des_{1-7}) and carrying two disulfide bonds (theor. monoisotopic $[M+H]^+$ 13448.78 *m/z*), but the low intensity of the ions did not allow performing MS/MS spectra to confirm the attribution.

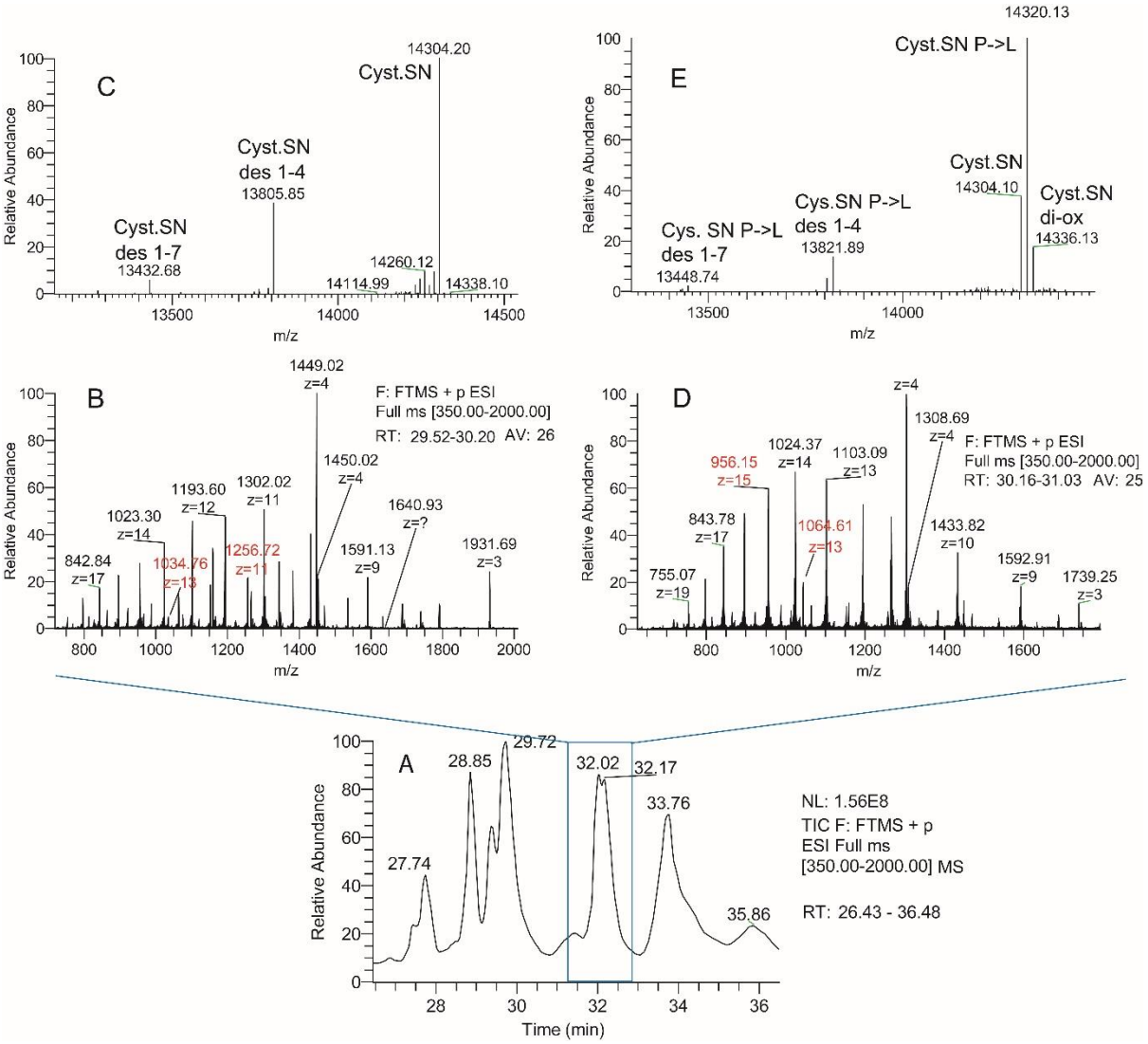


Figure 14. Characterization of cystatins SN proteoforms. Enlargement, in the elution time 26.6-36.4 min (panel A), of the HPLC high-resolution ESI-MS profile shown in Figure 3, boxed peaks correspond to S-type cystatins. High-resolution mass spectra of cystatins SN (panel B) and cystatin SN P_{11}→L (panel D) proteoforms with the corresponding deconvoluted spectra (respectively panels C, E). Red *m/z* values in panels B and D have been used for high-resolution MS/MS characterization of the proteoforms.

Figure 15. Cystatin SN Des$_{1-4}$. Annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+11H]^{11+}$ 1256.54 *m/z*.

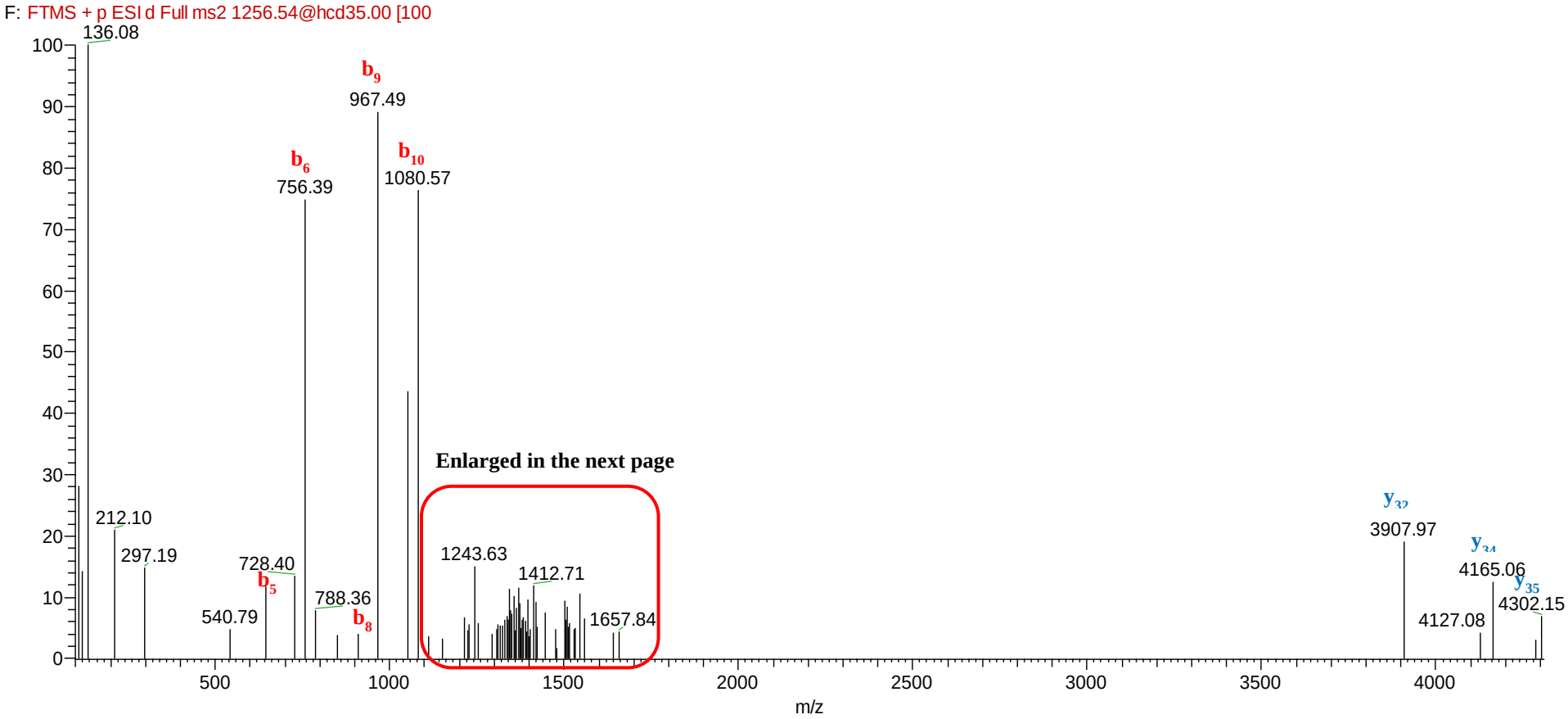


Figure 15a. Cystatin SN Des$_{1-4}$. Enlargement in the mass range 1040-1660 *m/z* of the annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+11H]^{11+}$ 1256.54 *m/z*.

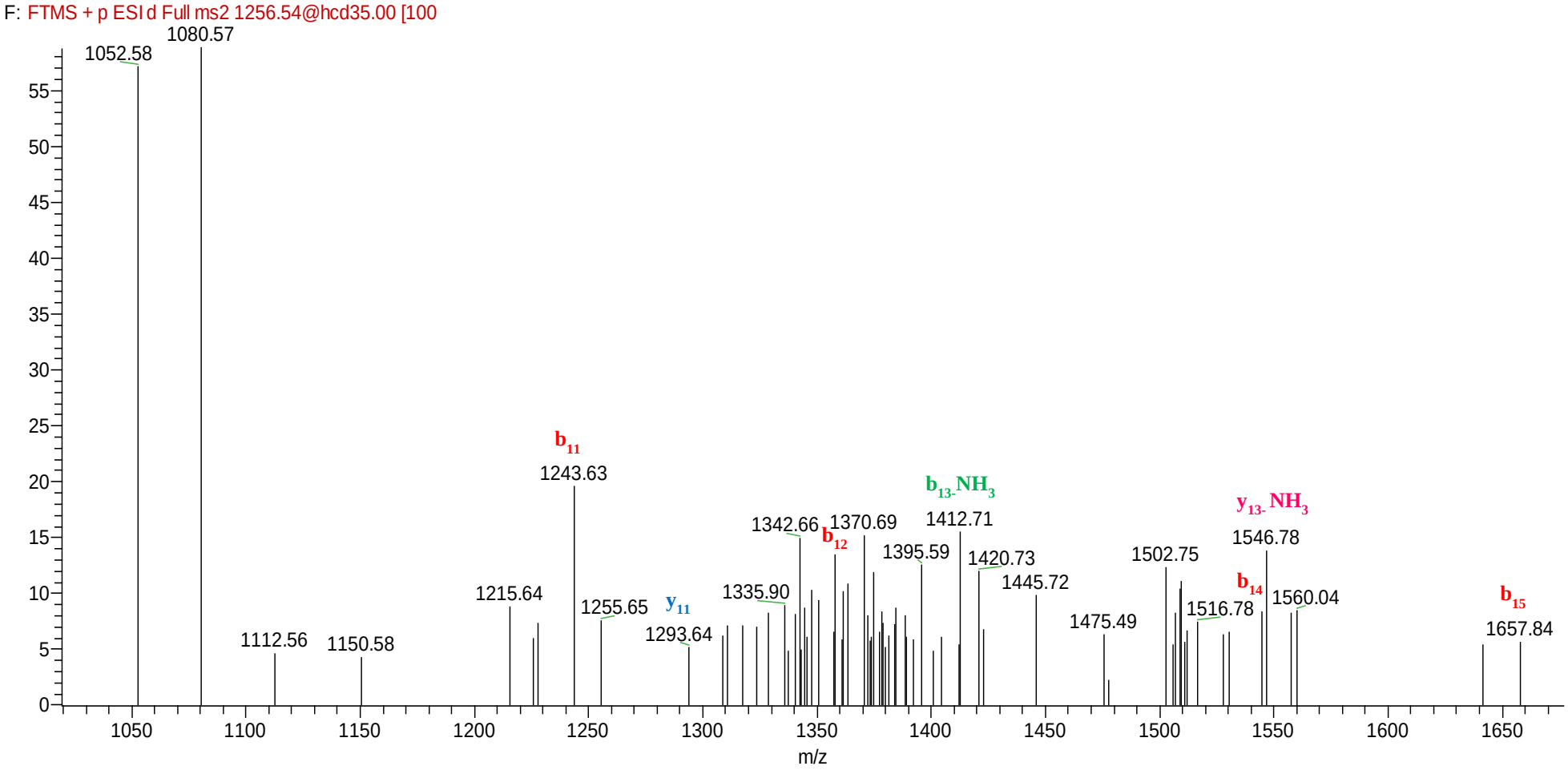


Figure 16. Cystatin SN Des$_{1-7}$. Annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+13H]^{13+}$ 1034.76 *m/z*.

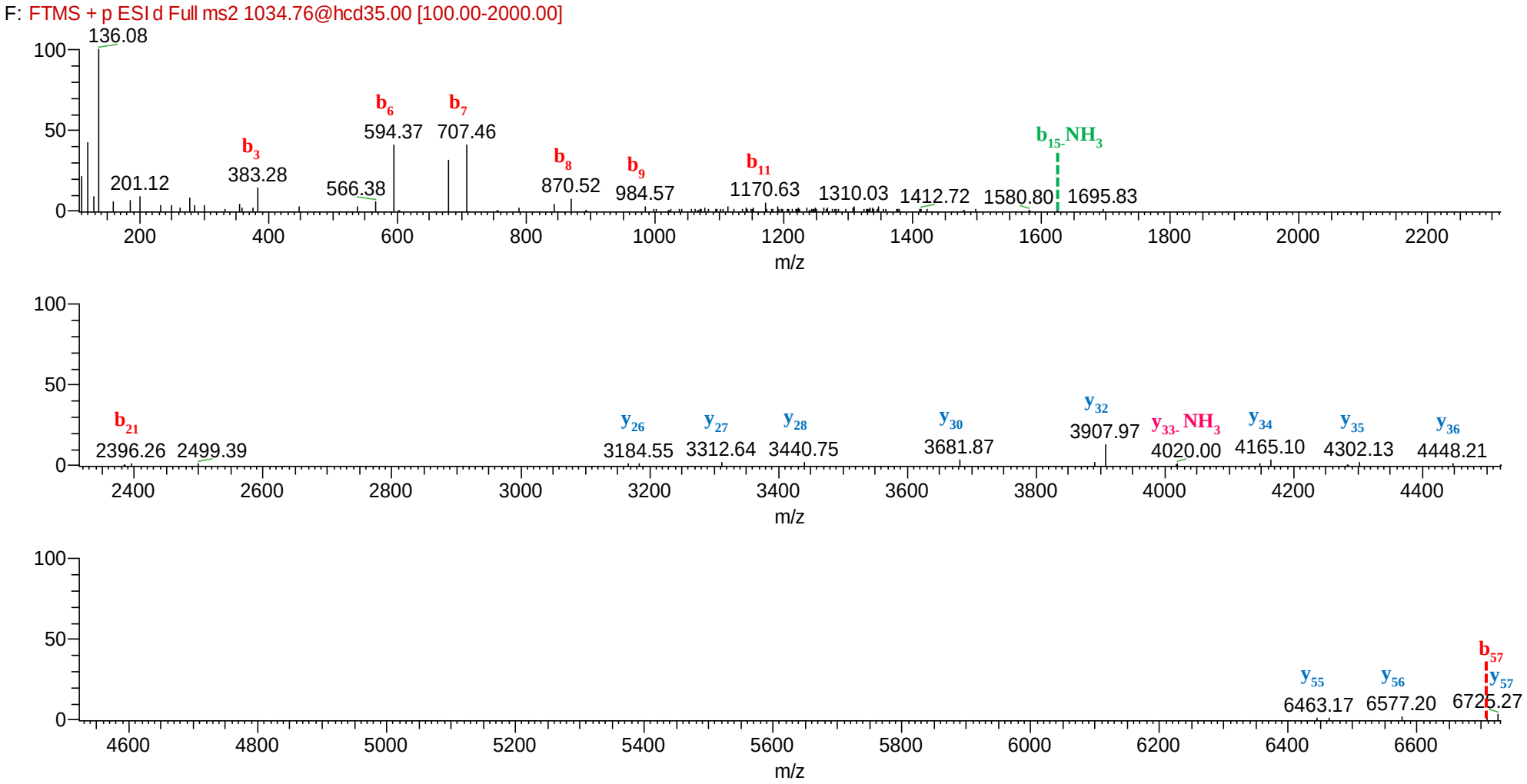


Figure 17. **Mono-oxidized Cystatin SN-W_{23}ox**. HCD annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+11H]^{11+}$ 1103.02 *m/z* in agreement with W_{23}-oxidized (+15.995 Da)

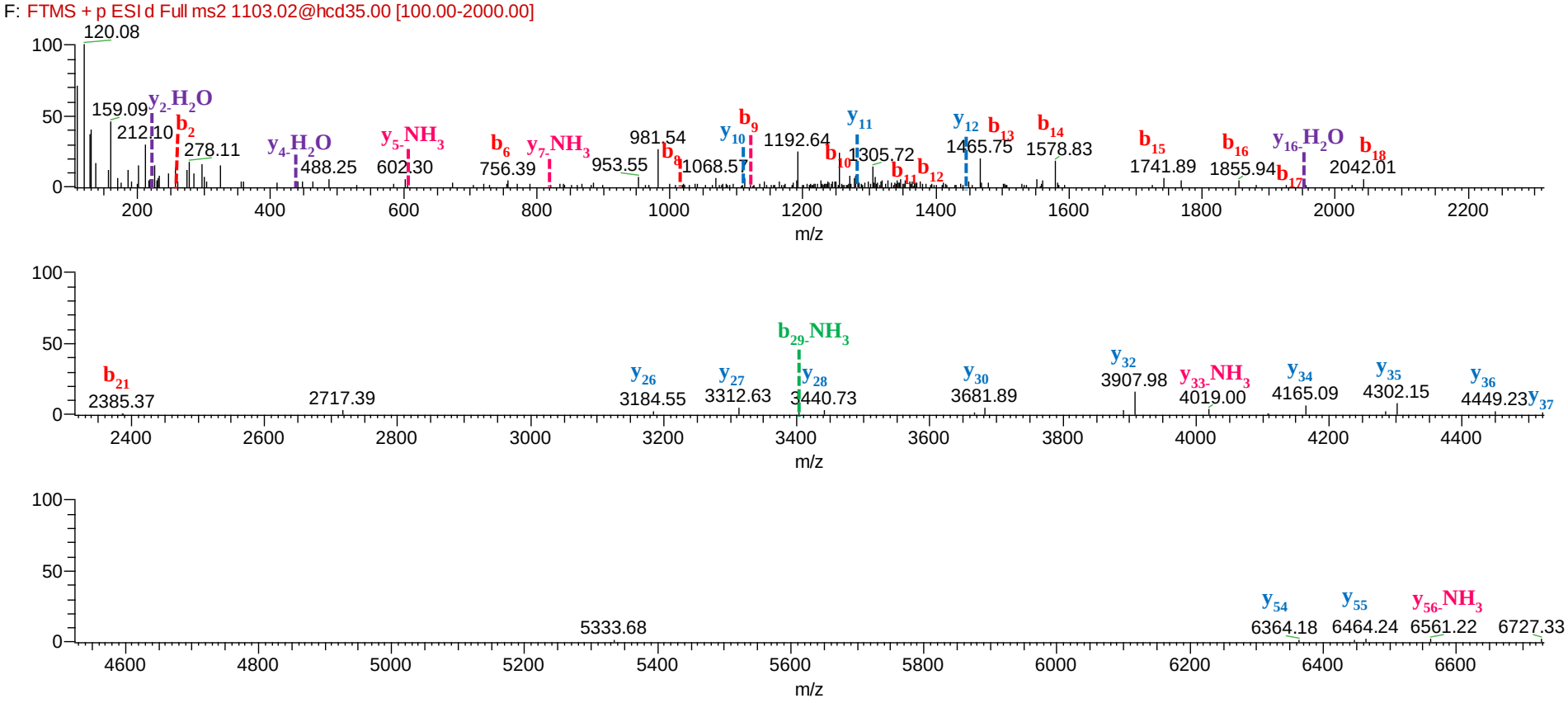


Figure 18. **Mono-oxidized Cystatin SN-W$_{23}$ox**. CID annotated MH$^+$ deconvoluted spectra of high-resolution MS/MS of the ion [M+12H]$^{12+}$ 1194.85 *m/z* in agreement with W$_{23}$-oxidized (+15.995 Da)

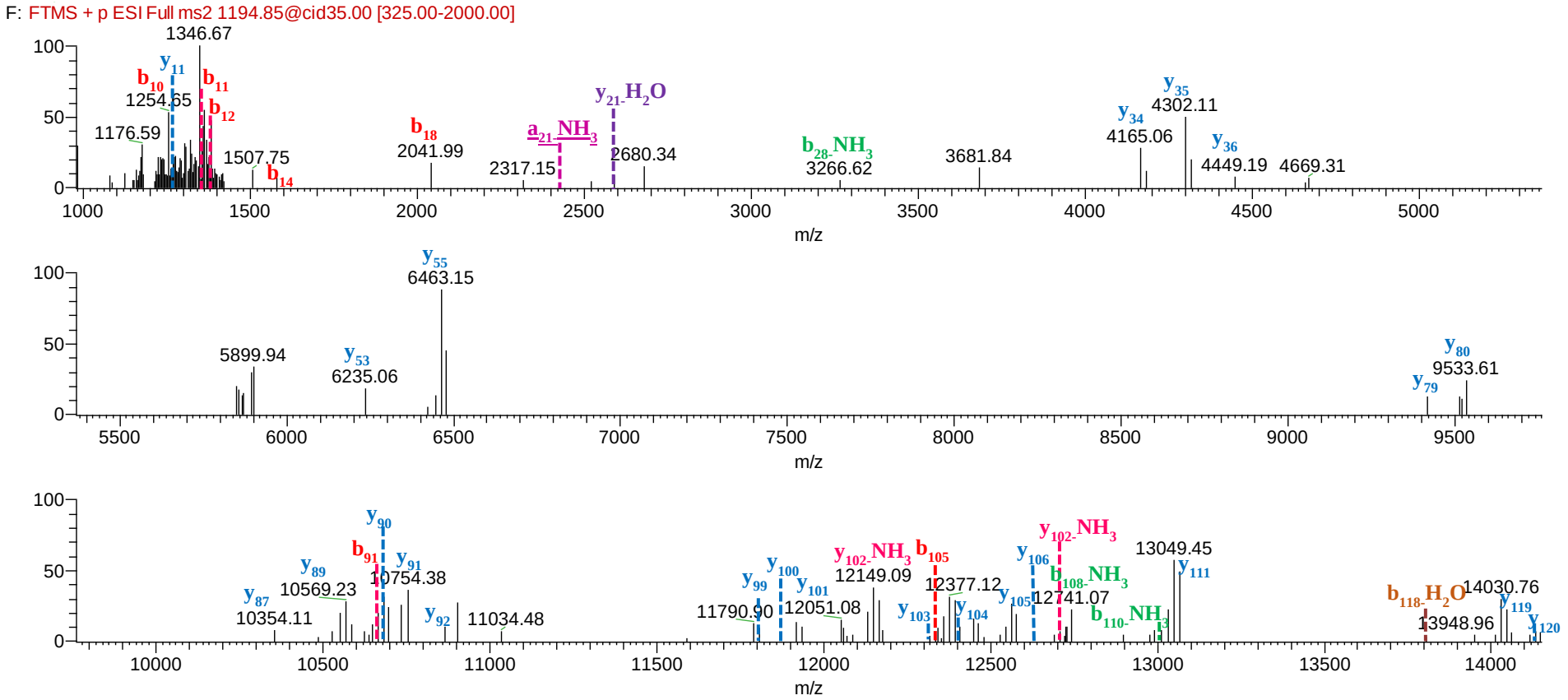


Figure 19. Cystatin SN $P_{11 \rightarrow}$L. Annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+15H]^{15+}$ 956.22 *m/z*.

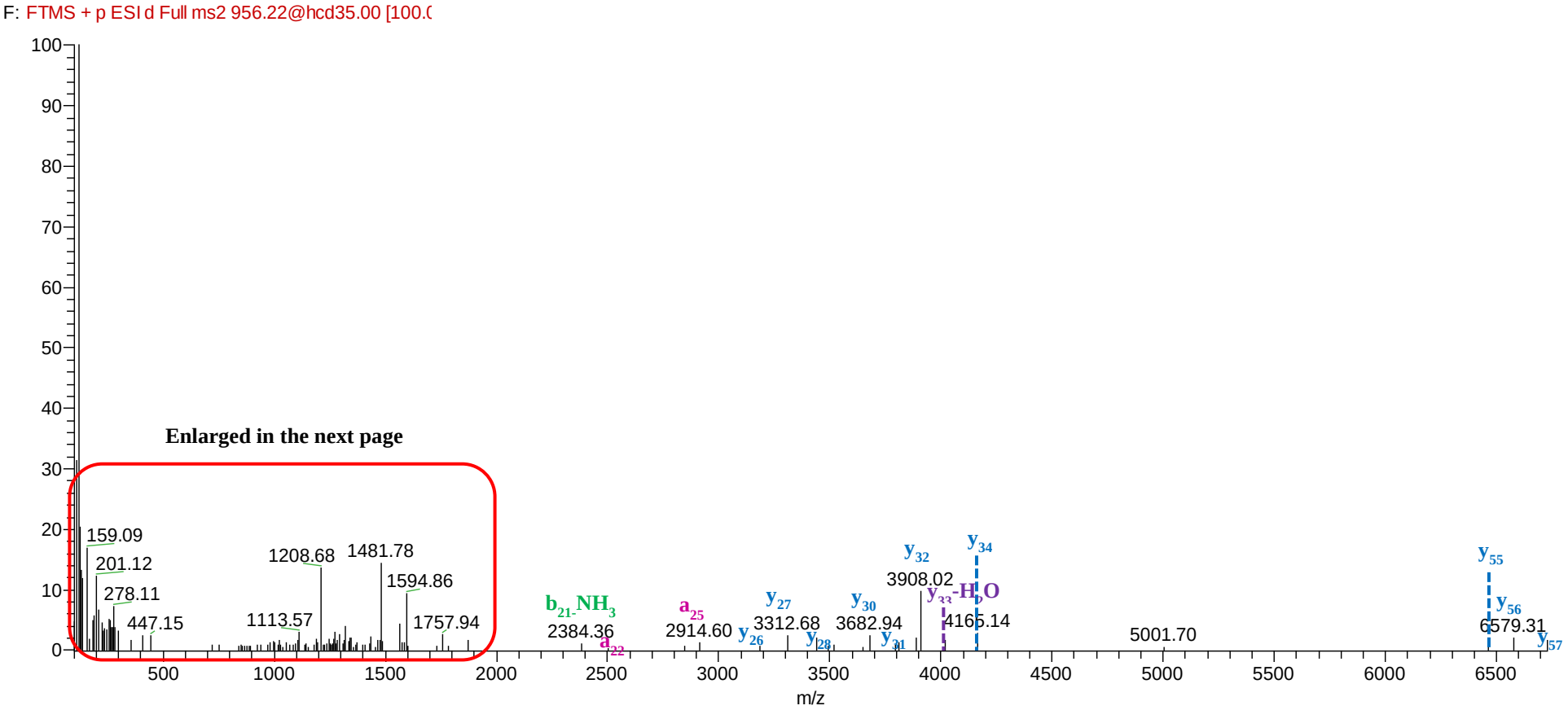


Figure 19a. Cystatin SN $P_{11 \rightarrow}L$. Enlargement in the mass range 300-2000 *m/z* of the annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+15H]^{15+}$ 956.22 *m/z*.

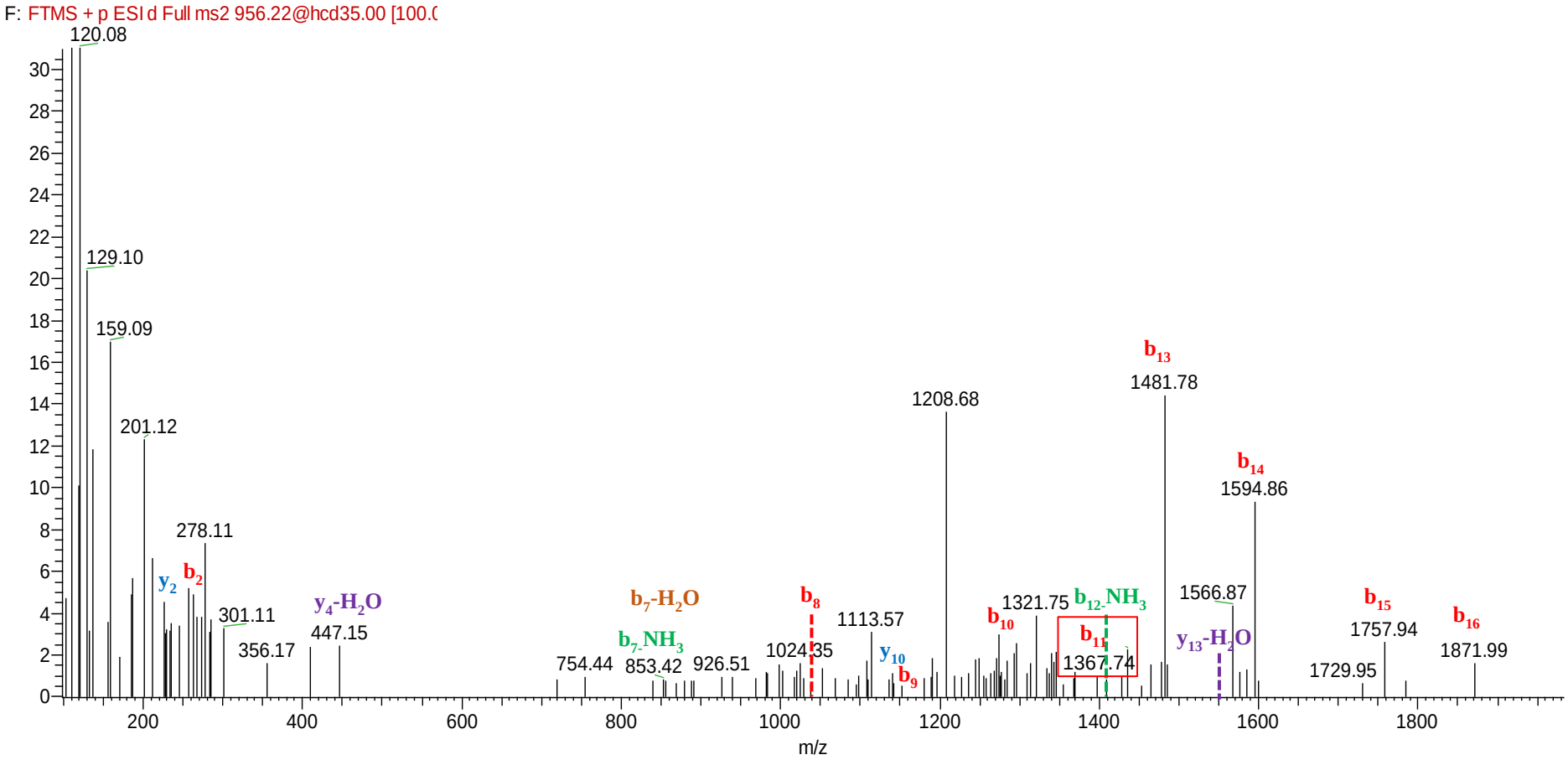


Figure 20. Cystatin SN $P_{11\rightarrow}$L $Des_{1\text{-}4}$. Annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+13H]^{13+}$ 1064.61 *m/z*.

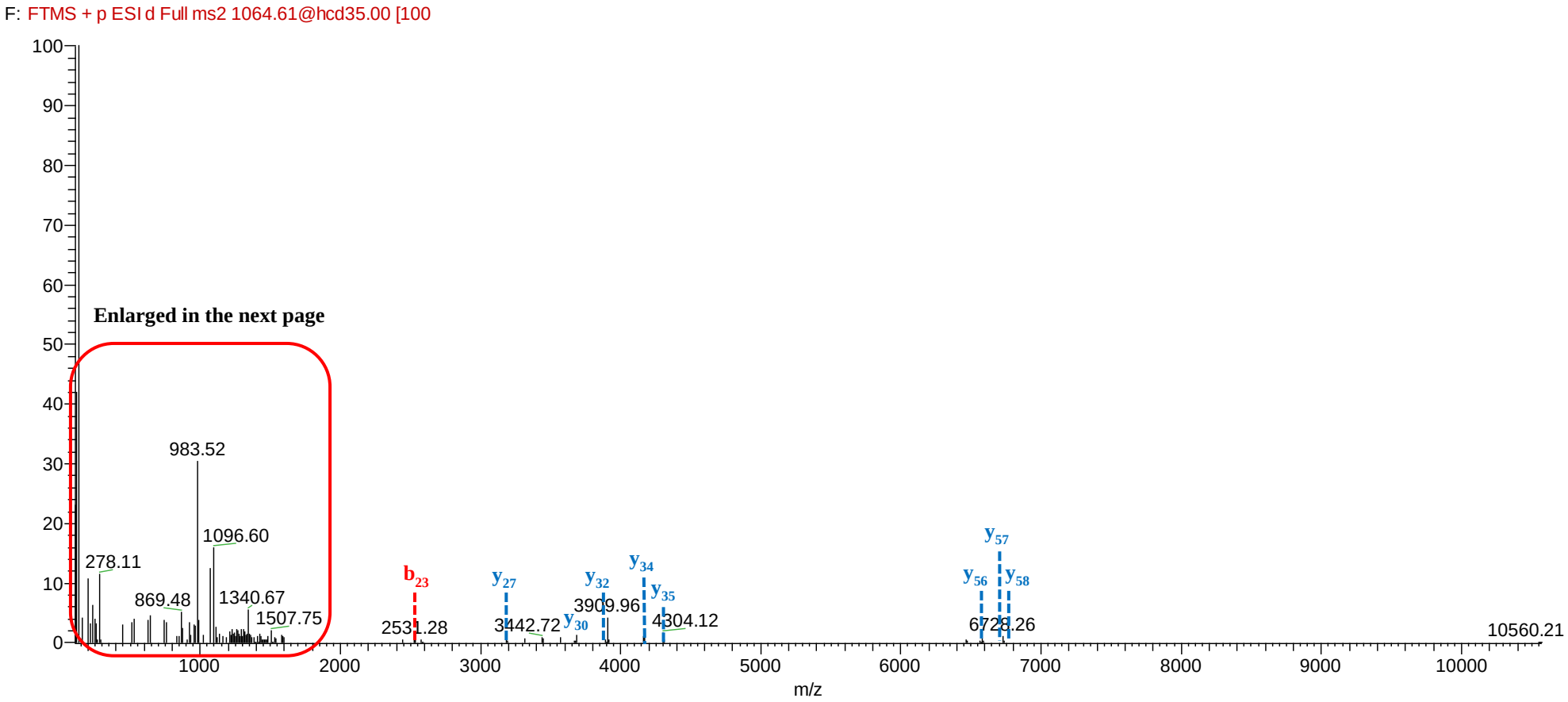
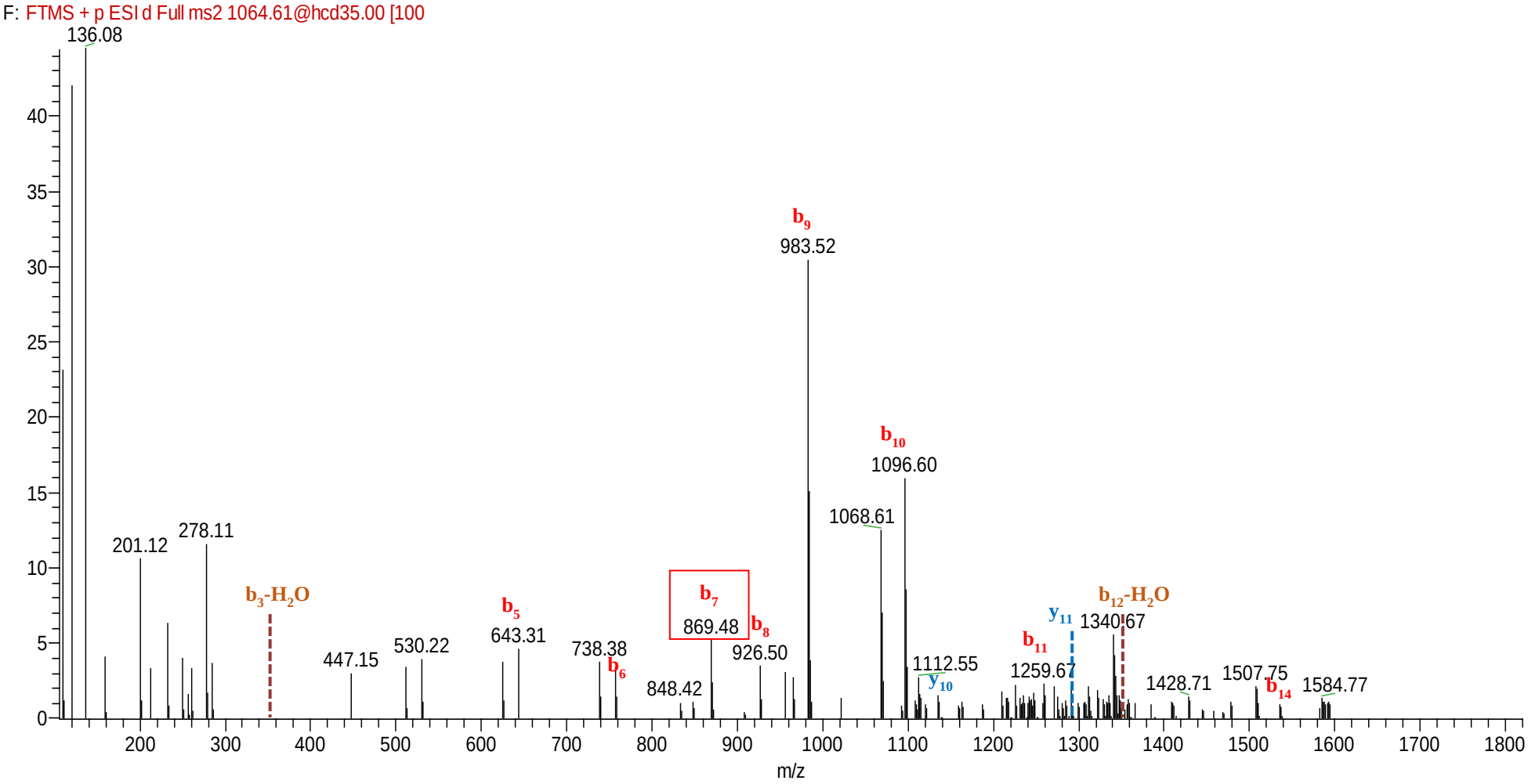


Figure 20a. Cystatin SN $P_{11\rightarrow}$L $Des_{1\text{-}4}$. Enlargement in the mass range 300-2000 *m/z* of the annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+13H]^{13+}$ 1064.61 *m/z*.



3.1.8 High-resolution top-down structural characterization of cystatins SA proteoform.

In the peak eluting between 32.5-33.5 minutes of the HPLC high-resolution ESI-MS profile shown in panel A of figure 21, we identified a new proteoform of cystatin SA whose high-resolution and deconvoluted mass spectra are shown respectively in panels B and C. The mass difference of -871.34 Da between cystatin SA (experimental monoisotopic ion $[M+H]^+$ at 14338.0 ± 0.2 *m/z*) and the experimental monoisotopic ion $[M+H]^+$ at 13466.7 ± 0.2 *m/z* could be attributed to a N-terminal cleavage of cystatin SA. The experimental monoisotopic $[M+H]^+$ value 13466.7 ± 0.2 *m/z* was recognised as the proteoform of cystatin SA lacking the first 7 amino acids from the N-terminus (cystatin SA Des_{1-7}) and carrying two disulfide bonds (theor. monoisotopic $[M+H]^+$ 13466.67 *m/z*). The sequence was confirmed by top-down high-resolution MS/MS experiments performed on the ion 1123.81 ± 0.02 *m/z* ($[M+12H]^{+12}$) as shown by the high-resolution MS/MS annotated spectra reported in figure 22 allowed confirming the sequence.

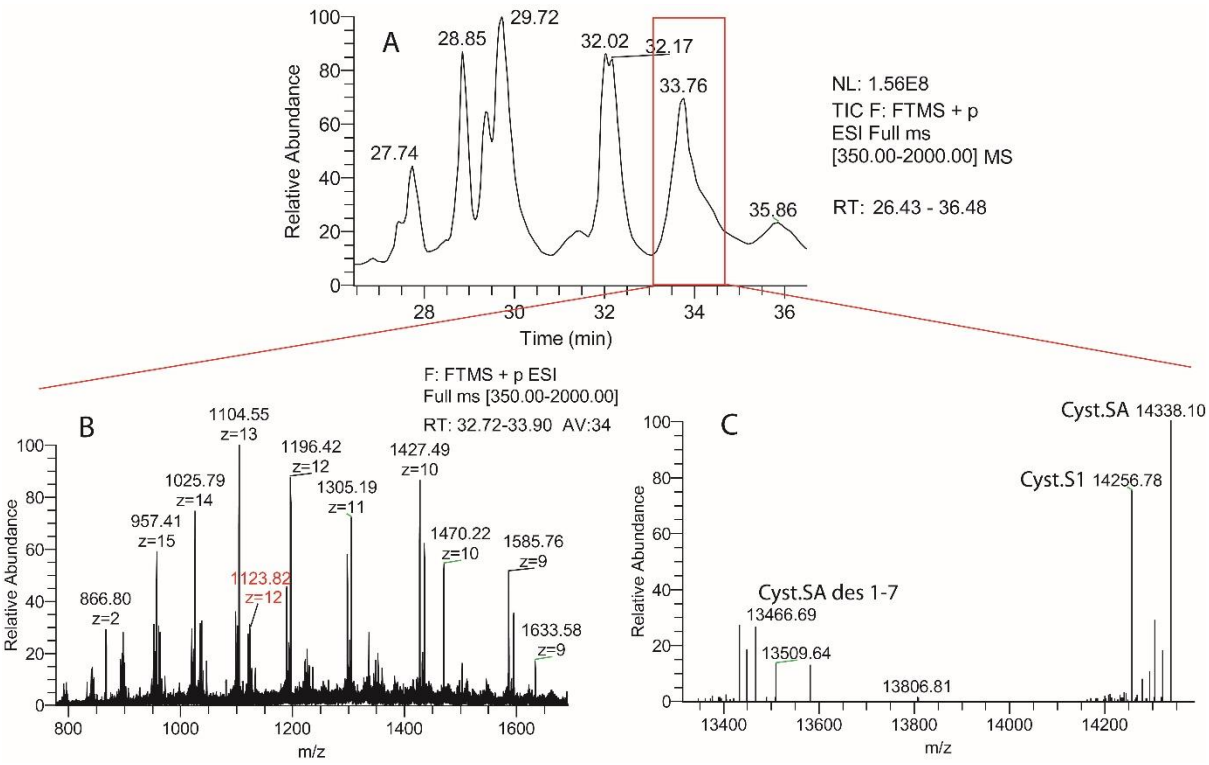
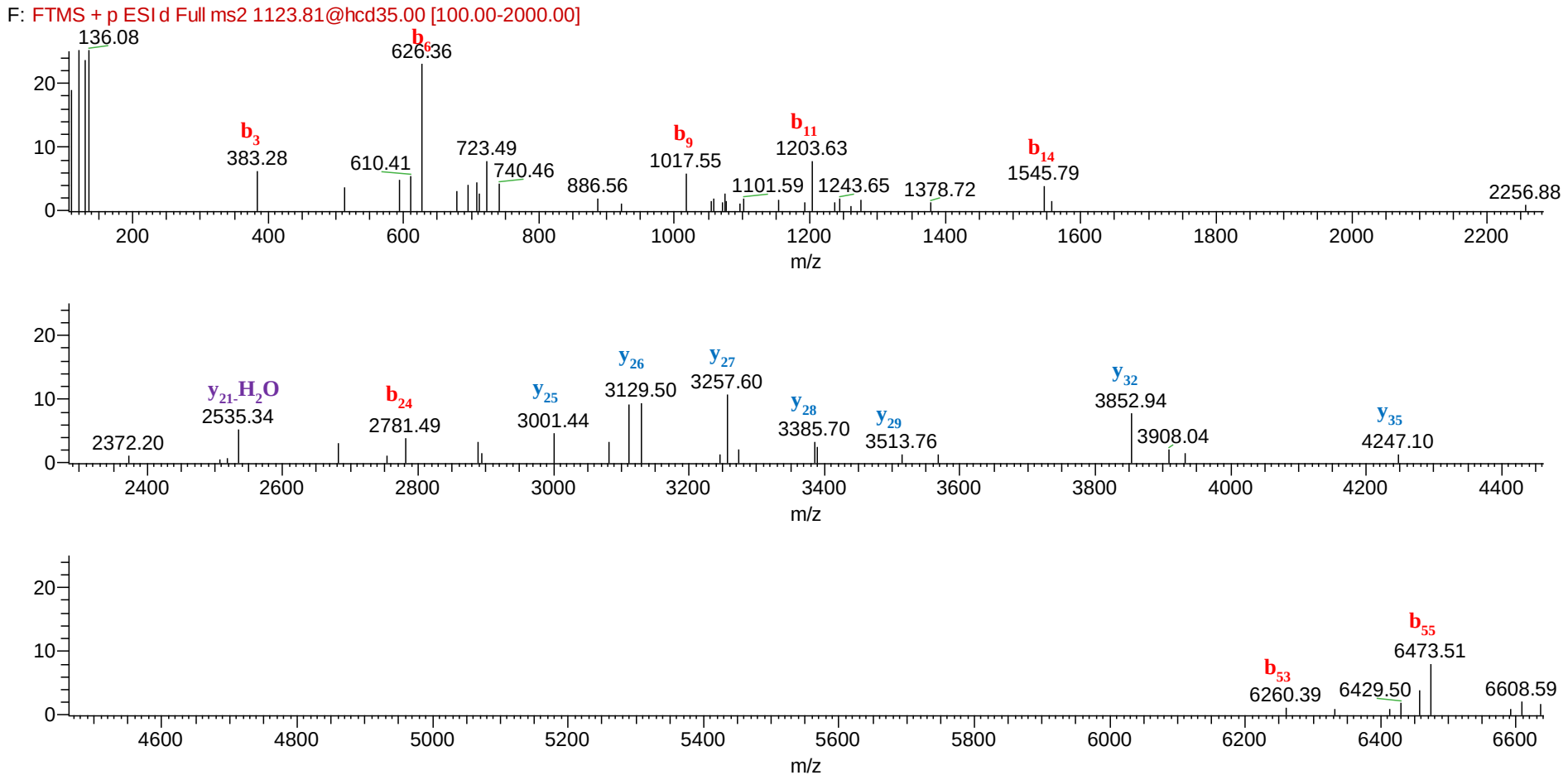


Figure 21. Characterization of cystatin SA proteoform. Enlargement, in the elution time 26.6-36.4 min (panel A), of the HPLC high-resolution ESI-MS profile shown in Figure 3, boxed peaks correspond to S-type cystatins. High-resolution mass spectra of cystatins SA (panel B) proteoforms with the corresponding deconvoluted spectra (panel C). Red *m/z* value in panel B have been used for high-resolution MS/MS characterization of the proteoform.

Figure 22. Cystatin SA Des$_{1\text{-}7}$. Annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+12H]^{12+}$ 1123.81 *m/z*.



3.1.9 High-resolution top-down structural characterization of oxidized proteoforms of cystatins S.

In the elution time 33-35 minutes of the HPLC low-resolution ESI-MS profile, two proteins with mass average 14272 ± 2 Da and 14288 ± 2 Da were detected, eluting slightly before cystatin S1 (Mav 14256 ± 2 Da). The mass difference of +16 and +32 Da with respect to cystatin S1 suggested that they could be mono-oxidized and di-oxidized forms of cystatin S1. In order to define the site of oxidation, the soluble fraction of saliva samples containing a high amount of the oxidized proteoforms, was submitted to top-down high-resolution MS/MS experiments using both HCD- and CID-based fragmentations. High resolution MS/MS experiments were performed on the mono-oxidized proteoforms of cystatin S1 (exp. monoisotopic ion $[M+H]^+$ at 14272.8 ± 0.02 *m/z*) in order to gain new insights into residues involved in oxidation. In particular the HCD MS/MS experiments performed on the ion 1299.26 ± 0.02 *m/z* ($[M+11H]^{+11}$) as shown by the high-resolution MS/MS annotated spectra reported in figure 23 and CID MS/MS experiments performed on the ion 1191.07 ± 0.02 *m/z* ($[M+12H]^{+12}$) as shown by the high-resolution MS/MS annotated spectra reported in figure 24 were consistent with tryptophan (W_{23}) as oxidized residue. Moreover, top-down high resolution MS/MS experiments performed with HCD- based fragmentation method on the ion 1192.32 ± 0.02 *m/z* ($[M+12H]^{+12}$) of the di-oxidized proteoforms of cystatin S1 (theor. monoisotopic $[M+H]^+$ 14288.76 *m/z*) were consistent with the presence of two isobaric proteoforms; one form di-oxidized at W_{23} as shown by the high-resolution MS/MS annotated spectra reported in figure 25 and the other form oxidized at W_{23} and W_{107} as shown by the high-resolution MS/MS annotated spectra reported in figure 26. Surprisingly, we were not able to obtain MS/MS data consistent with the oxidation of the Met_{111} of cystatin S.

Figure 23. **Mono-oxidized Cystatin S1-W_{23}ox**. HCD annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+12H]^{12+}$ 1299.36 *m/z* in agreement with W_{23}-oxidized (+15.995 Da)

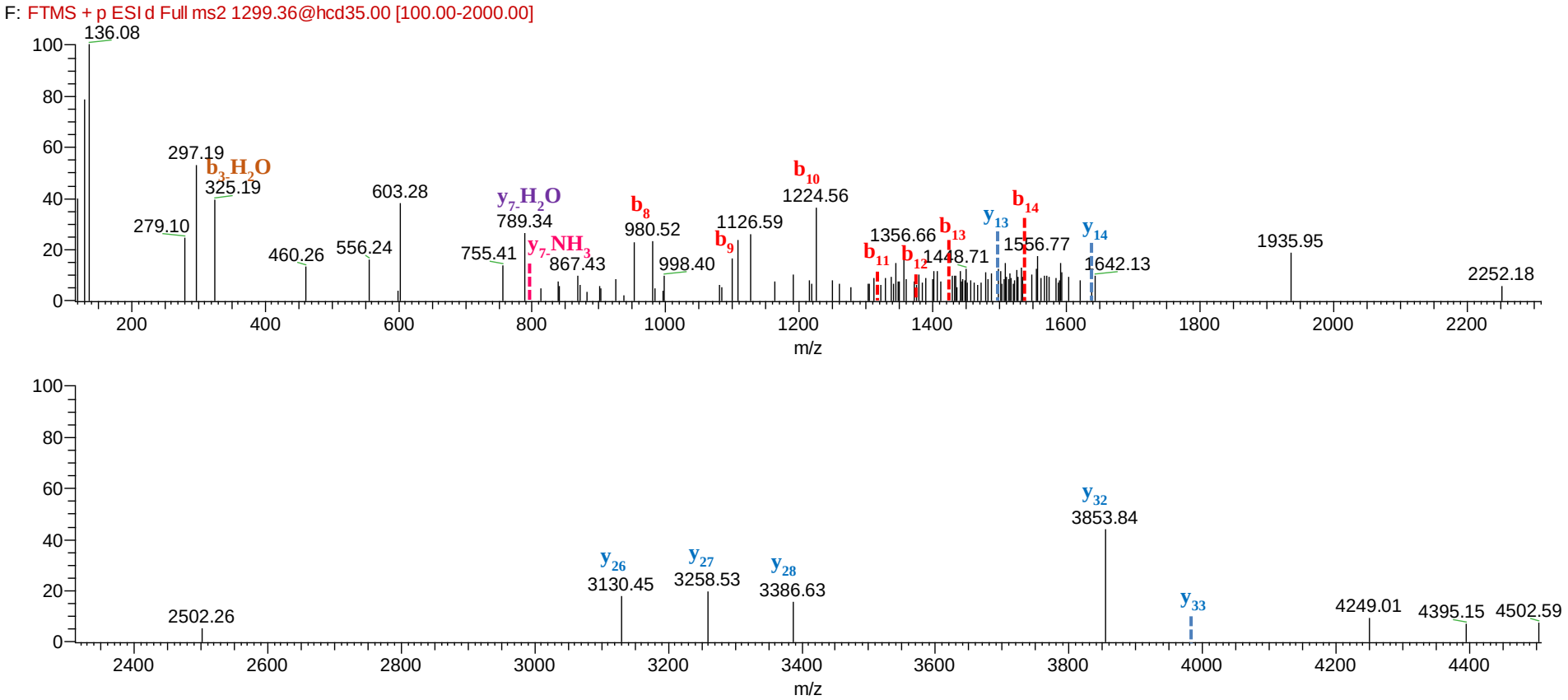


Figure 24. **Mono-oxidized Cystatin S1-W_{23}ox**. CID annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+12H]^{12+}$ 1191.07 *m/z* in agreement with W_{23}-oxidized (+15.995 Da)

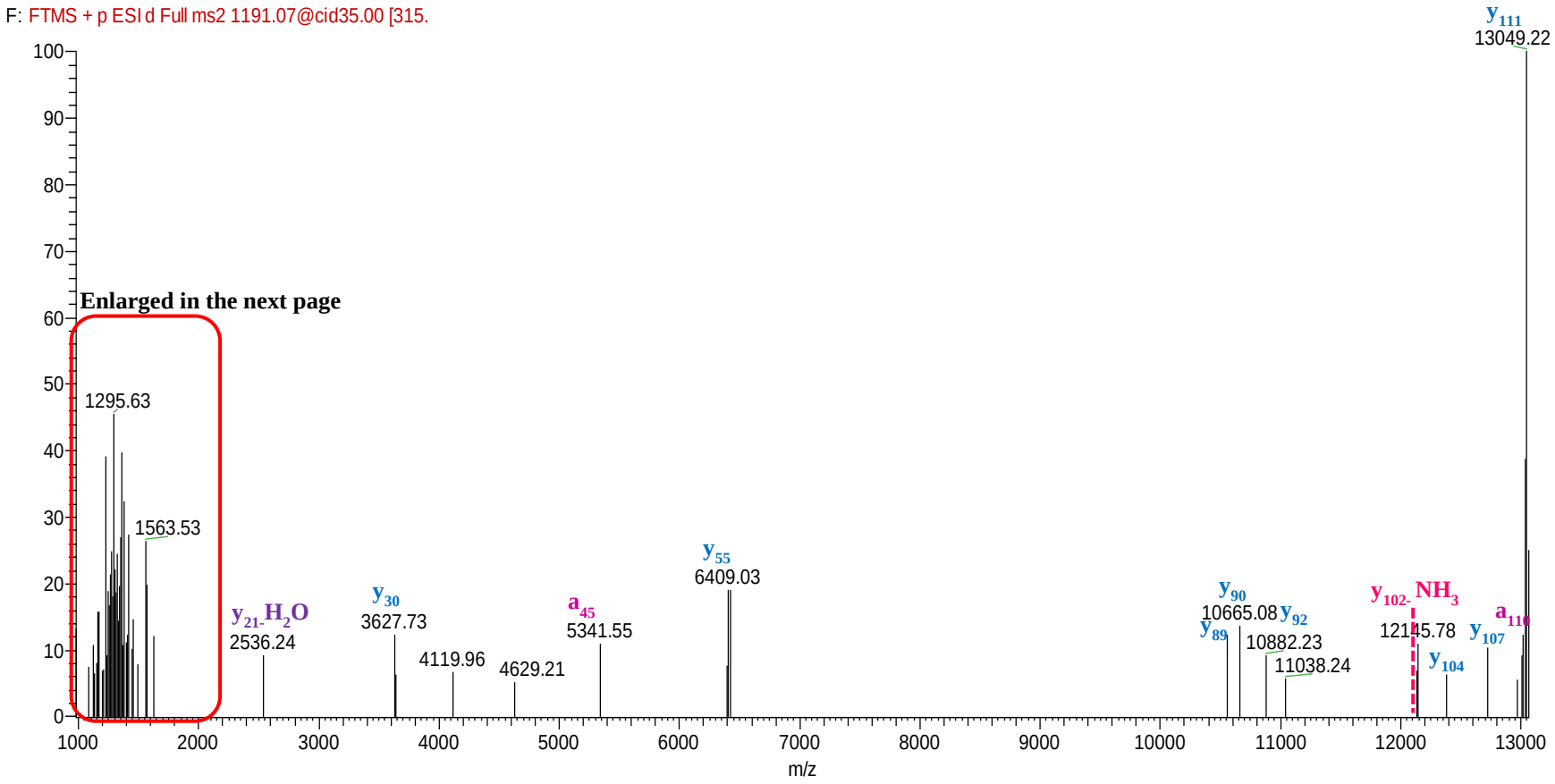


Figure 24a. Mono-oxidized Cystatin S1-W$_{23}$ox. Enlargement in the mass range 300-2000 *m/z* of the annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+12H]^{12+}$ 1191.07 *m/z*.

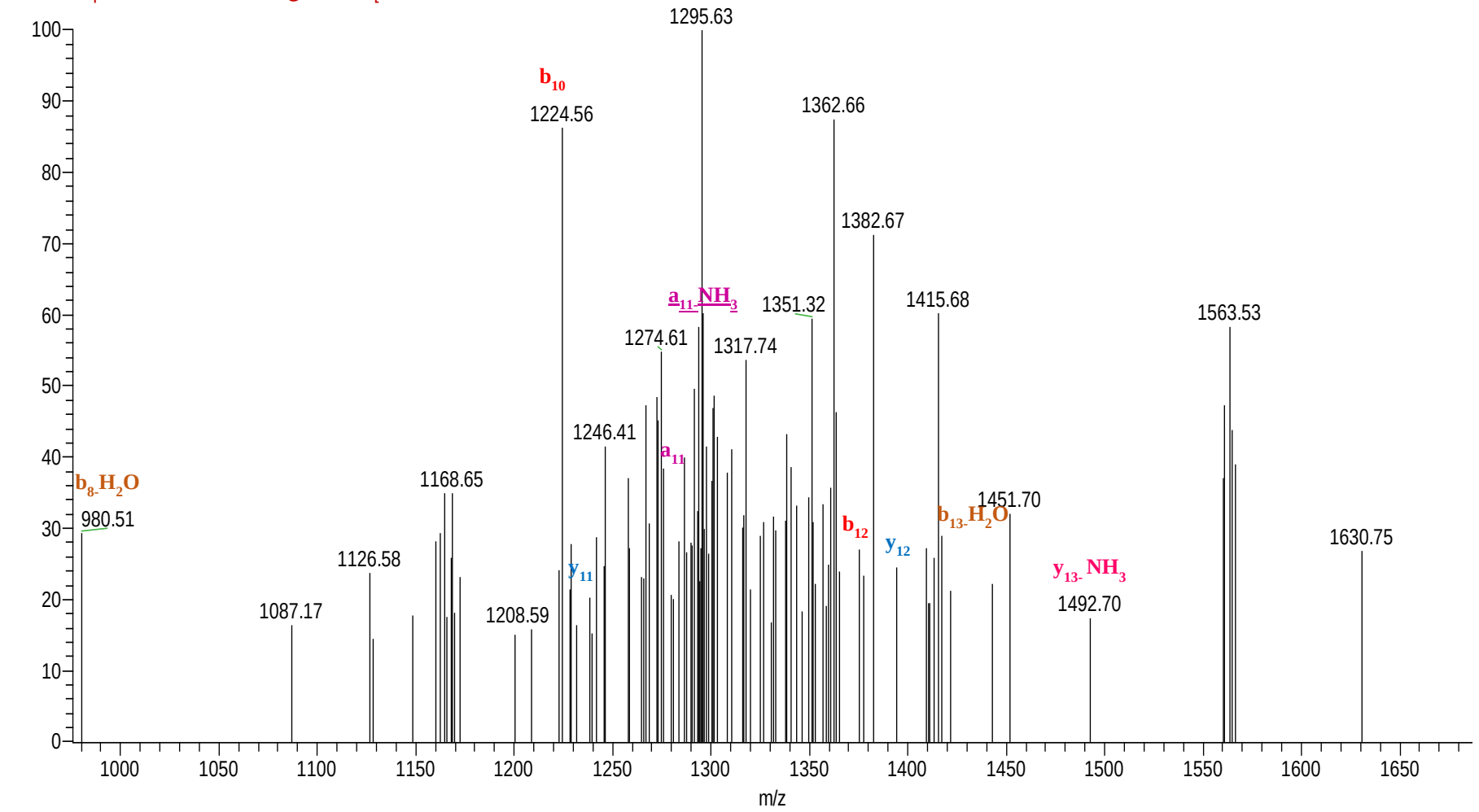


Figure 25. **Di-oxidized Cystatin S1-W_{23}di-ox**. HCD annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+12H]^{12+}$ 1192.32 *m/z* in agreement with W_{23} di-oxidized (+31.99 Da)

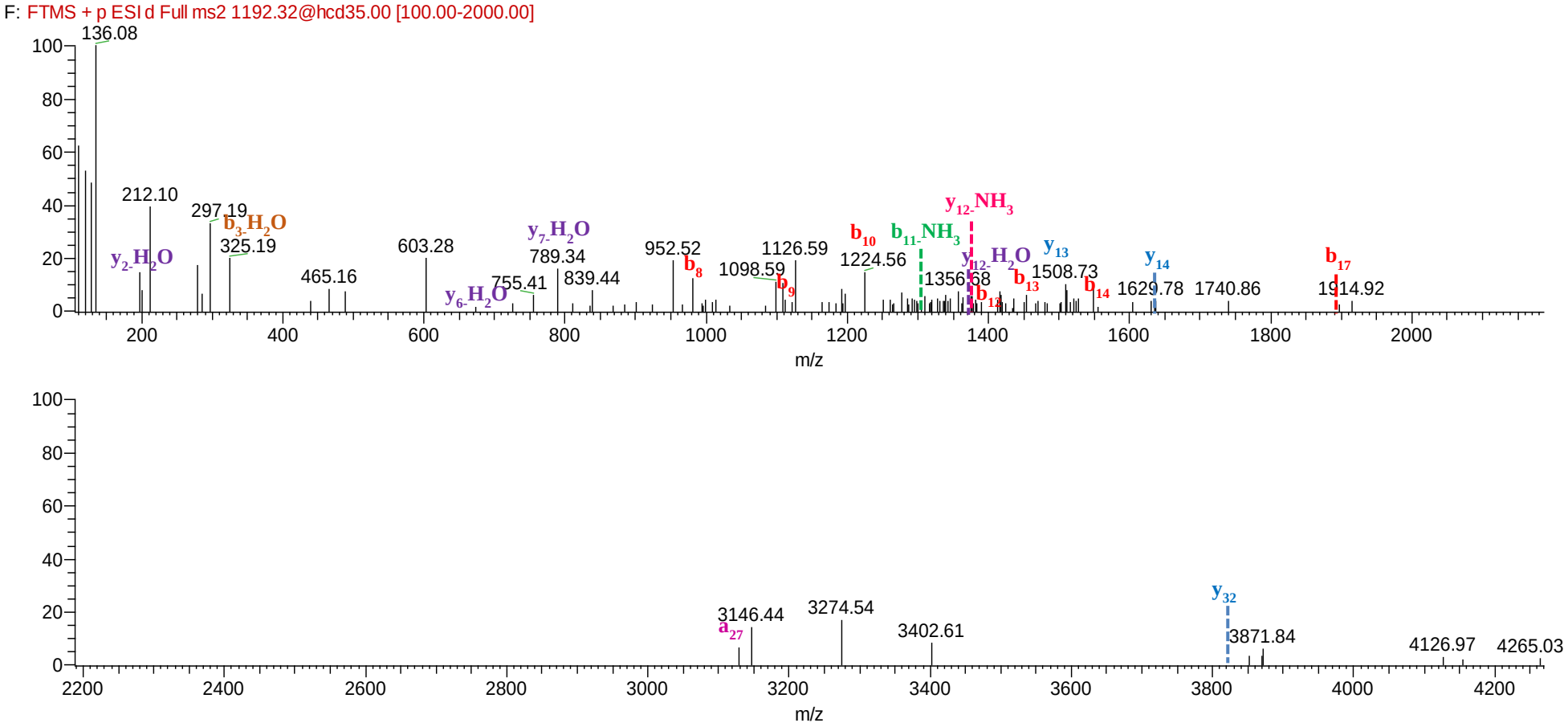
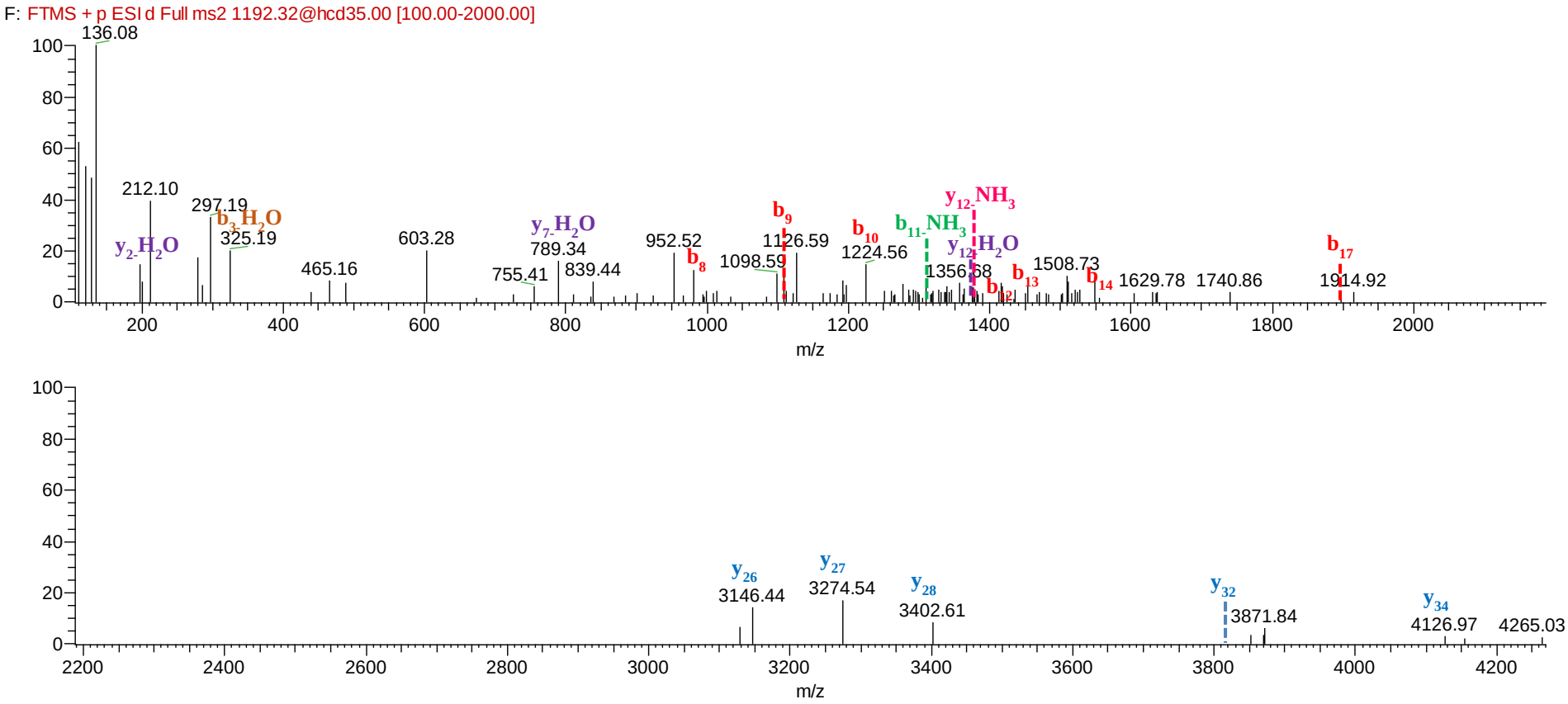


Figure 26. **Di-oxidized Cystatin S1-W_{23}ox W_{107}ox**. HCD annotated MH^+ deconvoluted spectra of high-resolution MS/MS of the ion $[M+12H]^{12+}$ 1192.32 *m/z* in agreement with W_{23} W_{107}-oxidized (+31.99 Da)



3.1.10 Quantitive analysis in Multiple Sclerosis

The quantitative analysis was performed by RP-HPLC low resolution ESI-MS and the relative abundance of the salivary proteins was determined by measuring the area of RP-HPLC-ESI-MS eXtracted Ion Current (XIC) peaks. This value is linearly proportional to the peptide concentration and it can be used to monitor relative abundances, under constant analytical conditions (Levin et al., 2007).

We analyzed the level of 102 peptides/proteins of which 68 secreted by salivary glands such as acidic proline-rich proteins, PC peptide, Db-s, Db-f, bPRP, statherins, histatins, S-type cystatins, cystatin C, cystatin D and P-B peptide and 34 not secreted from salivary glands (Castagnola et al., 2011a) such as cystatin A, cystatin B, β-thymosins 4 and 10, α-defensins, S100A7 (D_{27}), S100A8, S100A9 (long (L) and short (S) isoforms), S100A12, PIP, antileukoproteinase and pIgR fragments. We excluded from the analysis α-amylases and glycosylated bPRPs since the great heterogeneity of ESI-spectra did not allow their identification and quantification (Manconi et al., 2016b; Messana et al., 2004).

A comparison between patients and controls was performed by the software GraphPad Prism (version 6.0) for calculating means and standard deviations of peptides/proteins XIC peak areas and for statistical analysis.

Table 4 reports all peptides/proteins analyzed in this study divided by class, peptides/proteins name, XIC peak areas mean ± standard deviation (SD) ($x10^8$), frequency in MS patients and controls respectively and statistical significance expressed by *p* value for each protein/peptide.

Table 5 reports peptides/proteins found statistically variated between MS and control subjects; 14 were up-regulated in MS subjects with respect to control group and they were PC peptide fragments (PC Fr. 1-14, PC Fr. 24-44 and PC Fr. 36-44), cystatin A Thr_{96}→Met, cystatin SN Des_{1-4}, cystatin SN Pro_{11}→Leu, cystatin C, statherin 1P, statherin SV1, S100A7 (D_{27}), S100A8-SNO, antileukoproteinase and ASVD; while 6 were down-regulated in MS patients with respect to control group and they are cystatin S1 mono-ox, S1 di-ox, cystatin SN mono-ox, SN di-ox, cystatin SA and SA mono-ox as shown in figure 27.

Table 4. Proteins and peptides investigated, XIC peak areas mean ± standard deviation (SD) ($x10^8$), frequency in MS patients and controls respectively and statistical significance expressed by *p* value for each protein/peptide are reported.

Proteins/Peptides	MS		Controls		*p* value
	Mean ± SD	Frequency	Mean ± SD	Frequency	
aPRPs					
PRP-1 0P	0.74 ± 2.03	16/49	0.75 ± 1.60	21/54	ns
PRP-1 1P	11.22 ± 10.87	45/49	13.69 ± 10.50	54/54	ns
PRP-1 2P	90.07 ± 80.27	47/49	98.04 ± 73.87	54/54	ns
PRP-1 3P	3.15 ± 3.57	36/49	3.34 ± 3.69	48/54	ns
PRP-1 Tot	105.24 ± 91.92	47/49	115.80 ± 83.58	54/54	ns
PRP-3 0P	0.15 ± 0.33	13/49	0.15 ± 0.33	21/54	ns
PRP-3 1P	4.09 ± 3.74	44/49	5.11 ± 3.98	53/54	ns
PRP-3 2P	26.72 ± 28.23	46/49	31.17 ± 28.23	54/54	ns
PRP-3 3P	0.19 ± 0.48	8/49	0.17 ± 0.61	6/54	ns
PRP-3 Tot	31.16 ± 26.68	46/49	36.69 ± 31.75	54/54	ns
PRP-3 2P Des R_{106}	8.33 ± 12.26	47/49	5.92 ± 8.39	53/54	ns
P-C	18.55 ± 15.98	48/49	19.97 ± 14.68	54/54	ns
P-C Des P	0.35 ± 0.97	14/49	0.47 ± 0.74	20/54	ns
P-C Fr. 1-14	1.00 ± 1.03	47/49	0.52 ± 0.66	51/54	↑0.0005
P-C Fr. 1-25	2.00 ± 2.73	41/49	1.16 ± 1.07	49/54	ns
P-C Fr. 5-25	0.20 ± 0.40	29/49	0.12 ± 0.15	34/54	ns
P-C Fr. 15-44	0.74 ± 1.46	18/49	0.38 ± 0.74	14/54	ns
P-C Fr. 26-35	0.08 ± 0.21	15/49	0.06 ± 0.10	18/54	ns
P-C Fr. 26-44	1.18 ± 1.47	38/49	0.47 ± 0.62	37/54	↑0.004
P-C Fr. 36-44	0.19 ± 0.23	38/49	0.12 ± 0.23	27/54	↑0.02
P-C Fr. Tot	5.76 ± 6.46	49/49	3.30 ± 2.73	54/54	↑0.05
P-C Tot	24.31 ± 20.26	49/49	23.27 ± 15.27	54/54	ns
Pa 2-mer 4P	3.77 ± 10.48	10/49	5.71 ± 12.44	18/54	ns
Db-s 2P	7.89 ± 17.84	17/49	6.83 ± 14.06	18/54	ns
Db-s 1P	0.49 ± 1.50	10/49	0.29 ± 0.83	7/54	ns
Db-s Tot	8.56 ± 20.11	17/49	7.30 ± 15.03	18/54	ns
Db-f 2P	4.15 ± 9.52	16/49	4.01 ± 8.66	13/54	ns
Db-f 1P	0.40 ± 1.11	8/49	0.37 ± 0.86	10/54	ns
Db-f Tot	4.55 ± 10.50	16/49	4.38 ± 9.44	13/54	ns
bPRPs					
P-J	7.75 ± 9.64	41/49	8.37 ± 9.59	50/54	ns
IB-1	12.92 ± 16.35	36/49	14.86 ± 16.03	51/54	ns
P-F	5.67 ± 7.89	36/49	8.57 ± 10.83	47/54	ns
P-H	7.97 ± 9.18	43/49	9.77 ± 11.36	52/54	ns
Type 1 cystatins					
Cystatin A	3.34 ± 4.73	44/49	2.97 ± 3.35	52/54	ns

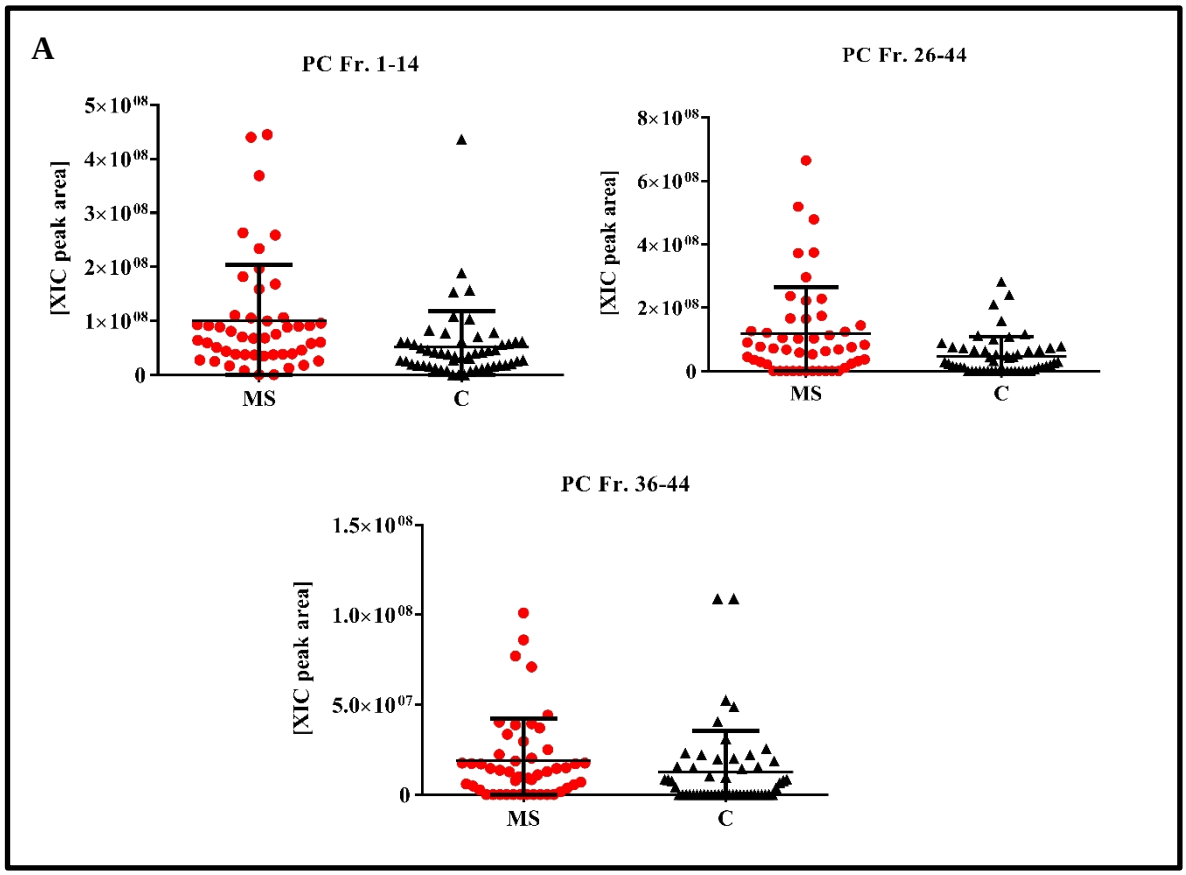
Proteins/Peptides	MS		Controls		*p* value
	Mean ± SD	Frequency	Mean ± SD	Frequency	
Cystatin A Acetyl	0.60 ± 0.84	31/49	0.81 ± 1.36	48/54	ns
Cystatin A Tot	3.94 ± 5.45	44/49	3.78 ± 4.15	52/54	ns
Cystatin A (T_{96}→M)	0.61 ± 1.45	14/49	0.12 ± 0.39	5/54	↑ 0.003
Cystatin B Acetyl S-CMC	0.44 ± 0.9	16/49	0.26 ± 0.46	10/54	ns
Cystatin B SSG	1.06 ± 0.97	46/49	0.83 ± 0.88	45/54	ns
Cystatin B S-cysteinyl	0.37 ± 0.53	28/49	0.20 ± 0.29	28/54	ns
Cystatin B S-S dimer	0.75 ± 1.28	23/49	0.50 ± 1.65	12/54	ns
Cystatin B Tot	2.62 ± 2.70	49/49	1.78 ± 2.58	47/54	↑ 0.004
Type 2 cystatins					
Cystatin S	0.86 ± 1.16	45/49	0.73 ± 0.82	47/54	ns
Cystatin S1	10.78 ± 13.51	47/49	10.66 ± 17.0	46/54	ns
Cystatin S1 mono-ox	0.14 ± 0.32	9/49	0.52 ± 0.73	30/54	↓ 0.002
Cystatin S1 di-ox	0.14 ± 0.47	6/49	0.39 ± 0.63	19/54	↓ 0.009
Cystatin S1 Tot	10.78 ± 13.51	48/49	10.66 ± 17.00	48/54	ns
Cystatin S2	3.72 ± 5.62	45/49	3.17 ± 5.34	44/54	ns
Cystatin S2 mono-ox	0.05 ± 0.30	2/49	0.08 ± 0.26	6/54	ns
Cystatin S2 di-ox	0.06 ± 0.38	2/49	0.03 ± 0.12	3/54	ns
Cystatin S2 Tot	3.82 ± 5.85	45/49	3.17 ± 5.35	44/54	ns
Cystatin SN	16.57 ± 22.27	44/49	24.03 ± 71.49	48/54	ns
Cystatin SN mono-ox	0.09 ± 0.22	8/49	0.45 ± 0.60	26/54	↓ 0.0003
Cystatin SN di-ox	0.02 ± 0.11	2/49	0.30 ± 0.56	16/54	↓ 0.0004
Cystatin SN Tot	16.68 ± 22.22	44/49	24.97 ± 71.53	48/54	ns
Cystatin SN Des_{1-4}	4.06 ± 10.93	22/49	0.83 ± 2.48	14/54	↑ 0.02
Cystatin SN (P_{11}→L)	2.00 ± 5.59	10/49	0.73 ± 3.51	3/54	↑ 0.03
Cystatin SN Des_{1-4} (P_{11}→L)	0.61 ± 2.79	5/49	0.13 ± 0.82	2/54	ns
Cystatin SA	2.57 ± 6.03	14/49	2.76 ± 4.97	29/54	↓ 0.03
Cystatin SA mono-ox	0.10 ± 0.32	6/49	0.36 ± 0.57	18/54	↓ 0.01
Cystatin SA Tot	2.67 ± 6.25	14/49	2.71 ± 4.01	29/54	↓ 0.04
Cystatin C	0.84 ± 2.24	18/49	0.19 ± 0.54	8/54	↑ 0.01
pGlu-cystatin DC_{26}→R Des_{1-5}	2.34 ± 10.0	16/49	0.60 ± 1.99	15/54	ns
Histatins					
Histatin 1 0P	0.17 ± 0.62	6/49	0.29 ± 0.76	13/54	ns
Histatin 1	2.36 ± 2.85	35/49	2.78 ± 2.89	38/54	ns
Histatin 3	1.16 ± 2.64	20/49	1.26 ± 1.80	27/54	ns
Histatin 5	3.20 ± 3.91	40/49	3.45 ± 3.29	47/54	ns
Histatin 6	1.14 ± 1.14	34/49	1.20 ± 1.25	38/54	ns
Histatin 3 Fr. 2-6	0.02 ± 0.04	13/49	0.005 ± 0.019	9/54	ns
Histatin 3 Fr. 1-11	0.08 ± 0.14	18/49	0.11 ± 0.14	30/54	ns
Histatin 3 Fr. 1-12	0.03 ± 0.07	9/49	0.04 ± 0.08	15/54	ns

Proteins/Peptides	MS		Controls		*p* value
	Mean ± SD	Frequency	Mean ± SD	Frequency	
Histatin 3 Fr. 1-13	0.02 ± 0.04	10/49	0.02 ± 0.04	13/54	ns
Histatin 3 Fr. 28-32	0.04 ± 0.13	17/49	0.02 ± 0.04	14/54	ns
Histatin Fr. Tot	4.53 ± 5.19	46/49	4.84 ± 4.52	51/54	ns
Statherins and PB					
Statherin 0P	0.01 ± 0.06	4/49	0.02 ± 0.05	8/54	ns
Statherin 1P	0.27 ± 0.37	29/49	0.45 ± 0.56	42/54	↑ 0.02
Statherin	11.55 ± 11.45	45/49	14.28 ± 11.93	49/54	ns
Statherin Tot	11.84 ± 11.74	45/49	14.75 ± 12.32	49/54	ns
Statherin Des_{1-9}	0.51 ± 0.60	35/49	0.59 ± 0.52	46/54	ns
Statherin Des_{1-10}	0.44 ± 0.45	37/49	0.46 ± 0.33	46/54	ns
Statherin Des_{1-13}	0.26 ± 0.30	36/49	0.29 ± 0.21	45/54	ns
Statherin SV1	4.69 ± 4.30	49/49	3.00 ± 3.68	52/54	↑ 0.03
Statherin Des T_{42} F_{43}	1.47 ± 1.50	48/49	1.89 ± 1.27	52/54	ns
Statherin Des D_1	0.94 ± 1.10	36/49	1.22 ± 1.68	48/54	ns
P-B	20.74 ± 21.57	49/49	21.63 ± 13.20	54/54	ns
P-B Des_{1-4}	2.44 ± 2.87	45/49	1.65 ± 1.49	44/54	ns
P-B Des_{1-5}	4.17 ± 7.59	49/49	3.34 ± 3.66	50/54	ns
P-B Des_{1-7}	6.19 ± 1.43	49/49	3.04 ± 1.39	54/54	ns
P-B Des_{1-12}	2.06 ± 2.56	43/49	1.43 ± 1.47	50/54	ns
P-B Fr. Tot	14.86 ± 18.71	49/49	9.46 ± 6.00	54/54	ns
P-B Tot	35.61 ± 32.32	49/49	31.09 ± 15.44	54/54	ns
α-defensins					
α-defensin 1	3.06 ± 2.74	48/49	2.55 ± 3.20	52/54	ns
α-defensin 2	2.17 ± 1.93	46/49	1.85 ± 2.28	52/54	ns
α-defensin 3	1.28 ± 1.55	36/49	0.98 ± 1.41	38/54	ns
α-defensin 4	0.62 ± 0.55	40/49	0.49 ± 0.59	33/54	ns
α-defensin Tot	7.13 ± 6.22	48/49	5.88 ± 7.05	52/54	ns
S100A					
S100A7 (D_{27})	0.39 ± 0.84	16/49	0.06 ± 0.02	6/54	↑ 0.01
S100A8	1.07 ± 4.61	13/49	1.13 ± 6.24	11/54	ns
S100A8-SSG	0.09 ± 0.34	5/49	0.006 ± 0.039	1/54	ns
S100A8-SNO	0.59 ± 1.56	13/49	0.13 ± 0.49	4/54	↑ 0.01
S100A8/A9-SS dimer	3.14 ± 8.14	10/49	1.84 ± 7.31	6/54	ns
Hyper-oxidized S100A8	0.28 ± 0.56	14/49	0.18 ± 0.41	13/54	ns
S100A8-SO3H/W54ox	0.19 ± 0.48	8/49	0.11 ± 0.30	8/54	ns
S100A8-SO_2H	0.14 ± 0.44	6/49	0.06 ± 0.25	5/54	ns
S100A8 Tot ox	4.44 ± 9.53	28/49	2.32 ± 7.24	25/54	ns
S100A8 Tot	5.51 ± 10.9	29/49	3.45 ± 12.98	27/54	ns
S100A9 short	2.52 ± 2.69	34/49	1.86 ± 3.01	28/54	ns

Proteins/Peptides	MS		Controls		*p* value
	Mean ± SD	Frequency	Mean ± SD	Frequency	
S100A9 short mono-ox	0.77 ± 1.01	23/49	1.22 ± 1.35	35/54	ns
S100A9 short+ox	3.30 ± 3.19	36/49	3.08 ± 4.00	40/54	ns
S100A9 short P	0.42 ± 0.92	12/49	0.57 ± 1.17	16/54	ns
S100A9 short P mono-ox	0.11 ± 0.27	8/49	0.18 ± 0.42	10/54	ns
S100A9 short P+ox	0.53 ± 1.10	12/49	0.75 ± 1.32	19/54	ns
S100A9 short Tot	3.83 ± 3.57	36/49	3.83 ± 5.02	40/54	ns
S100A9 long SSG	2.13 ± 3.06	24/49	1.56 ± 3.24	23/54	ns
S100A9 long SSG mono-ox	0.38 ± 0.75	15/49	0.46 ± 0.91	16/54	ns
S100A9 long SSG+ox	2.51 ± 3.34	29/49	2.01 ± 3.86	27/54	ns
S100A9 long SSG P+ox	0.33 ± 0.97	7/49	0.25 ± 0.93	7/54	ns
S100A9 long SSG Tot	2.84 ± 3.73	31/49	2.25 ± 4.67	27/54	ns
S100A9 long cyst Tot	0.26 ± 0.86	8/49	0.26 ± .12	6/54	ns
S100A12	0.81 ± 1.68	16/49	0.97 ± 2.40	11/54	ns
Antileukoproteinase					
Antileukoproteinase	0.35 ± 0.53	19/49	0.10 ± 0.19	16/54	↑ 0.004
Thymosin					
Thymosin β4	2.15 ± 2.59	37/49	1.71 ± 2.25	37/54	ns
pIgR Fragments					
AVAD	0.27 ± 0.25	36/49	0.26 ± 0.20	41/54	ns
ASVD	0.31 ± 0.30	45/49	0.17 ± 0.14	38/54	↑ 0.004

Table 5. Proteins and peptides investigated, XIC peak areas mean ± standard deviation (SD) ($x10^8$), frequency and statistically variated between MS and control subjects are reported.

Proteins/Peptides	MS		Controls		*p* value
	Mean ± SD	Frequency	Mean ± SD	Frequency	
aPRPs					
P-C Fr. 1-14	1.00 ± 1.03	47/49	0.52 ± 0.66	51/54	↑ 0.0005
P-C Fr. 26-44	1.18 ± 1.47	38/49	0.47 ± 0.62	37/54	↑ 0.004
P-C Fr. 36-44	0.19 ± 0.23	38/49	0.12 ± 0.23	27/54	↑ 0.02
P-C Fr. Tot	5.76 ± 6.46	49/49	3.30 ± 2.73	54/54	↑ 0.05
Type 1 cystatins					
Cystatin A (T_{96}→M)	0.61 ± 1.45	14/49	0.12 ± 0.39	5/54	↑ 0.003
Cystatin B Tot	2.62 ± 2.70	49/49	1.78 ± 2.58	47/54	↑ 0.004
Type 2 cystatins					
Cystatin S1 mono-ox	0.14 ± 0.32	9/49	0.52 ± 0.73	30/54	↓ 0.002
Cystatin S1 di-ox	0.14 ± 0.47	6/49	0.39 ± 0.63	19/54	↓ 0.009
Cystatin SN mono-ox	0.09 ± 0.22	8/49	0.45 ± 0.60	26/54	↓ 0.0003
Cystatin SN di-ox	0.02 ± 0.11	2/49	0.30 ± 0.56	16/54	↓ 0.0004
Cystatin SN Des_{1-4}	4.06 ± 10.93	22/49	0.83 ± 2.48	14/54	↑ 0.02
Cystatin SN (P_{11}→L)	2.00 ± 5.59	10/49	0.73 ± 3.51	3/54	↑ 0.03
Cystatin SA	2.57 ± 6.03	14/49	2.76 ± 4.97	29/54	↓ 0.03
Cystatin SA mono-ox	0.10 ± 0.32	6/49	0.36 ± 0.57	18/54	↓ 0.01
Cystatin C	0.84 ± 2.24	18/49	0.19 ± 0.54	8/54	↑ 0.01
Statherins and PB					
Statherin 1P	0.27 ± 0.37	29/49	0.45 ± 0.56	42/54	↑ 0.02
Statherin SV1	4.69 ± 4.30	49/49	3.00 ± 3.68	52/54	↑ 0.03
S100A					
S100A7 (D_{27})	0.39 ± 0.84	16/49	0.06 ± 0.02	6/54	↑ 0.01
S100A8-SNO	0.59 ± 1.56	13/49	0.13 ± 0.49	4/54	↑ 0.01
Antileukoproteinase					
Antileukoproteinase	0.35 ± 0.53	19/49	0.10 ± 0.19	16/54	↑ 0.004
pIgR					
ASVD	0.31 ± 0.30	45/49	0.17 ± 0.14	38/54	↑0.004



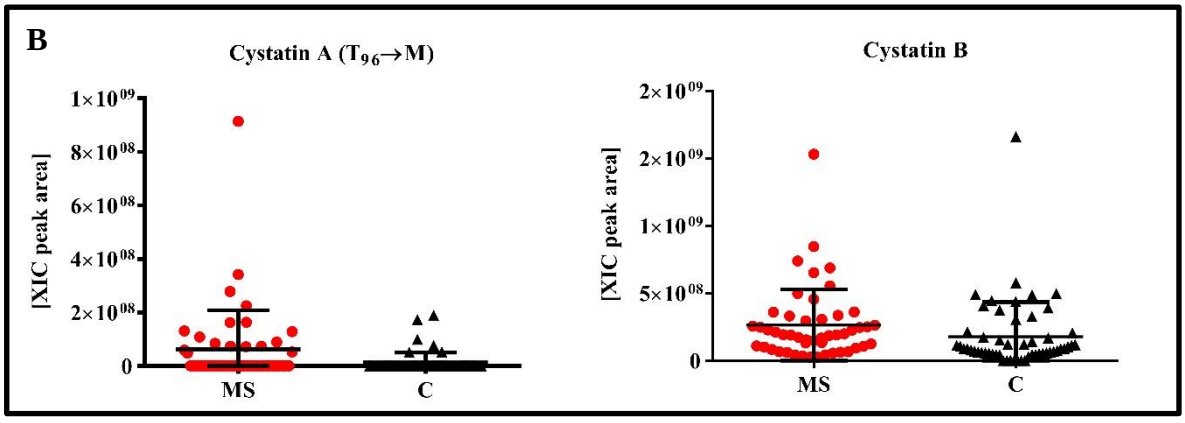
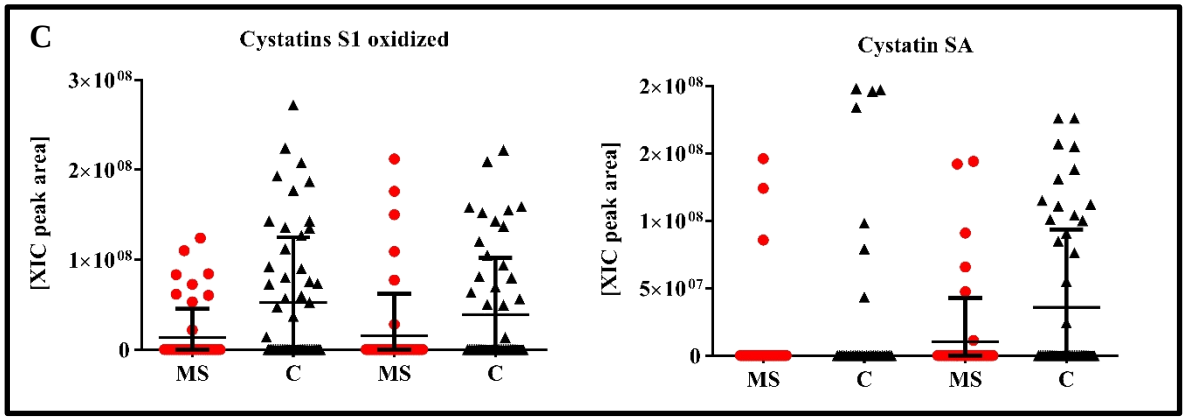
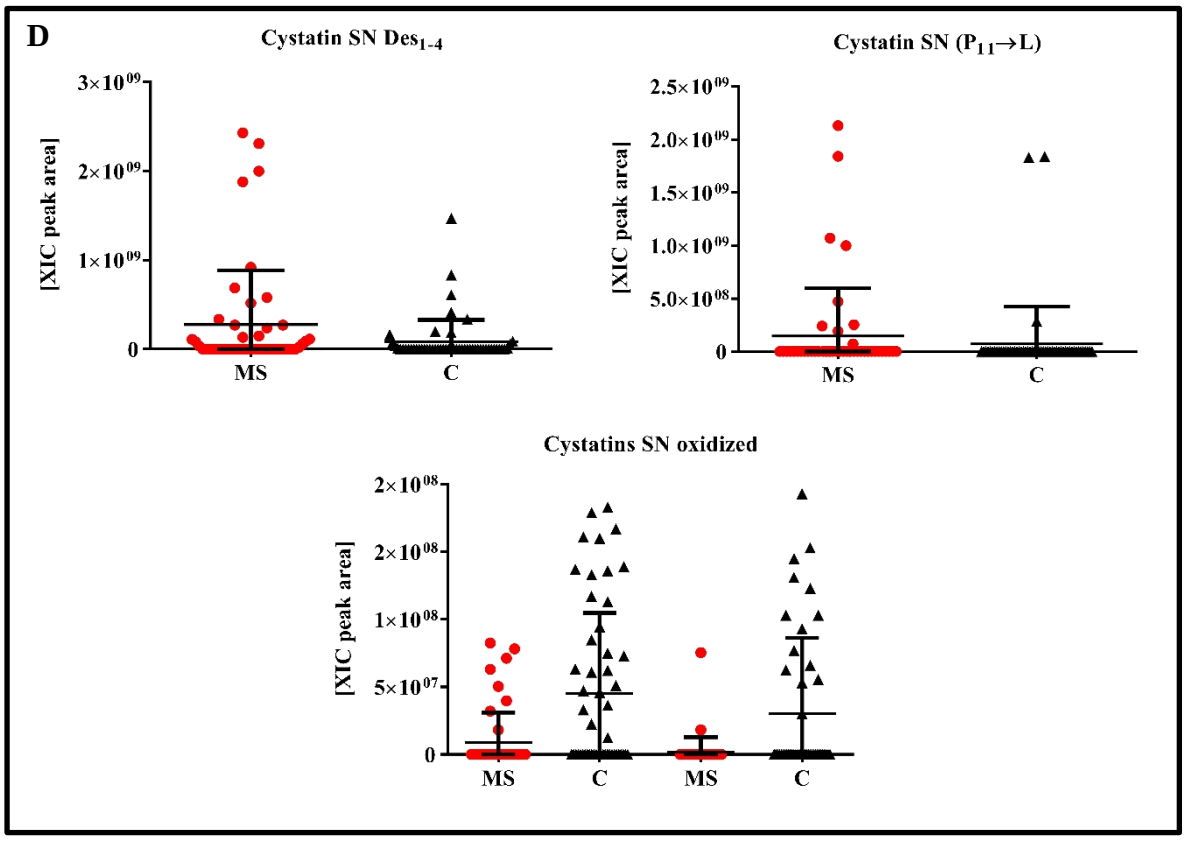
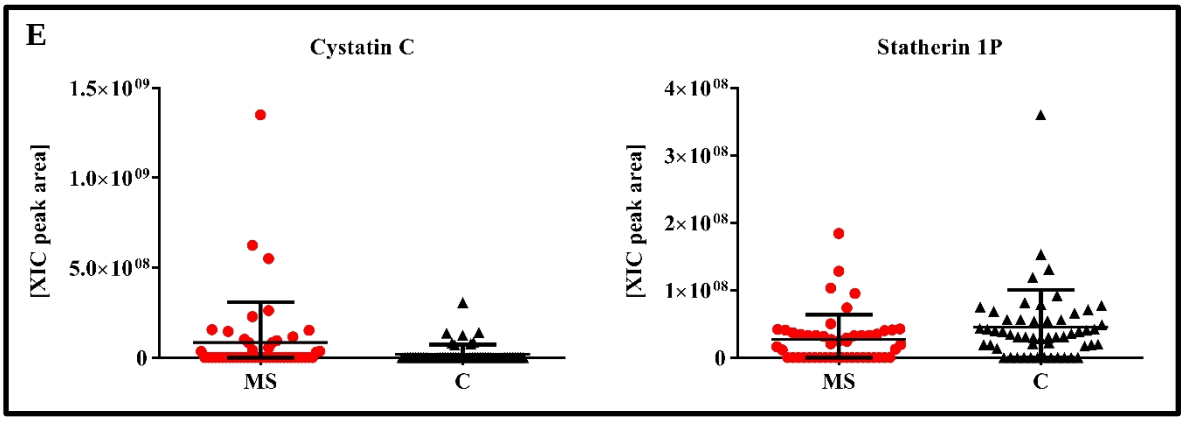


Figure 27. Distribution of the XIC peak area values measured in saliva and statistically variated from MS and C subjects of (A) PC fragments; (B) type 1 cystatins; (C) cystatins S1 oxidized proteoforms and cystatin SA; (D) cystatins SN proteoforms; (E) cystatin C and statherin 1P; (F) S100A7, S100A8-SNO, antileukoproteinase and ASVD. (Continue in the nexts pages).


C
Cystatins S1 oxidized
[XIC peak area]
3×10^08
2×10^08
1×10^08
0
MS
C
MS
C
Cystatin SA
[XIC peak area]
2×10^08
2×10^08
1×10^08
5×10^07
0
MS
C
MS
C


D
Cystatin SN Des1-4
[XIC peak area]
3×10^09
2×10^09
1×10^09
0
MS
C
Cystatin SN (P11→L)
[XIC peak area]
2.5×10^09
2.0×10^09
1.5×10^09
1.0×10^09
5.0×10^08
0
MS
C
Cystatins SN oxidized
[XIC peak area]
2×10^08
2×10^08
1×10^08
5×10^07
0
MS
C
MS
C


E
Cystatin C
[XIC peak area]
1.5×10^09
1.0×10^09
5.0×10^08
0
MS
C
Statherin 1P
[XIC peak area]
4×10^08
3×10^08
2×10^08
1×10^08
0
MS
C

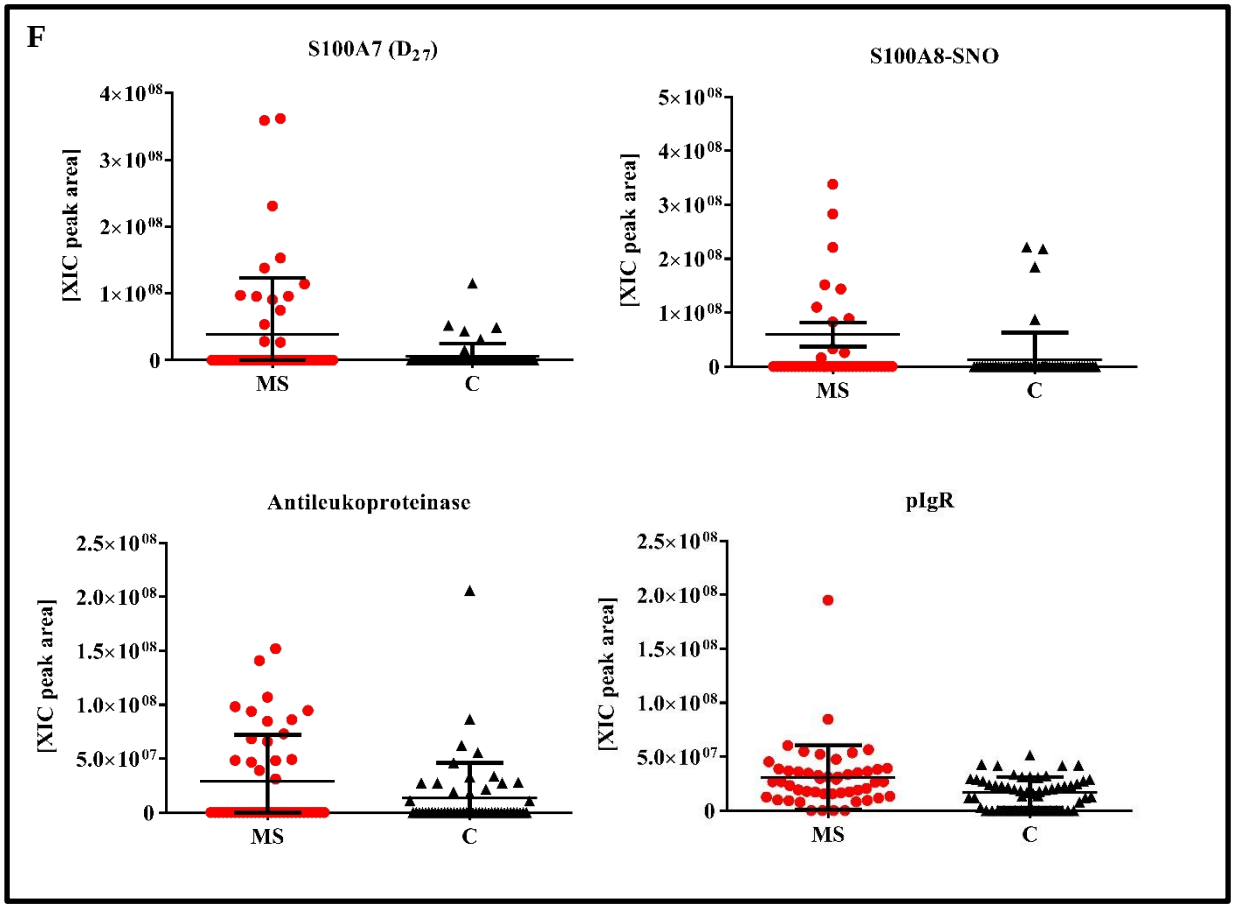


Figure 27. Distribution of the XIC peak area values measured in saliva and statistically variated from MS and C subjects of (A) PC fragments; (B) type 1 cystatins; (C) cystatins S1 oxidized proteoforms and cystatin SA; (D) cystatins SN proteoforms; (E) cystatin C and statherin 1P; (F) S100A7, S100A8-SNO, antileukoproteinase and pIgR.

3.1.11 Discussion

The comparative analysis of salivary proteome in MS patients with respect to controls allowed the identification and the structural characterization of new proteoforms of salivary cystatins never detected before in saliva. Moreover, this study highlighted quantitative alterations at the level of different peptides and proteins of specific glands secretion as well as some proteins non-specifically detectable in oral cavity.

The proteoforms detected and characterized in saliva for the first time during this study were cystatin A Thr_{96}→Met and its acetylated derivative; cystatin B N-terminally acetylated and CMC at Cys_3; N-terminally truncated cystatin D with the N-terminal Q converted to pyro-E and lacking the first 5 amino acid residues (pGlu-cystatin D Cys_{26}→R Des_{1-5}); N-terminally truncated forms of cystatin SN and SN P_{11}→L lacking the first 4 amino acids (cystatin SN Des_{1-4} and cystatin SN P_{11}→L Des_{1-4}) and the first 7 amino acids (cystatin SN Des_{1-7} and cystatin SN P_{11}→L Des_{1-7}); N-terminally truncated cystatin SA lacking the first 7 amino acids (cystatin SA Des_{1-7}); oxidized derivatives of cystatins SN and S1 at W_{23} and W_{107}.

The quantitative analysis performed on 102 salivary peptides/proteins, showed a high number of statistically variated proteins belonging to cystatins family. Among these, cystatin A Thr_{96}→Met, cystatin SN Des_{1-4} and SN and P_{11}→L; oxidized derivatives of cystatins SN and S1 were also found with altered level in MS with respect to controls group. Moreover, a higher number of protein statistically variated were found among those not specifically secreted from salivary glands such as S100A7 (D_{27}), S100A8-SNO, antileukoproteinase and ASVD.

Cystatins are a natural inhibitor of cysteine proteases and show inhibitory activity against human lysosomal cathepsins.

Cystatin A is a proteases inhibitor mainly involved in cellular proliferation and increased level of transcripts have been found in involved and uninvolved skin of psoriatic patients (Vasilopoulos et al., 2007), and in noncancerous tissue surrounding hepatocellular carcinoma cells (Lin et al., 2016). Polymorphisms in the gene for cystatin A have been associated with different inflammatory skin diseases. In particular the nonsense mutations causing the protein truncation (p.Lys22stop, p.Gln86stop) and the splice site mutation (p.Val23_Gln26del) were found responsible of the congenital exfoliative ichthyosis (Blaydon et al., 2011), and acral peeling skin syndrome (Muttardi et al., 2016), two skin diseases often associated with a defective epidermal barrier.

The subjects we analyzed, with the cystatin A T_{96}→M variant, were heterozygous since presented also the wild-type form of cystatin A and did not present any clinical evidence of skin disease. Even if the altered threonine at position 96 is a variable residue located at the protein C-terminus, not directly involved in the binding of cystatin A to potential cathepsin substrates (Renko et al., 2010), SIFT analysis predicted a deleterious effect of the T_{96}→M substitution on protein function (score 0.00).

It is interesting to note that the natural variant of cystatin A T_{96}→M is more frequent and with higher levels in MS subjects than the controls. Obviously, having been characterized for the first time in the present study, there is not previous bibliography that associates this variant with MS, but the group of Nobles et al. (Noben et al., 2006) was able to highlight an increased levels of cystatin A in the cerebral spinal fluid and plasma from patients affected by MS. It's important to note that they did not characterize cystatin A, therefore it can not be excluded that it is the same increase observed by us with the natural variant T_{96}→M.

Cystatin B plays an important role in innate immunity, mainly as a defence against bacterial infections (Zavasnik-Bergant, 2008).

The N-terminal region of cystatin B plays an important role in binding cysteine proteinases. Among the N-terminal residues, C_3 is the most important one for the interaction of the N-terminal region with papain and cathepsin H, and is also a major contributor to cathepsin L binding (Pavlova et al., 2003). C_3 of cystatin B in adult human whole saliva has been mostly found S-modified, while the S-unmodified derivatives were generally not detected (Cabras et al., 2012b), probably because the reactive C_3 is involved in binding tightly the target proteinases. Indeed, Turk et al. evidenced that dimer formation due to C_3 oxidation would lead to inactivation of cystatin B due to steric hindrance, and glutathionylation or cysteinylation of C_3 could behave in the same way (Turk et al., 1986).

We characterized for the first time a new S-modified derivative of cystatin B, namely S-(carboxymethyl)-cysteine (CMC), being the carboxymethylation a novel PTM not only for this protein but also for any human salivary protein. The high level of cystatin B observed in MS patients could suggest the creation of a self-defense mechanism against typical MS inflammation. On this regard, the presence of S-CMC derivatives of cystatin B and of the other salivary proteins prone to cysteine oxidation in whole saliva could represent a new biomarker of oral or systemic oxidative stress.

The natural variant of cystatin SN P_{11}→L has been previously detected in human saliva and the modified residue defined by top-down Fourier-transform Ion Cyclotron Resonance Mass Spectrometry (Whitelegge et al., 2007), but the truncated proteoforms of this natural variant (SN P_{11}→L Des_{1-4}, SN P_{11}→L Des_{1-7}), as well as those of the widespread cystatin SN (SN Des_{1-4}, SN Des_{1-7}), have been detected and characterized for the first time in this study.

The N-terminal sequence of full-length S-type cystatins is characterized by 11 amino acid variable residues beyond the well-conserved glycine at position 12 that, together with the QXVXG and PW motifs, forms a wedge-like structure in the cystatins molecules involved in the binding with papain-like enzymes (Stubbs et al., 1990). Nevertheless, S-type cystatins purified from human saliva have been found frequently truncated at their N-terminus (Saitoh et al., 1993) as well as human cystatin C, D (Popović et al., 1990) and N-terminally processed cystatins S have been observed also in rats (Nishiura et al., 1991). The N-terminal truncation drastically affects the inhibitory activity of cystatin C (Mason et al., 1998), since G_{12} has been shown to be responsible for the flexibility of the N-terminal region improving the binding with cysteine proteases (Bode et al., 1988). As the S-type cystatins it concerns, truncated proteoforms usually showed only modest differences in inhibitory activity against papain when compared with the full-length counterparts (Bobek et al., 1994), even if the proteoform of cystatin SA lacking the six N-terminal residues was found to be 1000-fold less active toward cathepsin L than the full-length cystatin proteoform (Baron et al., 1999). The level of N-terminal truncation of salivary cystatin SN (cystatin SN Des_{1-3}), connected to the activity of the involved proteinase, has been suggested as a biomarker for the pathology in oral squamous cell (Shintani et al., 2010). We observed an increase levels in the concentration of cystatin SN P_{11}→L and SN Des_{1-4} in subjects with MS compared to controls. In general, increased salivary cystatin concentration, which is also involved in oral cavity flogistic processes, may be the result of defense mechanisms designed to protect the mucous membranes from inflammatory damage which, in subjects with MS is typically observed in the SNC. There are no known natural variants of cystatin SN implicated in MS pathogenesis, so the significance of the increased level of cystatin SN P_{11}→L observed remains to be illustrated. The increase levels of cystatin SN Des_{1-4} observed in MS subjects may reflect an alteration in the oral micro-environment, such as fluctuations bacterial and proteolytic activity associated with it.

Unexpectedly, we observed reduced oxidation levels for cystatin S1, SN and SA, in patients with MS than in controls. Many studies report an increase in oxidative stress displaced in different districts and biological fluids, we do not know why there is this difference but the inter-individual variability of MS therapy effects may be present in different subjective or objective clinical response and different changes in expression of cytokines, chemokines, oxidative stress markers and other molecules. The effect of MS therapy on peripheral oxidative stress markers is highly complex due to strong inter-individual variability. Analysis of subgroups of patients is not possible due to the low number of patients included. Moreover, patients did not undergo a dental examination, and oral disease could also affect the oxidative stress markers (Karlìk et al., 2015).

Cystatin C expression level reductions have been described in the CNS in a very high number of neurological pathologies and in animal models with neurodegenerative diseases, underlining the prominent role of cystatin C in these conditions. However, the research conducted so far, in order to define the role of cystatin C in MS pathogenesis, has yielded contrasting results. Numerous studies have shown a decrease in cystatin C concentration in the cerebrospinal fluid of patients with MS (Nagai et al., 2000; Hansson et al., 2007), while recent studies suggest l absence of correlation between MS and serum cystatin C levels, highlighting how immunomodulatory and immunosuppressive therapies often used to control pathology induce a reduction in cathepsins concentration and a corresponding increase of cystatin C (Haves-Zburof et al., 2011). Similarly to what is described for plasma, the increase of the cystatin C levels that we observed in MS patients may be due to the pharmacological therapies currently in use at the time of collection of the samples and in this respect we will propose in future to evaluate the levels of this protein in different classes of MS subjects subdivided according to therapy. Although not changed in MS this protein will be used for future studies as well is a multifunctional protein and shows several activities not related to the antiproteinase activity, such as inhibition of proliferation, migration and invasion of colon carcinoma cells (Alvarez-Díaz et al., 2009), regulation of antigen presenting cells activity (Nashida et al., 2013) and modulation of gene expression related to its nuclear activity localization (Ferrer-Mayorga et al., 2015).

The cystatin D C_{26}→R lacking first 5 amino acid residues with the N-terminal glutamine converted to pyroglutamic acid (pGlu-cystatin D C_{26}→R Des_{1-5}) has been evidenced and characterized for the first time in the present study. While the N-terminal truncation on cystatin D C_{26}→R has been earlier observed in human saliva (cystatin D C_{26}→R Des_{1-4}

and cystatin D C_{26}→R Des_{1-8}) and characterized by top-down Fourier-transform ion cyclotron resonance mass-spectrometry approach (Ryan et al., 2010), the pyroglutamination has never been observed for salivary cystatins. The formation of pGlu can occur spontaneously or in the presence of glutamine cyclase, which can act on either N-terminal glutamine or glutamate. To date the presence of pGlu, among the salivary proteins, has been observed in acidic-PRP isoforms (Castagnola et al., 2003), α-amylases (Peng et al., 2012) and basic-PRPs IB1, IB4, PB, II2 (Kauffman et al., 1986). The pGlu moiety provides proteins resistance from degradation by amino peptidases and has a significant role in the functionality of peptides and proteins (Wirths et al., 2009). In fact, intact cystatin D is not commonly detectable and the pGlu-cystatin D C_{26}→R Des_{1-5} is by far the most abundant truncated proteoform in adult human saliva.

MS patients were characterized by high levels of some fragments of aPRP (PC Fr. 1-14, PC Fr. 24-44 and PC Fr. 36-44) and statherin (statherin 1P and SV_1) that are tipically secreted by salivary glands. It is recognized that these proteins play an important role in the creation of a protective environment for the teeth, in the modulation of the bacteria adhesion to the oral surfaces. Moreover, they are involved in the formation of the protein pellicles covering the oral surfaces (Delius et al., 2017). Therefore, an increased secretion of these peptides might be useful in the oral cavity of MS patients to protect the oral mucosa from damages caused by the inflammatory response. In fact, it's been report that the capacity of commensals to calibrate systemic immunity has profound consequences in the context of immunotherapy. For example, total body irradiation, used in defined settings of immunotherapy and bone marrow transplantation, is associated with gut damage and microbial translocation, providing an adjuvant effect to the transferred anti-tumoral T cells (Paulos et al., 2007).

In patients with MS, an increase levels in the concentration of two proteins belonging to the S100 family, S100A7 and S100A8-SNO (S-nitrosilate proteoform) were observed. The S100A7 protein has two EF-hand calcium-binding motifs (Watson et al., 1998), locates in the cytoplasm or nucleus of different cell types and, in addition to being involved in cell cycle regulation and cell differentiation, exhibits antimicrobial activity and immunomodulatory (Gläser et al., 2005). Abnormalities in expression levels of S100A7 were observed in the presence of imbalance of inflammatory and immune response (Mandal et al., 2007; West & Watson, 2010; Batycka-Baran et al., 2015). In particular, an increase of S100A7 levels were observed in subjects with systemic sclerosis, an autoimmune disease characterized by progressive fibrosis of the skin and

internal organs (Giusti et al., 2016). Similarly, it could be hypothesized that increased S100A7 protein levels in MS patients are due to a reduced ability of the body to contain the inflammatory response typical of the disease.

The S100A8 protein is highly expressed and secreted by neutrophils, activated macrophages and endothelial cells during inflammatory processes (Lim et al., 2008). Nitrosilation is an important post-translational modification that regulates NO nitrogen monoxide transport, cell signaling mechanisms and homeostasis. In fact, the S100A8-SNO nitrosylate proteomy seems to play a role in regulating interactions between leukocytes and endothelial cells in the micro-circulation and suppression of mediated mast cell inflammation (Lim et al., 2008). Therefore, an increase in S100A8-SNO protein levels in MS subjects may be related to an attempt to reduce cell-mediated inflammatory response.

Particularly interesting is the increased concentration of the antileukoproteinase protein that we have observed in the saliva of MS patients. Antileukoproteinase was observed in various types of secretions, including seminal fluid, cervical mucus and bronchial secretions. It is a serine protease inhibitor (Thompson & Ohlsson 1986; Eisenberg et al., 1990) and is involved in resolving inflammatory response by suppressing the activity of proteases by attenuating the innate immune response and activating and induction of proliferation of lymphocytes B (Wang et al., 2003; Xu et al., 2007). In CNS, antileukoproteinase increases as a result of ischemia (Wang et al., 2003) and spinal cord injury (Urso et al., 2007). In addition to these functions, recently has been demonstrated a preminent role of antileukoproteinase in the repair of neuronal tissues through increased neural stem cell proliferation and their differentiation in oligodendrocytes, specialized nerve cells deputy to the production of myelin. This study, conducted on murine models of MS with Autoimmune Experimental Encephalitis, showed a strong increase in antiileukoproteinase in macrophages, microglia, neuronal cells and astrocytes, confirming its involvement in in-vivo repair of CNS (Mueller et al., 2008) and suggesting its possible utility in the development of effective therapeutic treatments to counter the destruction of the myelin sheath and the cells responsible for its production. High secretion of this protein in saliva of subjects with MS is therefore in line with these recent observations, representing a protective response from the body against tissue damage caused by a persistent flogosis and could be a an important diagnostic tool for MS.

3.2 Salivary proteome in Autoimmune Hepatitis subjects

3.2.1 Patients population

A total of 41 AIH patients were recruited (52.2 ± 15.4 years old, males n = 7, females n = 34) from the Center for the Study of Liver Diseases, Department of Medical Sciences "M. Aresu", University of Cagliari, Sardinia, Italy. The diagnosis of AIH was based on the criteria reviewed by the International Autoimmune Hepatitis Study Group (IAIHG) in 1999 (Boberg, 2002).

Furthemore table 7 reports demographic features and clinical data of AIH patients. The serum levels of alkaline phosphatase (ALP), γ-glutamyl transpeptidase (γ-GT), AST, ALT, total bilirubin (TBIL) and serum immunoglobulin G (IgG) were measured in each patient at the time of saliva collection. Patients were also tested for antinuclear antibodies (ANA), anti-smooth antibodies (ASMA) and renal microsomal antigen antibodies (LKM). Seropositivity was observed in 73% of patients for ANA, 51% for ASMA and 7% of cases for LKM.

The presence of other concurrent autoimmune diseases was investigated; 21 out of 41 subjects presented at list one of the following disease: Hashimoto's thyroiditis, rheumatoid arthritis, Crohn's disease, ulcerative recto colitis, systemic lupus erythematosus, Sjogren's syndrome, celiac disease and Basedow's disease as shown in table 6.

Table 6. Precence of autoimmune diseases concomitant with AIH.

Autoimmune disease	AIH	%
Hashimoto's thyroiditis	11	26.8
Rheumatoid arthritis	4	9.7
Crohn's disease	1	2.4
Recto colitis	2	4.9
Systemic lupus erythematosus	1	2.4
Sjogren's syndrome	1	2.4
Celiac disease	3	7.3
Basedow's disease	1	2.4

Furthermore, 8 patients presented with liver cirrhosis, which was diagnosed by histological examination in 5 cases and by ultrasound and/or laboratory findings in the remaining 3 subjects. The histological stage of fibrosis was assessed in 29 patients out of

41. At the time of inclusion in this study 36 (88%) patients were under immunosuppressive treatment.

Table 7. Demographic features and clinical data of AIH peatients.

Parameters		
Age, Average (range)	Years	52 (17-86)
Gender, n (%)	Female	34 (83)
BMI, Average (range)	Kg/m^2	24 (17-38)
Cirrhosis (diagnosis on imaging techniques or blood tests), n (%) **Cirrhosis (histological diagnosis),** n (%)		3 (7.3) 5 (12)
Histological stage according to DESMET, n (%)	• I • II • III • IV • Not available	11 (26.8) 7 (17) 6 (14.6) 5 (12) 12 (29.2)
Positivity to autoantibodies, n (%)	• ANA • ASMA • LKM	30 (73) 21 (51.2) 3 (7.3)
AST, Median (range)	IU/l	25 (13-78)
ALT, Median (range)	IU/l	22 (12-67)
GGT, Median (range)	IU/l	28 (6-167)
ALP, Median (range)	IU/l	70 (28-216)
Gammaglobulins, Median (range)	g/dl	1.39 (0.69-2.51)
Albumine, Median (range)	g/dl	4 (1.2-4.83)
Prothrombin time, Median (range)	INR	1.03 (0.92-1.06)
Total Bilirubin, Median (range)	mg/dl	0.69 (0.25-2.19)
Platelets, Median (range)	$10^9/l$	212.5 (91-423)

3.2.2 Quantitative analysis in Autoimmune Hepatitis

For this preliminary study we determined the levels of proteins and peptides secreted by salivary glands, such as S-type cystatins, histatins and statherins and their naturally occurring proteoforms deriving from post-translational modifications.

The acidic soluble fraction of whole saliva of AIH patients and 41 control subjects was analyzed by top-down RP-HPLC low resolution ESI-MS proteomic approach. Protein/peptide quantification was based on relative abundance by measuring the area eXtracted ion current (XIC) peaks as shown in table 2.

For statistical analysis, the XIC peak areas of AIH patients and controls subjects were compared using the software Grafpad Prism for calculating means and standard deviations of peptides/proteins. The results of statistical analysis were reported in table 8.

S-type cystatins (considered the sum of all S, S1and S2 proteoforms present in saliva) have an increased level in AIH subjects with respect to control. In particular, the mono- and di-oxidated proteoforms of cystatin S1 showed an higher frequency and level in AIH patients with respect to control with a p value of <0.0001 and 0.007 respectively (figure 28, panel A), while the non oxidized proteoform was found not statistically variated between the two groups.

On the contrary, the oxidized proteoforms of cystatin S2 did not show any statistically significant variation between the two groups, but an increased level was observed in AIH patients only for the non oxidized proteoform with a p value of 0.02.

Likewise to cystatin S1 also cystatins SN showed an increased levels of oxidized proteoform (p value 0.04) but not for the non oxidized one in AIH patients respect to control. Regarding cystatin SN, it is whorty to note that also the truncated proteoform cystatin SN Des_{1-4} was found increased in AIH with a p value of 0.03 (figure 28, panel B).

Increased levels of truncated proteoforms were also observed for histatins family. In particular histatin 5 and histatin 6, both fragments belonging to histatin 3, were found with significantly increased level in AIH patients respect to controls with a p value of 0.04 and 0.005, respectively (figure 28, panel D). The increase levels of histatins 5 and 6 would not seem to derive from a major cleavage of histatin 3, seems we observed also an increased level of this protein in subjects affected by AIH respect to control even though it is not statistically significant. Moreover also histatin 1 was found with increased level in AIH and statistically significant with a p value of 0.04 (figure 28, panel D).

The di-phosphorylated and mono-phosphorylated statherin were also found significantly increased in AIH subjects with respect to controls with a p value of 0.02 and 0.04, respectively (figure 28, panel C).

Table 8. Proteins and peptides investigated, XIC peak areas mean ± standard deviation (SD) ($x10^9$), frequency and statistically variated between MS and control subjects are reported.

Proteins/Peptides	AIH		Controls		*p* value
	Mean ± SD	Frequency	Mean ± SD	Frequency	
Type 2 Cystatins					
Cystatin S	0.13 ± 0.15	26/41	0.085 ± 0.11	32/41	ns
Cystatin S1	1.78 ± 1.87	37/41	1.15 ± 1.36	37/41	ns
Cystatin S1 mono-ox	0.34 ± 0.43	30/41	0.044 ± 0.071	16/41	<0.0001
Cystatin S1 di-ox	0.13 ± 0.21	20/41	0.026 ± 0.054	10/41	0.007
Cystatin S1 Total	2.25 ± 2.33	37/41	1.22 ± 1.37	37/41	0.01
Cystatin S2	0.78 ± 1.07	32/41	0.38 ± 0.60	35/41	0.02
Cystatin S2 mono-ox	0.11 ± 0.25	9/41	0.0073 ± 0.023	4/41	ns
Cystatin S2 di-ox	0.044 ± 0.13	6/41	0.0020 ± 0.0090	2/41	ns
Cystatin S2 Total	0.93 ± 1.39	32/41	0.38 ± 0.60	35/41	0.03
Cystatin S Total	3.31 ± 3.75	37/41	1.69 ± 1.87	37/41	0.01
Cystatin SN	3.12 ± 3.97	34/41	2.23 ± 3.79	37/41	ns
Cystatin SN mono-ox	0.50 ± 0.85	22/41	0.11 ± 0.27	18/41	0.04
Cystatin SN di-ox	0.028 ± 0.13	3/41	0.016 ± 0.041	6/41	ns
Cystatin SN Des_{1-4}	0.55 ± 1.17	23/41	0.35 ± 1.25	14/41	0.03
Cystatin SA	0.42 ± 0.77	19/41	0.33 ± 0.65	19/41	ns
Cystatin SA mono-ox	0.037 ± 0.18	4/41	0.027 ± 0.053	10/41	ns
Histatins					
Histatin 1 0P	0.037 ± 0.095	10/41	0.021 ± 0.046	9/41	ns

Proteins/Peptides	AIH		Controls		*p* value
	Mean ± SD	Frequency	Mean ± SD	Frequency	
Histatin 1	0.71 ± 0.98	33/41	0.29 ± 0.33	27/41	0.04
Histatin 3	0.40 ± 0.77	23/41	0.12 ± 0.19	17/41	ns
Histatin 5	0.83 ± 1.15	33/41	0.35 ± 0.39	32/41	0.04
Histatin 6	0.31 ± 0.38	32/41	0.11 ± 0.15	24/41	0.005
Histatin Total	2.28 ± 3.15	35/41	0.89 ± 0.99	36/41	0.02
Statherin					
Statherin 0P	0.0026 ± 0.0086	4/41	0.0016 ± 0.0051	5/41	ns
Statherin 1P	0.076 ± 0.082	34/41	0.043 ± 0.047	31/41	0.04
Statherin	3.14 ± 3.56	40/41	1.73 ± 2.22	36/41	0.02
Statherin Total	3.22 ± 3.63	40/41	1.77 ± 2.26	36/41	0.01

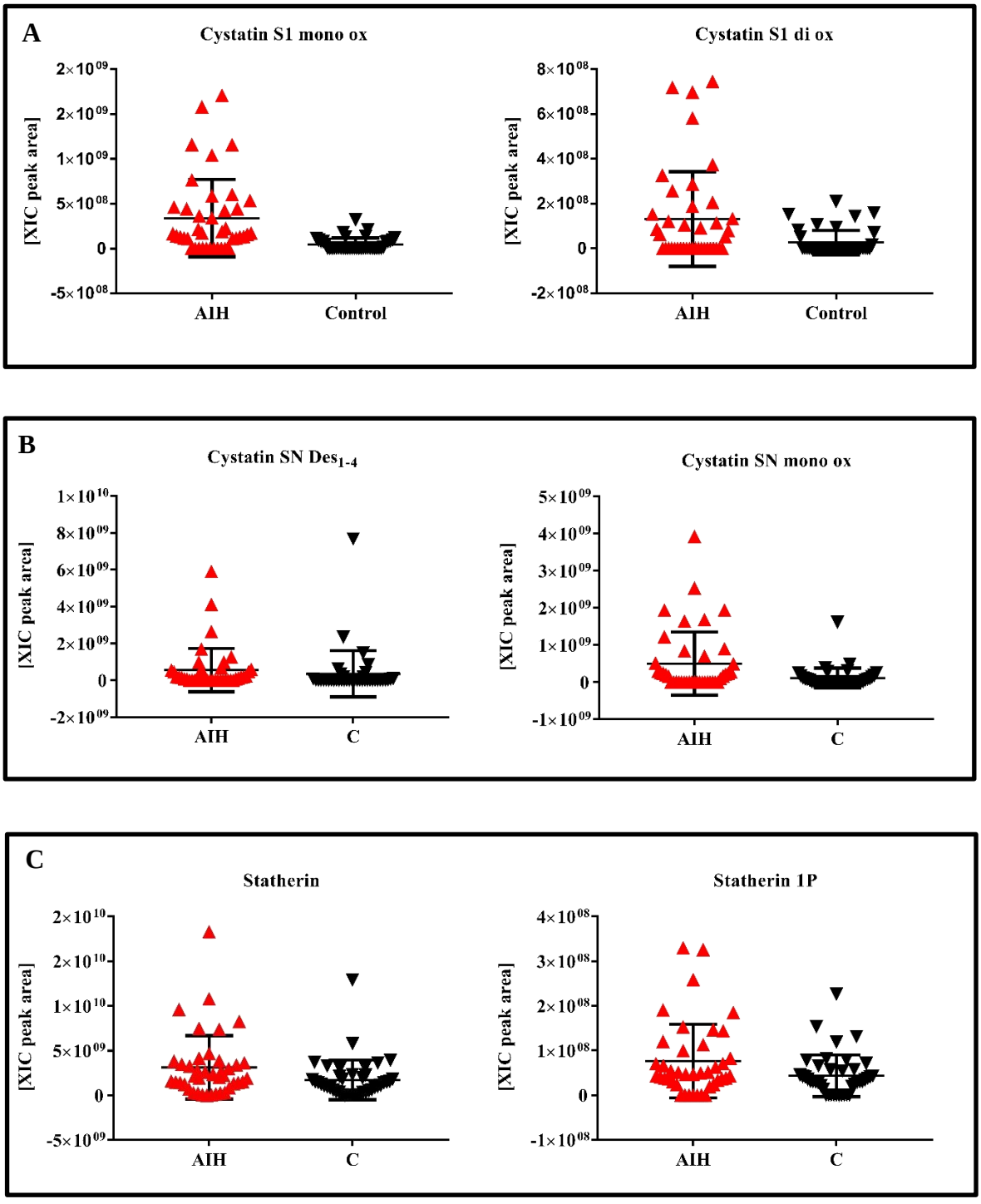


Figure 28. Distribution of the XIC peak area values measured in saliva and statistically variated from AIH and C subjects of (A) cystatin S1 oxidized proteoforms; (B) cystatin SN proteoforms; (C) statherin and statherin 1P; (D) histatins proteoforms. (Continue in the nexts pages).

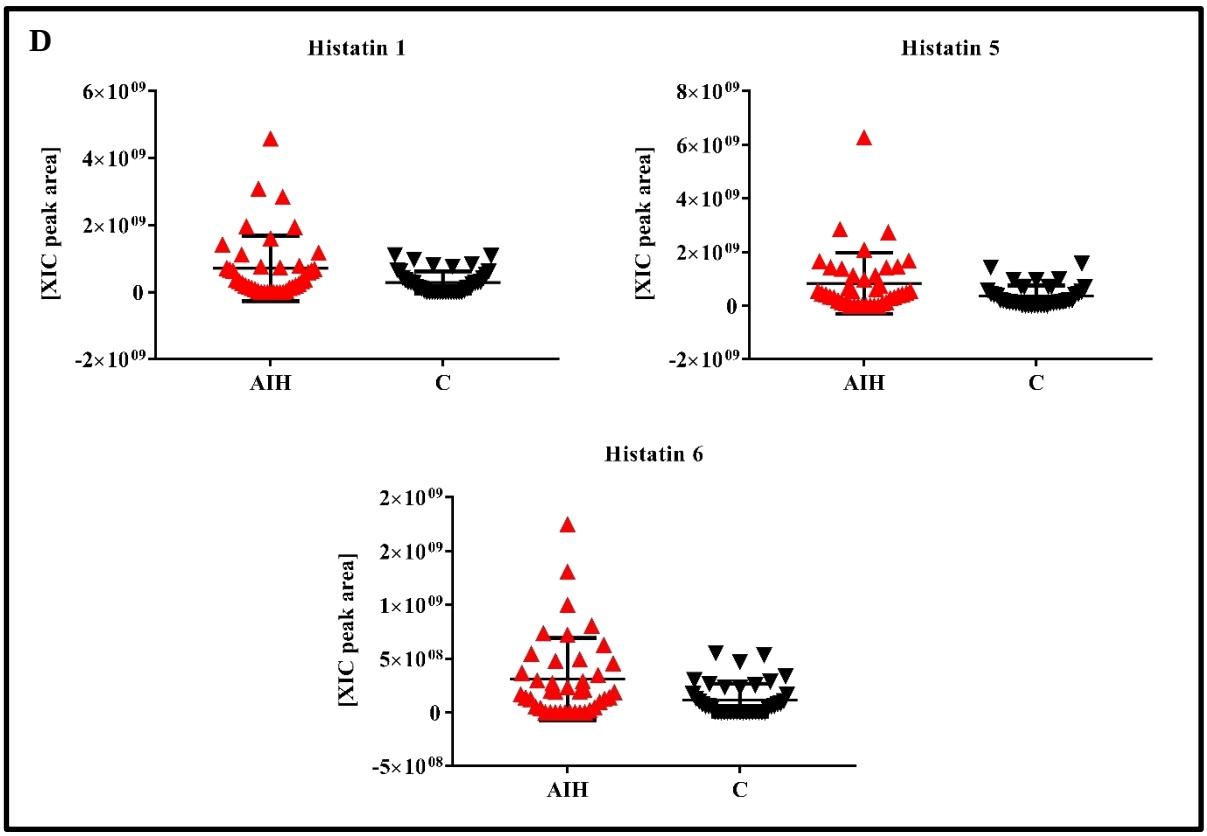


Figure 28. Distribution of the XIC peak area values measured in saliva and statistically variated from AIH and C subjects of (A) cystatin S1 oxidized proteoforms; (B) cystatin SN proteoforms; (C) statherin and statherin 1P; (D) histatins proteoforms.

3.2.3 Discussion

The purpose of this study was to explore qualitative difference among the proteins secreted from salivary glands in AIH subjects respect to control.

The general characteristics of the patients enrolled in this study reflect the fact that AIH has a strong female predominance with a *ratio* of 3:1 (34 of the patients were female) and it is usually concomitant with different extrahepatic autoimmune diseases.

We were able to compare the levels of proteins and peptides secreted by salivary glands, such as S-type cystatins, histatins and statherin. In particular, the mono- and di-oxidated proteoforms of cystatin S1, cystatins SN mono-oxidized, cystatin SN Des_{1-4} were found increased in AIH patients. Increased levels of truncated proteoforms were also observed for histatins family, in particular histatin 1, 5 and 6. The di-phosphorylated and mono-phosphorylated statherin were found significantly increased in AIH subjects with respect to controls.

One important role of the proteins that have been observed increased in AIH patients, is related to their activity in the creation of protective environment for the teeth and in the modulation of the bacteria adhesion to the oral surfaces. Moreover, they are involved in the formation of the protein pellicles covering the oral surfaces (Delius et al., 2017). Therefore, the increase levels of these proteins, which are also involved in the processes of oral inflammation, could reflect an alteration of the oral microenvironment.

To date there are no studies that highlight the variation of salivary proteome composition in autoimmune diseases related to the liver. Moreover, the patients recruited for this study had other immunological diseases concurrently, therefore the results obtained can also be influenced by the characteristics of other syndromes. For instance an increased level of salivary cystatins has been observed in Sjögren's syndrome, an autoimmune condition affecting the lacrimal and salivary glands. The inflammation of the salivary glands leads to reduction in salivary output, which imposes a significant impact on oral health (Tucci, Quatraro, & Silvestris, 2005). By collecting parotid saliva from Sjögren's syndrome patients and controls Carpenter et al. evidenced that cystatins could be a specifity marker for this pathology (Carpenter et al., 2000).

In the same way, the concomitant presence in AIH patients of systemic lupus erythematosus, which is a chronic autoimmune disease where autoantibodies are directed against all organs, including the salivary glands (Loyola Rodriguez et al., 2016), may giustify the increase levels of oxidized cystatins that we observed in AIH patients respect to control. On this regard, recently has been evidenced a decrease antioxidant capacity in

the saliva and serum of patients with systemic lupus erythematosus (Moori, Ghafoori, & Sariri, 2016).

Further studies have to be carried out to better explain the high overexpression of proteins involved in the protection of the oral cavity in subjects affected by AIH. A possible cause could be the simultaneous presence of autoimmune diseases involving the oral cavity in half of patients with AIH which could either have damage the oral mucosa more or modified the natural bacterial flora present in the oral cavity or both, generating an over-expression of the protein classes involved in the protection of the oral cavity.

4. Reference

Abe, N., Kadowaki, T., Okamoto, K., Nakayama, K., Ohishi, M., & Yamamoto, K. (1998). Biochemical and functional properties of lysine-specific cysteine proteinase (Lys-gingipain) as a virulence factor of Porphyromonas gingivalis in periodontal disease. *Journal of Biochemistry*, *123*(2), 305–12.

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature*, *422*(6928), 198–207.

Aebersold, R., & Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. *Nature*, *537*(7620), 347–355.

Agarwal, K., Czaja, A. J., & Donaldson, P. T. (2007). A functional Fas promoter polymorphism is associated with a severe phenotype in type 1 autoimmune hepatitis characterized by early development of cirrhosis. *Tissue Antigens*, *69*(3), 227–35.

Agarwal, K., Czaja, A. J., Jones, D. E., & Donaldson, P. T. (2000). Cytotoxic T lymphocyte antigen-4 (CTLA-4) gene polymorphisms and susceptibility to type 1 autoimmune hepatitis. *Hepatology*, *31*(1), 49–53.

Ahmad, M., Piludu, M., Oppenheim, F. G., Helmerhorst, E. J., & Hand, A. R. (2004). Immunocytochemical localization of histatins in human salivary glands. *The Journal of Histochemistry and Cytochemistry*, *52*(3), 361–70.

Ai J., Tang Q., Wu Y., Xu Y., Feng T., Zhou R., Chen Y., Gao X., Zhu Q., Yue X., Pan Q., Xu S., Li J., Huang M., Daugherty-Holtrop J., He Y., Xu H. E., Fan J., Ding J., & Geng M. (2011). The role of polymeric immunoglobulin receptor in inflammation-induced tumor metastasis of human hepatocellular carcinoma. *Journal of National Cancer Institute*; 103:1696-712.

Alakurtti, K., Weber, E., Rinne, R., Theil, G., Haan, G. J., Lindhout, D., Salmikangas, P., Saukko, P., Lahtinen, U., & Lehesjoki, A. E. (2005). Loss of lysosomal association of cystatin B proteins representing progressive myoclonus epilepsy, EPM1, mutations. *European Journal of Human Genetics*, *13*(2), 208–215.

Alvarez-Díaz, S., Valle, N., García, J. M., Peña, C., Freije, J. M. P., Quesada, V., Astudillo, A., Bonilla, F., López-Otín, C., & Muñoz, A. (2009). Cystatin D is a candidate tumor suppressor gene induced by vitamin D in human colon cancer cells. *The Journal of Clinical Investigation*, *119*(8), 2343–58.

Alvarez-Fernandez, M., Liang, Y.-H., Abrahamson, M., & Su, X.-D. (2005). Crystal structure of human cystatin D, a cysteine peptidase inhibitor with restricted inhibition profile. *The Journal of Biological Chemistry*, *280*(18), 18221–8.

Anderson, N. L., & Anderson, N. G. (1998). Proteome and proteomics: New technologies, new concepts, and new words. *Electrophoresis*, *19*(11), 1853–1861.

Andrassy, M., Igwe, J., Autschbach, F., Volz, C., Remppis, A., Neurath, M. F., Schleicher, E., Humpert, P. M., Wendt, T., Liliensiek, B., Morcos, M., Schiekofer, S., Thiele, K., Chen, J., Kientsch-Engel, R., Schmidt, A. M., Stremmel, W., Stern, D. M. Katus, H. A. Nawroth, P. P., & Bierhaus, A. (2006). Posttranslationally modified proteins as mediators of sustained intestinal inflammation. *The American Journal of Pathology*, *169*(4), 1223–37.

Asano M., Komiyama K. (2011). Polymeric immunoglobulin receptor. *Journal of Oral Science*; 53:147-56.

Ascherio, A., Rubertone, M., Spiegelman, D., Levin, L., Munger, K., Peck, C., & Lennette, E. (2005). Notice of retraction: "Multiple sclerosis and Epstein-Barr virus" (JAMA. 2003;289:1533-1536). *JAMA*, *293*(20), 2466.

Azen, E. A., Amberger, E., Fisher, S., Prakobphol, A., & Niece, R. L. (1996). PRB1, PRB2, and PRB4 coded polymorphisms among human salivary concanavalin-A binding, II-1, and Po proline-rich proteins. *American Journal of Human Genetics*, *58*(1), 143–53.

Azen, E. A., Latreille, P., & Niece, R. L. (1993). PRBI gene variants coding for length and null polymorphisms among human salivary Ps, PmF, PmS, and Pe proline-rich proteins (PRPs). *American Journal of Human Genetics*, *53*(1), 264–78.

Azen, E. A., Minaguchi, K., Latreille, P., & Kim, H. S. (1990). Alleles at the PRB3 locus coding for a disulfide-bonded human salivary proline-rich glycoprotein (Gl 8) and a null in an Ashkenazi Jew. *American Journal of Human Genetics*, *47*(4), 686–97.

Badamchian, M., Damavandy, A. A., Damavandy, H., Wadhwa, S. D., Katz, B., & Goldstein, A. L. (2007). Identification and quantification of thymosin beta4 in human saliva and tears. *Annals of the New York Academy of Sciences*, *1112*(1), 458–65.

Balbín, M., Freije, J. P., Abrahamson, M., Velasco, G., Grubb, A., & López-Otín, C. (1993). A sequence variation in the human cystatin D gene resulting in an amino acid (Cys/Arg) polymorphism at the protein level. *Human Genetics*, *90*(6), 668–9.

Baron, A.; DeCarlo, A.; Featherstone, J. (1999). Functional aspects of the human salivary cystatins in the oral environment. *Oral Diseases*, *5*(3), 234–40.

Batycka-Baran, A., Hattinger, E., Zwicker, S., Summer, B., Zack Howard, O. M., Thomas, P., Szepietowski, J. C., Ruzicka, T., Prinz, J. C., & Wolf, R. (2015). Leukocyte-derived koebnerisin (S100A15) and psoriasin (S100A7) are systemic mediators of inflammation in psoriasis. *Journal of Dermatological Science*, *79*(3), 214–21.

Bencharit, S., Altarawneh, S. K., Baxter, S. S., Carlson, J., Ross, G. F., Border, M. B., Mack, C. R., Byrd, W. C., Dibble, C. F., Barros, S., Loewy, Z., & Offenbacher, S. (2012). Elucidating role of salivary proteins in denture stomatitis using a proteomic approach. *Molecular bioSystems*, *8*(12), 3216–23.

Bennick, A. (1982). Salivary proline-rich proteins. *Molecular and Cellular Biochemistry, 45*(2), 83–99.

Bennick, A., Cannon, M., & Madapallimattam, G. (1981). Factors affecting the adsorption of salivary acidic proline-rich proteins to hydroxyapatite. *Caries Research, 15*(1), 9–20.

Bergey, E. J., Levine, M. J., Reddy, M. S., Bradway, S. D., & Al-Hashimi, I. (1986). Use of the photoaffinity cross-linking agent N-hydroxysuccinimidyl-4-azidosalicylic acid to characterize salivary-glycoprotein-bacterial interactions. *The Biochemical Journal, 234*(1), 43–8.

Blaydon, D. C.; Nitoiu, D.; Eckl, K. M.; Cabral, R. M.; Bland, P.; Hausser, I.; Van Heel, D. A.; Rajpopat, S.; Fischer, J.; Oji, V.; Zvulunov, A.; Traupe, H.; Hennies, H. C.; & Kelsell, D. P. (2011). Mutations in CSTA, encoding cystatin A, underlie exfoliative ichthyosis and reveal a role for this protease inhibitor in cell-cell adhesion. *Am. J. Hum. Genet. 89*(4), 564–71.

Bobek, L. A.; Ramasubbu, N.; Wang, X.; Weaver, T. R.; & Levine, M. J. (1994). Biological activities and secondary structures of variant forms of human salivary cystatin SN produced in *Escherichia coli*. *Gene, 151*(1–2), 303–8.

Boberg, K. M. (2002). Prevalence and epidemiology of autoimmune hepatitis. *Clinics in Liver Disease, 6*(3), 635–47.

Bode, W.; Engh, R.; Musil, D.; Thiele, U.; Huber, R.; Karshikov, A.; Brzin, J.; Kos, J.; Turk, V. (1988). The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. EMBO J. *7*(8), 2593–9.

Bogdanov, B., & Smith, R. D. (2005). Proteomics by FTICR mass spectrometry: Top down and bottom up. *Mass Spectrometry Reviews, 24*(2), 168–200.

Browne, P., Chandraratna, D., Angood, C., Tremlett, H., Baker, C., Taylor, B. V, & Thompson, A. J. (2014). Atlas of Multiple Sclerosis 2013: A growing global problem with widespread inequity. *Neurology, 83*(11), 1022–4.

Cabras, T., Boi, R., Pisano, E., Iavarone, F., Fanali, C., Nemolato, S., Faa, G., Castagnola, M., & Messana, I. (2012a). HPLC-ESI-MS and MS/MS structural characterization of multifucosylated N-glycoforms of the basic proline-rich protein IB-8a CON1+ in human saliva. *Journal of Separation Science, 35*(9), 1079–86.

Cabras, T., Iavarone, F., Manconi, B., Olianas, A., Sanna, M. T., Castagnola, M., & Messana, I. (2014). Top-down analytical platforms for the characterization of the human salivary proteome. *Bioanalysis, 6*(4), 563–81.

Cabras, T., Inzitari, R., Fanali, C., Scarano, E., Patamia, M., Sanna, M. T., Pisano, E., Giardina, B., Castagnola, M., & Messana, I. (2006). HPLC-MS characterization of cyclo-

statherin Q-37, a specific cyclization product of human salivary statherin generated by transglutaminase 2. *Journal of Separation Science*, *29*(17), 2600–8.

Cabras, T., Manconi, B., Iavarone, F., Fanali, C., Nemolato, S., Fiorita, A., Scarano, E., Passali, G. C., Manni, A., Cordaro, M., Paludetti, G., Faa, G., Messana, I., & Castagnola, M. (2012b). RP-HPLC–ESI-MS evidenced that salivary cystatin B is detectable in adult human whole saliva mostly as S-modified derivatives: S-Glutathionyl, S-cysteinyl and S–S 2-mer. *Journal of Proteomics*, *75*(3), 908–913.

Cabras, T., Pisano, E., Boi, R., Olianas, A., Manconi, B., Inzitari, R., Fanali, C., Giardina, B., Castagnola, M., & Messana, I. (2009). Age-Dependent Modifications of the Human Salivary Secretory Protein Complex research articles. *Journal of Proteome Research*, 8, 4126–4134.

Cabras, T., Pisano, E., Montaldo, C., Giuca, M. R., Iavarone, F., Zampino, G., Castagnola, M., & Messana, I. (2013). Significant modifications of the salivary proteome potentially associated with complications of Down syndrome revealed by top-down proteomics. *Molecular & Cellular Proteomics*, *12*(7), 1844–52.

Cabras, T., Sanna, M., Manconi, B., Fanni, D., Demelia, L., Sorbello, O., Iavarone, F., Castagnola, M., Faa, G., & Messana, I. (2015). Proteomic investigation of whole saliva in Wilson's disease. *Journal of Proteomics*, 128, 154–163.

Cai, K., & Bennick, A. (2004). Processing of acidic proline-rich proprotein by human salivary gland convertase. *Archives of Oral Biology*, *49*(11), 871–879.

Cai, W., Tucholski, T. M., Gregorich, Z. R., & Ge, Y. (2016). Top-down Proteomics: Technology Advancements and Applications to Heart Diseases. *Expert Review of Proteomics*, *13*(8), 717–730.

Carlsson, H., Yhr, M., Petersson, S., Collins, N., Polyak, K., & Enerbäck, C. (2005). Psoriasin (S100A7) and calgranulin-B (S100A9) induction is dependent on reactive oxygen species and is downregulated by Bcl-2 and antioxidants. *Cancer Biology & Therapy*, *4*(9), 998–1005.

Carpenter, G. H., Proctor, G. B., Pankhurst, C. L., O'Donohue, J., Scott, D., & Hunnable, M. P. (2000). Sialochemical markers of salivary gland involvement with Sjögren's syndrome secondary to rheumatoid arthritis and primary biliary cirrhosis. *Journal of Oral Pathology & Medicine*, *29*(9), 452–9.

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

Castagnola, M., Cabras, T., Iavarone, F., Vincenzoni, F., Vitali, A., Pisano, E., Nemolato, S., Scarano, E., Fiorita, A., Vento, G., Tirone, C., Romagnoli, C., Cordaro, M., Paludetti,

G., Faa, G., & Messana, I. (2012). Top-down platform for deciphering the human salivary proteome. *The Journal of Maternal-Fetal & Neonatal Medicine*, *25*(25S5), 1476–7058.

Castagnola, M., Inzitari, R., Fanali, C., Iavarone, F., Vitali, A., Desiderio, C., Vento, G., Tirone, C., Romagnoli, C., Cabras, T., Manconi, B., Sanna, M. T., Boi, R., Pisano, E., Olianas, A., Pellegrini, M., Nemolato, S., Heizmann, C. W., Faa, G., Messana, I. (2011a). The surprising composition of the salivary proteome of preterm human newborn. *Molecular & Cellular Proteomics*, *10*(1), M110.003467.

Castagnola, M., Inzitari, R., Rossetti, D. V., Olmi, C., Cabras, T., Piras, V., Nicolussi, P., Sanna, M. T., Pellegrini, M. Giardina, B., & Messana, I. (2004). A cascade of 24 histatins (histatin 3 fragments) in human saliva. Suggestions for a pre-secretory sequential cleavage pathway. *The Journal of Biological Chemistry*, *279*(40), 41436–43.

Castagnola, M., Messana, I., Inzitari, R., Fanali, C., Cabras, T., Morelli, A., Pecoraro, A. M., Neri, G., Torrioli, M. G., & Gurrieri, F. (2008). Hypo-phosphorylation of salivary peptidome as a clue to the molecular pathogenesis of autism spectrum disorders. *Journal of Proteome Research*, *7*(12), 5327–32.

Castagnola, M., Picciotti, P. M., Messana, I., Fanali, C., Fiorita, a, Cabras, T., Calò, L., Pisano, E., Passali, G. C., Iavarone, F., Paludetti, G., & Scarano, E. (2011). Potential applications of human saliva as diagnostic fluid. *Acta Otorhinolaryngologica Italica*, *31*(6), 347–57.

Castagnola, M.; Cabras, T.; Inzitari, R.; Zuppi, C.; Rossetti, D. V.; Petruzzelli, R.; Vitali, A.; Loy, F.; Conti, G.; Fadda, M. B. (2003). Determination of the post-translational modifications of salivary acidic proline-rich proteins. *Eur. J. Morphol. 41*(2), 93–8.

Chaly, Y. V, Paleolog, E. M., Kolesnikova, T. S., Tikhonov, I. I., Petratchenko, E. V, & Voitenok, N. N. (2000). Neutrophil alpha-defensin human neutrophil peptide modulates cytokine production in human monocytes and adhesion molecule expression in endothelial cells. *European Cytokine Network*, *11*(2), 257–66.

Chan, M., & Bennick, A. (2001). Proteolytic processing of a human salivary proline-rich protein precursor by proprotein convertases. *European Journal of Biochemistry*, *268*(12), 3423–31.

Chiappin, S., Antonelli, G., Gatti, R., & De Palo, E. F. (2007). Saliva specimen : A new laboratory tool for diagnostic and basic investigation. *Clinica Chimica Acta*, *383*(1–2), 30–40.

Chugh, S., Suen, C., & Gramolini, A. (2010). Proteomics and mass spectrometry: what have we learned about the heart? *Current Cardiology Reviews*, *6*(2), 124–33.

Compston, D. A., Batchelor, J. R., & McDonald, W. I. (1976). B-lymphocyte alloantigens associated with multiple sclerosis. *Lancet*, *2*(7998), 1261–5.

Cookson, S., Constantini, P. K., Clare, M., Underhill, J. A., Bernal, W., Czaja, A. J., & Donaldson, P. T. (1999). Frequency and nature of cytokine gene polymorphisms in type 1 autoimmune hepatitis. *Hepatology, 30*(4), 851–6.

Crutchfield, C. A., Thomas, S. N., Sokoll, L. J., & Chan, D. W. (2016). Advances in mass spectrometry-based clinical biomarker discovery. *Clinical Proteomics, 13*(1), 1.

Cuevas-Córdoba, B., & Santiago-García, J. (2014). Saliva: A Fluid of Study for OMICS. *OMICS: A Journal of Integrative Biology, 18*(2), 87–97.

Cui, W., Rohrs, H. W., & Gross, M. L. (2011). Top-down mass spectrometry: recent developments, applications and perspectives. *The Analyst, 136*(19), 3854–64.

Czaja, A. J. (2002). Autoimmune hepatitis. *Clinics in Liver Disease, 6*(3), xi–xii.

Dagley, L. F., Emili, A., & Purcell, A. W. (2013). Application of quantitative proteomics technologies to the biomarker discovery pipeline for multiple sclerosis. *Proteomics. Clinical Applications, 7*(1–2), 91–108.

Davies, M. E., & Barrett, A. J. (1984). Immunolocalization of human cystatins in neutrophils and lymphocytes. *Histochemistry, 80*(4), 373–7.

Delius, J., Trautmann, S., Médard, G., Kuster, B., Hannig, M., & Hofmann, T. (2017). Label-free quantitative proteome analysis of the surface-bound salivary pellicle. *Colloids and Surfaces B: Biointerfaces*, 152, 68–76.

Dell'Avvento, S., Sotgiu, M. A., Manca, S., Sotgiu, G., & Sotgiu, S. (2016). Epidemiology of multiple sclerosis in the pediatric population of Sardinia, Italy. *European Journal of Pediatrics, 175*(1), 19–29.

Dickinson, D. P. (2002). Cysteine Peptidases of Mammals: Their Biological Roles and Potential Effects in the Oral Cavity and Other Tissues in Health and Disease. *Critical Reviews in Oral Biology & Medicine, 13*(3), 238–275.

Donato, R. (2003). Intracellular and extracellular roles of S100 proteins. *Microscopy Research and Technique, 60*(6), 540–551.

Driscoll, J., Zuo, Y., Xu, T., Choi, J. R., Troxler, R. F., & Oppenheim, E. (1995). Functional Comparison of Native and Recombinant Human Salivary Histatin 1. *Journal of Dental Research, 74*(12), 1837–1844.

Dyment, D. A., Yee, I. M. L., Ebers, G. C., Sadovnick, A. D., & Canadian Collaborative Study Group. (2006). Multiple sclerosis in stepsiblings: recurrence risk and ascertainment. *Journal of Neurology, Neurosurgery, and Psychiatry, 77*(2), 258–9.

Ebers, G. C., Sadovnick, A. D., & Risch, N. J. (1995). A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. *Nature, 377*(6545), 150–1.

Ebers, G. C., Yee, I. M., Sadovnick, A. D., & Duquette, P. (2000). Conjugal multiple sclerosis: population-based prevalence and recurrence risks in offspring. Canadian Collaborative Study Group. *Annals of Neurology, 48*(6), 927–31.

Eckert, R. L., Broome, A.-M., Ruse, M., Robinson, N., Ryan, D., & Lee, K. (2004). S100 proteins in the epidermis. *The Journal of Investigative Dermatology, 123*(1), 23–33.

Edgeworth, J., Gorman, M., Bennett, R., Freemont, P., & Hogg, N. (1991). Identification of p8,14 as a highly abundant heterodimeric calcium binding protein complex of myeloid cells. *The Journal of Biological Chemistry, 266*(12), 7706–13.

Ellias, M. F., Zainal Ariffin, S. H., Karsani, S. A., Abdul Rahman, M., Senafi, S., & Megat Abdul Wahab, R. (2012). Proteomic analysis of saliva identifies potential biomarkers for orthodontic tooth movement. *The Scientific World Journal*, 2012, 647240.

Fahey, J. V, & Wira, C. R. (2002). Effect of menstrual status on antibacterial activity and secretory leukocyte protease inhibitor production by human uterine epithelial cells in culture. *The Journal of Infectious Diseases, 185*(11), 1606–13.

Farquhar, C., VanCott, T. C., Mbori-Ngacha, D. A., Horani, L., Bosire, R. K., Kreiss, J. K., Richardson, B., & John-Stewart, G. C. (2002). Salivary secretory leukocyte protease inhibitor is associated with reduced transmission of human immunodeficiency virus type 1 through breast milk. *The Journal of Infectious Diseases, 186*(8), 1173–6.

Ferrer-Mayorga, G., Alvarez-Díaz, S., Valle, N., De Las Rivas, J., Mendes, M., Barderas, R., Canals, F., Tapia, O., Casal, J. I., Lafarga, M., & Muñoz, A. (2015). Cystatin D locates in the nucleus at sites of active transcription and modulates gene and protein expression. *Journal of Biological Chemistry, 290*(44), 26533–26548.

Ferrer-Mayorga, G.; Alvarez-Díaz, S.; Valle, N.; De Las Rivas, J.; Mendes, M.; Barderas, R.; Canals, F.; Tapia, O.; Casal, J. I.; Lafarga, M.; & Muñoz, A. (2015). Cystatin D locates in the nucleus at sites of active transcription and modulates gene and protein expression. *J. Biol. Chem, 290* (44), 26533–48.

Fu J. L., Wang Y. R., Li G. Z., Zhou Y., Liu P. (2012) Change in expression of the intestinal polymeric immunoglobulin receptor in acute liver necrosis. *Journal of Gastroenterology and Hepatology Research*; 1:69-73.

Gillece-Castro, B. L., Prakobphol, A., Burlingame, A. L., Leffler, H., & Fisher, S. J. (1991). Structure and bacterial receptor activity of a human salivary proline-rich glycoprotein. *The Journal of Biological Chemistry*, 266(26), 17358–68.

Gipson, T. S., Bless, N. M., Shanley, T. P., Crouch, L. D., Bleavins, M. R., Younkin, E. M., Sarma, V., Gibbs, D. F., Tefera, W., & Ward, P. A. (1999). Regulatory effects of endogenous protease inhibitors in acute lung inflammatory injury. *Journal of Immunology, 162*(6), 3653–62.

Giusti, L., Baldini, C., Ciregia, F., Giannaccini, G., Giacomelli, C., de Feo, F., Sedie, A., Delle Riente, L., Lucacchini, A., Bazzichi, L., &Bombardieri, S. (2010). Is GRP78/BiP a potential salivary biomarker in patients with rheumatoid arthritis? *Proteomics. Clinical Applications*, *4*(3), 315–324.

Giusti, L., Bazzichi, L., Baldini, C., Ciregia, F., Mascia, G., Giannaccini, G., Del Rosso, M., Bombardieri, S., & Lucacchini, A. (2007). Specific proteins identified in whole saliva from patients with diffuse systemic sclerosis. *Journal of Rheumatology*, *34*(10), 2063–2069.

Giusti, L., Sernissi, F., Donadio, E., Ciregia, F., Giacomelli, C., Giannaccini, G., Mazzoni, M. R., Lucacchini, A., & Bazzichi, L. (2016). Salivary psoriasin (S100A7) correlates with diffusion capacity of carbon monoxide in a large cohort of systemic sclerosis patients. *Journal of Translational Medicine*, *14*(1), 262.

Gläser, R., Harder, J., Lange, H., Bartels, J., Christophers, E., & Schröder, J.-M. (2005). Antimicrobial psoriasin (S100A7) protects human skin from *Escherichia coli* infection. *Nature Immunology*, *6*(1), 57–64.

Goebel, C., Mackay, L. G., Vickers, E. R., & Mather, L. E. (2000). Determination of defensin HNP-1, HNP-2, and HNP-3 in human saliva by using LC/MS. *Peptides*, *21*(6), 757–65.

Goyette, J., & Geczy, C. L. (2011). Inflammation-associated S100 proteins: new mechanisms that regulate function. *Amino Acids*, *41*(4), 821–42.

Grassl, N., Kulak, N. A., Pichler, G., Geyer, P. E., Jung, J., Schubert, S., Sinitcyn, P., Cox, J., & Mann, M. (2016). Ultra-deep and quantitative saliva proteome reveals dynamics of the oral microbiome. *Genome Medicine*, *8*(1), 44.

Greabu, M., Battino, M., Mohora, M., Totan, A., Didilescu, A., Spinu, T., Totan, C., Miricescu, D., & Radulescu, R. (2009). Saliva--a diagnostic window to the body, both in health and in disease. *Journal of Medicine and Life*, *2*(2), 124–132.

Gregorich, Z. R., & Ge, Y. (2014). Top-down proteomics in health and disease: Challenges and opportunities. *Proteomics*, *14*(10), 1195–1210.

Grubb, A. O. (2000). Cystatin C--properties and use as diagnostic marker. *Advances in Clinical Chemistry*, 35, 63–99.

Hagerman, A. E., & Butler, L. G. (1981). The specificity of proanthocyanidin-protein interactions. *The Journal of Biological Chemistry*, *256*(9), 4494–7.

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

Hannappel, E. (2010). Thymosin beta4 and its posttranslational modifications. *Annals of the New York Academy of Sciences*, 1194(1), 27–35.

Hansson, S. F., Simonsen, A. H., Zetterberg, H., Andersen, O., Haghighi, S., Fagerberg, I., Andréasson, U., Westman-Brinkmalm, A., Wallin, A., Rüetschi, U., & Blennow, K. (2007). Cystatin C in cerebrospinal fluid and multiple sclerosis. *Annals of Neurology, 62*(2), 193–196.

Hatton, M. N., Loomis, R. E., Levine, M. J., & Tabak, L. A. (1985). The role of carbohydrate in the lubricating property of a salivary glycoprotein-albumin complex. *Biochem. J, 230*(3), 817–820.

Haves-Zburof, D., Paperna, T., Gour-Lavie, A., Mandel, I., Glass-Marmor, L., & Miller, A. (2011). Cathepsins and their endogenous inhibitors cystatins: expression and modulation in multiple sclerosis. *Journal of Cellular and Molecular Medicine, 15*(11), 2421–9.

Hay, D. I., Carlson, E. R., Schluckebier, S. K., Moreno, E. C., & Schlesinger, D. H. (1987). Inhibition of calcium phosphate precipitation by human salivary acidic proline-rich proteins: structure-activity relationships. *Calcified Tissue International, 40*(3), 126–32.

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

Helmerhorst, E. J., & Oppenheim, F. G. (2007). Saliva: a dynamic proteome. *Journal of Dental Research, 86*(8), 680–93.

Helmerhorst, E. J., Venuleo, C., Beri, A., & Oppenheim, F. G. (2005). Candida glabrata is unusual with respect to its resistance to cationic antifungal proteins. *Yeast, 22*(9), 705–14.

Hernán, M. A., Jick, S. S., Logroscino, G., Olek, M. J., Ascherio, A., & Jick, H. (2005). Cigarette smoking and the progression of multiple sclerosis. *Brain: A Journal of Neurology, 128*(Pt 6), 1461–5.

Hirtz, C., Chevalier, F., Sommerer, N., & Raingeard, I. (2006). Salivary Protein Profiling in Type 1 Diabetes Using Two-Dimensional Electrophoresis and Mass Spectrometry. *Clinical Proteomics*, 117–127.

Hopsu-Havu, V. K., Joronen, I. A., Järvinen, M., Rinne, A., & Aalto, M. (1984). Cysteine proteinase inhibitors produced by mononuclear phagocytes. *Cell and Tissue Research, 236*(1), 161–4.

Hu, S., & Wong, D. T. (2007). Oral cancer proteomics. *Current Opinion in Molecular Therapeutics, 9*(5), 467–476.

Hubbard, M. J., & Cohen, P. (1993). On target with a new mechanism for the regulation of protein phosphorylation. *Trends in Biochemical Sciences, 18*(5), 172–7.

Huff, T., Müller, C. S., Otto, A. M., Netzker, R., & Hannappel, E. (2001). beta-Thymosins, small acidic peptides with multiple functions. *The International Journal of Biochemistry & Cell Biology*, *33*(3), 205–20.

Humphrey, S. P., & Williamson, R. T. (2001). A review of saliva: Normal composition, flow, and function. *The Journal of Prosthetic Dentistry*, *85*(2), 162–169.

Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., Bumgarner, R., Goodlett, D. R., Aebersold, R., Hood, L. (2001). Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science*, *292*(5518), 929–34.

Inzitari, R., Cabras, T., Onnis, G., Olmi, C., Mastinu, A., Sanna, M. T., Pellegrini, M., Castagnola, M., & Messana, I. (2005). Different isoforms and post-translational modifications of human salivary acidic proline-rich proteins. *Proteomics*, *5*(3), 805–815.

Inzitari, R., Cabras, T., Pisano, E., Fanali, C., Manconi, B., Scarano, E., Fiorita, A., Paludetti, G., Manni, A., Nemolato, S., Faa, G., Castagnola, M., Messana, I. (2009). HPLC-ESI-MS analysis of oral human fluids reveals that gingival crevicular fluid is the main source of oral thymosins beta(4) and beta(10). *Journal of Separation Science*, *32*(1), 57–63.

Inzitari, R., Cabras, T., Rossetti, D. V., Fanali, C., Vitali, A., Pellegrini, M., Paludetti, G., Manni, A., Giardina, B., Messana, I., & Castagnola, M. (2006). Detection in human saliva of different statherin and P-B fragments and derivatives. *Proteomics*, *6*(23), 6370–9.

Inzitari, R., Vento, G., Capoluongo, E., Boccacci, S., Fanali, C., Cabras, T., Romagnoli, C., Giardina, B., Messana, I., & Castagnola, M. (2007). Proteomic analysis of salivary acidic Proline-Rich Proteins in human preterm and at-term newborns. *Journal of Proteome Research*, *6*(4), 1371–1377.

Isemura, S. (2000). Nucleotide Sequence of Gene PBII Encoding Salivary Proline-Rich Protein P-B1. *J Biochem*, *127*(3), 393–398.

Isemura, S., Saitoh, E., Sanada, K., & Minakata, K. (1991). Identification of full-sized forms of salivary (S-type) cystatins (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of cystatin S) in human whole saliva and determination of phosphorylation sites of cystatin S. *Journal of Biochemistry*, *110*(4), 648–54.

Jarai, T., Maasz, G., Burian, A., Bona, A., Jambor, E., Gerlinger, I., & Mark, L. (2012). Mass spectrometry-based salivary proteomics for the discovery of head and neck squamous cell carcinoma. *Pathology Oncology Research*, *18*(3), 623–8.

Jaros, J. A. J., Guest, P. C., Bahn, S., & Martins-de-Souza, D. (2013). Affinity depletion of plasma and serum for mass spectrometry-based proteome analysis. *Methods in Molecular Biology*, 1002, 1–11.

Järvinen, M., & Rinne, A. (1982). Human spleen cysteineproteinase inhibitor. Purification, fractionation into isoelectric variants and some properties of the variants. *Biochimica et Biophysica Acta, 708*(2), 210–7.

Järvinen, M., Pernu, H., Rinne, A., Hopsu-Havu, V. K., & Altonen, M. (1983). Localization of three inhibitors of cysteineproteinases in the human oral mucosa. *Acta Histochemica, 73*(2), 279–82.

Jin, F. Y., Nathan, C., Radzioch, D., & Ding, A. (1997). Secretory leukocyte protease inhibitor: a macrophage product induced by and antagonistic to bacterial lipopolysaccharide. *Cell, 88*(3), 417–26.

Jou, Y. J., Lin, C. Der, Lai, C. H., Chen, C. H., Kao, J. Y., Chen, S. Y., Tsai, M. H., Huang, S., Hua Lin, C. W. (2010). Proteomic identification of salivary transferrin as a biomarker for early detection of oral cancer. *Analytica Chimica Acta, 681*(1–2), 41–48.

Kaetzel C. S. (2005) The polymeric immunoglobulin receptor: bridging innate and adaptive immune responses at mucosal surfaces. *Immunological Reviews; 206*:83-99.

Karlík, M., Valkovič, P., Hančinová, V., Krížová, L., Tóthová, Ľ., & Celec, P. (2015). Markers of oxidative stress in plasma and saliva in patients with multiple sclerosis. *Clinical Biochemistry, 48*(1–2), 24–28.

Kauffman, D.; Hofmann, T.; Bennick, A.; Keller, P. (1986). Basic proline-rich proteins from human parotid saliva: complete covalent structures of proteins IB-1 and IB-6. *Biochemistry, 25* (9), 2387–92.

Kikuchi, T., Abe, T., Hoshi, S., Matsubara, N., Tominaga, Y., Satoh, K., & Nukiwa, T. (1998). Structure of the murine secretory leukoprotease inhibitor (Slpi) gene and chromosomal localization of the human and murine SLPI genes. *American Journal of Respiratory Cell and Molecular Biology, 19*(6), 875–80.

Kingsmore, S. F. (2006). Multiplexed protein measurement: technologies and applications of protein and antibody arrays. *Nature Reviews. Drug Discovery, 5*(4), 310–20.

Klein JJ, Goldstein AL, White A. (1965). Enhancement of in vivo incorporation of labeled precursors into DNA and total protein of mouse lymph nodes after administration of thymic extracts. *Proc Natl Acad Sci USA*, 53:812-7.

Kopitar-Jerala, N., Schweiger, A., Myers, R. M., Turk, V., & Turk, B. (2005). Sensitization of stefin B-deficient thymocytes towards staurosporin-induced apoptosis is independent of cysteine cathepsins. *FEBS Letters, 579*(10), 2149–55.

Kos, J., Krasovec, M., Cimerman, N., Nielsen, H. J., Christensen, I. J., & Brünner, N. (2000). Cysteine proteinase inhibitors stefin A, stefin B, and cystatin C in sera from patients with colorectal cancer: relation to prognosis. *Clinical Cancer Research, 6*(2), 505–11.

Koutroukides, T. A., Guest, P. C., Leweke, F. M., Bailey, D. M. D., Rahmoune, H., Bahn, S., & Martins-de-Souza, D. (2011). Characterization of the human serum depletome by label-free shotgun proteomics. *Journal of Separation Science, 34*(13), 1621–6.

Kratchmarova, I., Blagoev, B., Haack-Sorensen, M., Kassem, M., & Mann, M. (2005). Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. *Science, 308*(5727), 1472–7.

Krawitt, E. L. (1998). Can you recognize autoimmune hepatitis? *Postgraduate Medicine, 104*(2), 145–9, 152.

Kurtzke, J. F. (1975). A reassessment of the distribution of multiple sclerosis. *Part one. Acta Neurologica Scandinavica, 51*(2), 110–36.

Kurtzke, J. F. (1993). Epidemiologic evidence for multiple sclerosis as an infection. *Clinical Microbiology Reviews, 6*(4), 382–427.

Lamkin, M. S., & Oppenheim, F. G. (1993). Structural features of salivary function. *Critical Reviews in Oral Biology and Medicine: An Official Publication of the American Association of Oral Biologists, 4*(3–4), 251–9.

Lang, H. L. E., Jacobsen, H., Ikemizu, S., Andersson, C., Harlos, K., Madsen, L., Hjorth, P., Sondergaard, L., Svejgaard, A., Wucherpfennig, K., Stuart, D. I., Bell, J. I., Jones, E. Y., & Fugger, L. (2002). A functional and structural basis for TCR cross-reactivity in multiple sclerosis. *Nature Immunology, 3*(10), 940–3.

Lavery, A. M., Verhey, L. H., & Waldman, A. T. (2014). Outcome measures in relapsing-remitting multiple sclerosis: capturing disability and disease progression in clinical trials. *Multiple Sclerosis International*, 2014, 262350.

Lee, Y.-H., & Wong, D. T. (2009). Saliva: an emerging biofluid for early detection of diseases. *American Journal of Dentistry, 22*(4), 241–8.

Lehrer, R. I., & Lu, W. (2012). α-Defensins in human innate immunity. *Immunological Reviews, 245*(1), 84–112.

Lehtinen, M. K., Tegelberg, S., Schipper, H., Su, H., Zukor, H., Manninen, O., Kopra, O., Joensuu, T., Hakala, P., Bonni, A., & Lehesjoki, A. E. (2009). Cystatin B deficiency sensitizes neurons to oxidative stress in progressive myoclonus epilepsy, EPM1. *The Journal of Neuroscience, 29*(18), 5910–5.

Levin, Y., Schwarz, E., Wang, L., Leweke, F. M., & Bahn, S. (2007). Label-free LC-MS/MS quantitative proteomics for large-scale biomarker discovery in complex samples. *Journal of Separation Science, 30*(14), 2198–2203.

Levine, M. J., Reddy, M. S., Tabak, L. A., Loomis, R. E., Bergey, E. J., Jones, P. C., Cohen, R. E., Stinson, M. W., & Al-Hashimi, I. (1987). Structural aspects of salivary glycoproteins. *Journal of Dental Research, 66*(2), 436–41.

Li, H., Li, G., Zhao, X., Wu, Y., Ma, W., Liu, Y., Gong, F., & Liang, S. (2013). Complementary serum proteomic analysis of autoimmune hepatitis in mice and patients. *Journal of Translational Medicine*, *11*(1), 146.

Lim, S. Y., Raftery, M. J., Goyette, J., Hsu, K., & Geczy, C. L. (2009). Oxidative modifications of S100 proteins: functional regulation by redox. *Journal of Leukocyte Biology*, *86*(3), 577–87.

Lin, R., Zhou, L., Zhang, J., & Wang, B. (2015). Abnormal intestinal permeability and microbiota in patients with autoimmune hepatitis. *International Journal of Clinical and Experimental Pathology*, *8*(5), 5153–60.

Lin, Y. Y.; Chen, Z. W.; Lin, Z. P.; Lin, L. B.; Yang, X. M.; Xu, L. Y.; & Xie, Q. (2016). Tissue Levels of Stefin A and Stefin B in Hepatocellular Carcinoma. *Anat. Rec. 299*(4), 428–38.

Littera, R., Chessa, L., Onali, S., Figorilli, F., Lai, S., Secci, L., La Nasa, G., Caocci, G., Arras, M., Melis, M., Cappellini, S., Balestrieri, C., Serra, G., Conti, M., Zolfino, T., Casale, M., Casu, S., Pasetto, M. C., Barca, L., Salustro, C., Matta, L., Scioscia, R., Zamboni, F., Faa, G., Orrù, S., & Carcassi, C. (2016). Exploring the Role of Killer Cell Immunoglobulin-Like Receptors and Their HLA Class I Ligands in Autoimmune Hepatitis. *PloS One*, *11*(1), e0146086.

Löhr, H. F., Schlaak, J. F., Lohse, A. W., Böcher, W. O., Arenz, M., Gerken, G., & Meyer Zum Büschenfelde, K. H. (1996). Autoreactive CD4+ LKM-specific and anticlonotypic T-cell responses in LKM-1 antibody-positive autoimmune hepatitis. *Hepatology*, *24*(6), 1416–21.

Loyola Rodriguez, J. P., Galvan Torres, L. J., Martinez Martinez, R. E., Abud Mendoza, C., Medina Solis, C. E., Ramos Coronel, S., Garcia Cortes, J. O., & Domínguez Pérez, R. A. (2016). Frequency of dental caries in active and inactive systemic lupus erythematous patients: salivary and bacterial factors. *Lupus*, *25*(12), 1349–56.

Lu, Y., & Bennick, A. (1998). Interaction of tannin with human salivary proline-rich proteins. *Archives of Oral Biology*, *43*(9), 717–28.

Lupi, A., Messana, I., Denotti, G., Schininà, M. E., Gambarini, G., Fadda, M. B., Vitali, A., Cabras, T., Piras, V., Patamia, M., Cordaro, M., Giardina, B., & Castagnola, M. (2003). Identification of the human salivary cystatin complex by the coupling of high-performance liquid chromatography and ion-trap mass spectrometry. *Proteomics*, *3*(4), 461–467.

Lyons, K. M., Stein, J. H., & Smithies, O. (1988). Length polymorphisms in human proline-rich protein genes generated by intragenic unequal crossing over. *Genetics*, *120*(1), 267–78.

Madapallimattam, G., & Bennick, A. (1990). Phosphopeptides derived from human salivary acidic proline-rich proteins. Biological activities and concentration in saliva. *The Biochemical Journal, 270*(2), 297–304.

Maeda, N. (1985). Inheritance of the human salivary proline-rich proteins: A reinterpretation in terms of six loci forming two subfamilies. *Biochemical Genetics, 23*(5–6), 455–464.

Maeda, N., Kim, H. S., Azen, E. A., & Smithies, O. (1985). Differential RNA splicing and post-translational cleavages in the human salivary proline-rich protein gene system. *The Journal of Biological Chemistry, 260*(20), 11123–30.

Mahad, D. H., Trapp, B. D., & Lassmann, H. (2015). Pathological mechanisms in progressive multiple sclerosis. *The Lancet. Neurology, 14*(2), 183–93.

Manconi, B., Cabras, T., Sanna, M., Piras, V., Liori, B., Pisano, E., Iavarone, F., Vincenzoni, F., Cordaro, M., Faa, G., Castagnola, M., & Messana, I. (2016b). N- and O-linked glycosylation site profiling of the human basic salivary proline-rich protein 3M. *Journal of Separation Science, 39*(10), 1987–1997.

Manconi, B., Castagnola, M., Cabras, T., Olianas, A., Vitali, A., Desiderio, C., Sanna, M. T., & Messana, I. (2016a). The intriguing heterogeneity of human salivary proline-rich proteins: Short title: Salivary proline-rich protein species. *Journal of Proteomics*, 134, 47–56.

Mandal, S., Curtis, L., Pind, M., Murphy, L. C., & Watson, P. H. (2007). S100A7 (psoriasin) influences immune response genes in human breast cancer. *Experimental Cell Research, 313*(14), 3016–25.

Mann, M., & Jensen, O. N. (2003). Proteomic analysis of post-translational modifications. *Nature Biotechnology, 21*(3), 255–61.

Marenholz, I., Heizmann, C. W., & Fritz, G. (2004). S100 proteins in mouse and man: from evolution to function and pathology (including an update of the nomenclature). *Biochemical and Biophysical Research Communications, 322*(4), 1111–22.

Marenholz, I., Lovering, R. C., & Heizmann, C. W. (2006). An update of the S100 nomenclature. *Biochimica et Biophysica Acta, 1763*(11), 1282–3.

Marrie, R. A. (2004). Environmental risk factors in multiple sclerosis aetiology. The Lancet. *Neurology, 3*(12), 709–18.

Marrosu, M. G., Muntoni, F., Murru, M. R., Costa, G., Pischedda, M. P., Pirastu, M., Sotgiu, S., Rosati, G., & Cianchetti, C. (1992). HLA-DQB1 genotype in Sardinian multiple sclerosis: evidence for a key role of DQB1 *0201 and *0302 alleles. *Neurology, 42*(4), 883–6.

Martini, D., Gallo, A., Vella, S., Sernissi, F., Cecchettini, A., Luciano, N., Polizzi, E., Conaldi, P. G., Mosca, M., & Baldini, C. (2017). Cystatin S-a candidate biomarker for severity of submandibular gland involvement in Sjögren's syndrome. *Rheumatology, 56*(6), 1031–1038.

Mason, R. W.; Sol-Church, K.; & Abrahamson, M. (1998) Amino acid substitutions in the N-terminal segment of cystatin C create selective protein inhibitors of lysosomal cysteine proteinases. *Biochemical J. 330*(2), 833–8.

Messana, I., Cabras, T., Iavarone, F., Manconi, B., Huang, L., Martelli, C., Olianas, A., Sanna, M. T., Pisano, E., Sanna, M., Arba, M., D'Alessandro, A., Desiderio, C., Vitali, A., Pirolli, D., Tirone, C., Lio, A., Vento, G., Romagnoli, C., Cordaro, M., Manni, A., Gallenzi, P., Fiorita, A., Scarano, E., Calò, L., Passali, G. C., Picciotti, P. M., Paludetti, G., Fanos, V., Faa, G., & Castagnola, M. (2015). Chrono-proteomics of human saliva: variations of the salivary proteome during human development. *Journal of Proteome Research, 14*(4), 1666–77.

Messana, I., Cabras, T., Iavarone, F., Vincenzoni, F., Urbani, A., & Castagnola, M. (2013). Unraveling the different proteomic platforms. *Journal of Separation Science, 36*(1), 128–139.

Messana, I., Cabras, T., Inzitari, R., Lupi, A., Zuppi, C., Olmi, C., Fadda, M. B., Cordaro, M., Giardina, B., & Castagnola, M. (2004). Characterization of the human salivary basic proline-rich protein complex by a proteomic approach. *Journal of Proteome Research, 3*(4), 792–800.

Messana, I., Cabras, T., Pisano, E., Sanna, M. T., Olianas, A., Manconi, B., Pellegrini, M., Paludetti, G., Scarano, E., Fiorita, A., Agostino, S., Contucci, A. M., Calò, L., Picciotti, P. M., Manni, A., Bennick, A., Vitali, A., Fanali, C., Inzitari, R., & Castagnola, M. (2008a). Trafficking and postsecretory events responsible for the formation of secreted human salivary peptides: a proteomics approach. *Molecular & Cellular Proteomics, 7*(5), 911–26.

Messana, I., Inzitari, R., Fanali, C., Cabras, T., & Castagnola, M. (2008b). Facts and artifacts in proteomics of body fluids. What proteomics of saliva is telling us? *Journal of Separation Science, 31*(11), 1948–1963.

Minaguchi, K., Madapallimattam, G., & Bennick, A. (1988). The presence and origin of phosphopeptides in human saliva. *The Biochemical Journal, 250*(1), 171–7.

Moore, B. W. (1965). A soluble protein characteristic of the nervous system. *Biochemical and Biophysical Research Communications, 19*(6), 739–44.

Moori, M., Ghafoori, H., & Sariri, R. (2016). Nonenzymatic antioxidants in saliva of patients with systemic lupus erythematosus. *Lupus, 25*(3), 265–71.

Moreno, E. C., Varughese, K., & Hay, D. I. (1979). Effect of human salivary proteins on the precipitation kinetics of calcium phosphate. *Calcified Tissue International, 28*(1), 7–16.

Muratori, P., Czaja, A.-J., Muratori, L., Pappas, G., Maccariello, S., Cassani, F., Granito, A., Ferrari, R., Mantovani, V., Lenzi, M., & Bianchi, F. B. (2005). Genetic distinctions between autoimmune hepatitis in Italy and North America. *World Journal of Gastroenterology, 11*(12), 1862–6.

Muttardi, K.; Nitoiu, D.; Kelsell, D. P.; O'Toole, E. A.; & Batta, K. (2016). Acral peeling skin syndrome associated with a novel CSTA gene mutation. *Clin. Exp. Dermatol. 41*(4), 394–8.

Nagai, A., Murakawa, Y., Terashima, M., Shimode, K., Umegae, N., Takeuchi, H., & Kobayashi, S. (2000). Cystatin C and cathepsin B in CSF from patients with inflammatory neurologic diseases. *Neurology, 55*(12), 1828–32.

Nagata, K., Nakao, M., Shibata, S., Shizukuishi, S., Nakamura, R., & Tsunemitsu, A. (1983). Purification and characterization of galactosephilic component present on the cell surfaces of Streptococcus sanguis ATCC 10557. *Journal of Periodontology, 54*(3), 163–72.

Nashida, T., Sato, R., Haga-Tsujimura, M., Yoshie, S., Yoshimura, K., Imai, A., & Shimomura, H. (2013). Antigen-presenting cells in parotid glands contain cystatin D originating from acinar cells. *Archives of Biochemistry and Biophysics, 530*(1), 32–9.

Nashida, T.; Sato, R.; Haga-Tsujimura, M.; Yoshie, S.; Yoshimura, K.; Imai, A.; & Shimomura, H. (2013). Antigen-presenting cells in parotid glands contain cystatin D originating from acinar cells. *Arch. Biochem. Biophys. 530*(1), 32–9.

Navazesh, M. (1993). Methods for collecting saliva. *Annals of the New York Academy of Sciences*, 694, 72–7.

Navazesh, M., & Christensen, C. M. (1982). A comparison of whole mouth resting and stimulated salivary measurement procedures. *Journal of Dental Research, 61*(10), 1158–62.

Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research, 31*(13), 3812–4.

Nikolov, M., Schmidt, C., & Urlaub, H. (2012). Quantitative Mass Spectrometry-Based Proteomics: An Overview. *Methods in Molecular Biology*, 893, 85–100.

Nilsson, C., Lindvall-Axelsson, M., & Owman, C. (1992). Neuroendocrine regulatory mechanisms in the choroid plexus-cerebrospinal fluid system. *Brain Research. Brain Research Reviews, 17*(2), 109–38.

Nishiura, T.; Ishibashi, K.; & Abe, K. (1991). Isolation of three forms of cystatin from submandibular saliva of isoproterenol-treated rats, its properties and kinetic data. *Biochim. Biophys. Acta. 1077*(3), 346–54.

Noben, J.-P., Dumont, D., Kwasnikowska, N., Verhaert, P., Somers, V., Hupperts, R., Stinissen, P., & Robben, J. (2006). Lumbar cerebrospinal fluid proteome in multiple sclerosis: characterization by ultrafiltration, liquid chromatography, and mass spectrometry. *Journal of Proteome Research, 5*(7), 1647–57.

Olafsson, I., & Grubb, A. (2000). Hereditary cystatin C amyloid angiopathy. *Amyloid, 7*(1), 70–9.

Olerup, O., & Hillert, J. (1991). HLA class II-associated genetic susceptibility in multiple sclerosis: a critical evaluation. *Tissue Antigens, 38*(1), 1–15.

Oppenheim, F. G., Salih, E., Siqueira, W. L., Zhang, W., & Helmerhorst, E. J. (2007). Salivary proteome and its genetic polymorphisms. *Annals of the New York Academy of Sciences*, 1098, 22–50.

Oppenheim, F. G., Xu, T., McMillian, F. M., Levitz, S. M., Diamond, R. D., Offner, G. D., & Troxler, R. F. (1988). Histatins, a novel family of histidine-rich proteins in human parotid saliva. *The Journal of Biological Chemistry, 263*(16), 7472–77.

Paulos, C. M., Wrzesinski, C., Kaiser, A., Hinrichs, C. S., Chieppa, M., Cassard, L., Palmer, D. C., Boni, A., Muranski, P., Yu, Z., Gattinoni, L., Antony, P. A., Rosenberg, S. A., & Restifo, N. P. (2007). Microbial translocation augments the function of adoptively transferred self/tumor-specific CD8+ T cells via TLR4 signaling. *The Journal of Clinical Investigation, 117*(8), 2197–204.

Pavlova, A.; & Björk, I. (2003). Grafting of features of Cystatins C or B into the N-Terminal region or second binding loop of Cystatin A (Stefin A) substantially enhances inhibition of cysteine proteinases. *Biochemistry, 42*(38), 11326–33.

Peluso, G., De Santis, M., Inzitari, R., Fanali, C., Cabras, T., Messana, I., Castagnola, M., & Ferraccioli, G. F. (2007). Proteomic study of salivary peptides and proteins in patients with Sjögren's syndrome before and after pilocarpine treatment. *Arthritis and Rheumatism, 56*(7), 2216–2222.

Peng, Y.; Chen, X.; Sato, T.; Rankin, S. A.; Tsuji, R. F.; & Ge, Y. (2012). Purification and high-resolution top-down mass spectrometric characterization of human salivary α-amylase. *Anal. Chem. 84*(7), 3339–46.

Petricoin, E. F., Ardekani, A. M., Hitt, B. A., Levine, P. J., Fusaro, V. A., Steinberg, S. M., Mills, G. B., Simone, C., Fishman, D. A., Kohn, E. C., Liotta, L. A. (2002). Use of proteomic patterns in serum to identify ovarian cancer. *Lancet, 359*(9306), 572–7.

Pisano, E., Cabras, T., Montaldo, C., Piras, V., Inzitari, R., Olmi, C., Castagnola, M., Messana, I. (2005). Peptides of human gingival crevicular fluid determined by HPLC-ESI-MS. *European Journal of Oral Sciences, 113*(6), 462–8.

Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., Fujihara, K., Havrdova, E., Hutchinson, M., Kappos, L., Lublin, F. D., Montalban, X., O'Connor, P., Sandberg-Wollheim, M., Thompson, A. J., Waubant, E., Weinshenker, B., & Wolinsky, J. S. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of Neurology, 69*(2), 292–302.

Popović, T.; Brzin, J.; Ritonja, A.; & Turk, V. (1990). Different forms of human cystatin C. *Biol. Chem. Hoppe-Seyler, 371*(7), 575–80.

Racanelli, V., & Rehermann, B. (2006). The liver as an immunological organ. *Hepatology, 43*(2 Suppl 1), S54-62.

Raj, P. A., Johnsson, M., Levine, M. J., & Nancollas, G. H. (1992). Salivary statherin. Dependence on sequence, charge, hydrogen bonding potency, and helical conformation for adsorption to hydroxyapatite and inhibition of mineralization. *The Journal of Biological Chemistry, 267*(9), 5968–76.

Ramachandran P., Boontheung P., Xie Y., Sondej M., Wong D. T., Loo J. A. (2006) Identification of N-linked glycoproteins in human saliva by glycoprotein capture and mass spectrometry. *Journal of Proteome Research*; 5:1493-503.

Ramasubbu, N., Reddy, M. S., Bergey, E. J., Haraszthy, G. G., Soni, S. D., & Levine, M. J. (1991). Large-scale purification and characterization of the major phosphoproteins and mucins of human submandibular-sublingual saliva. *The Biochemical Journal, 280*(Pt 2), 341–52.

Räsänen, O., Järvinen, M., & Rinne, A. (1978). Localization of the human SH-protease inhibitor in the epidermis. Immunofluorescent studies. *Acta Histochemica, 63*(2), 193–6.

Ravasi, T., Hsu, K., Goyette, J., Schroder, K., Yang, Z., Rahimi, F., Miranda, Les P. A., Paul F., Hume, D. A., & Geczy, C. (2004). Probing the S100 protein family through genomic and functional analysis. *Genomics, 84*(1), 10–22.

Renko, M.; Požgan, U.; Majera, D.; & Turk, D. (2010). Stefin A displaces the occluding loop of cathepsin B only by as much as required to bind to the active site cleft. *FEBS J. 277*(20), 4338–45.

Rifai, N., Gillette, M. A., & Carr, S. A. (2006). Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nature Biotechnology, 24*(8), 971–83.

Rinne, A., Alavaikko, M., Järvinen, M., Martikainen, J., Karttunen, T., & Hopsu-Havu, V. (1983). Demonstration of immunoreactive acid cysteine-proteinase inhibitor in

reticulum cells of lymph node germinal centres. *Virchows Archiv. B, Cell Pathology Including Molecular Pathology, 43*(2), 121–6.

Ruhl, S., Sandberg, A. L., & Cisar, J. O. (2004). Salivary receptors for the proline-rich protein-binding and lectin-like adhesins of oral actinomyces and streptococci. *Journal of Dental Research, 83*(6), 505–10.

Ryan, C. M.; Souda, P.; Halgand, F.; Wong, D. T.; Loo, J. A.; Faull, K. F.; & Whitelegge, J. P. (2010). Confident assignment of intact mass tags to human salivary cystatins using top-down Fourier-transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom. 21*(6), 908–17.

Sabatini, L. M., Carlock, L. R., Johnson, G. W., & Azen, E. A. (1987). cDNA cloning and chromosomal localization (4q11-13) of a gene for statherin, a regulator of calcium in saliva. *American Journal of Human Genetics, 41*(6), 1048–60.

Saitoh, E.; & Isemura, S. (1993). Molecular biology of human salivary cysteine proteinase inhibitors. *Crit. Rev. Oral Biol. Med. 4*(3–4), 487–93.

Sanna, M., Firinu, D., Manconi, P. E., Pisanu, M., Murgia, G., Piras, V., Castagnola, M., Messana, I., del Giacco, S. R., & Cabras, T. (2015). The salivary proteome profile in patients affected by SAPHO syndrome characterized by a top-down RP-HPLC-ESI-MS platform. *Mol. BioSyst., 11*(6), 1552–1562.

Santamaria-Kisiel, L., Rintala-Dempsey, A. C., & Shaw, G. S. (2006). Calcium-dependent and -independent interactions of the S100 protein family. *Biochemical Journal, 396*(2), 201–214.

Schafer, C. A., Schafer, J. J., Yakob, M., Lima, P., Camargo, P., & Wong, D. T. W. (2014). Saliva diagnostics: utilizing oral fluids to determine health status. *Monographs in Oral Science*, 24, 88–98.

Schlesinger, D. H., & Hay, D. I. (1977). Complete covalent structure of statherin, a tyrosine-rich acidic peptide which inhibits calcium phosphate precipitation from human parotid saliva. *The Journal of Biological Chemistry, 252*(5), 1689–95.

Schmidt, C. (2016). Biology: A degenerative affliction. *Nature, 540*(7631), S2–S3.

Schwartz, S. S., Hay, D. I., & Schluckebier, S. K. (1992). Inhibition of calcium phosphate precipitation by human salivary statherin: structure-activity relationships. *Calcified Tissue International, 50*(6), 511–7.

Sedaghat, F., & Notopoulos, A. (2008). S100 protein family and its application in clinical practice. *Hippokratia, 12*(4), 198–204.

Serafini, B., Rosicarelli, B., Franciotta, D., Magliozzi, R., Reynolds, R., Cinque, P., Andreoni, L., Trivedi, P., Salvetti, M., Faggioni, A., & Aloisi, F. (2007). Dysregulated

Epstein-Barr virus infection in the multiple sclerosis brain. *The Journal of Experimental Medicine*, *204*(12), 2899–2912.

Shang, X., Cheng, H., & Zhou, R. (2008). Chromosomal mapping, differential origin and evolution of the S100 gene family. *Genetics, Selection, Evolution*, *40*(4), 449–64.

Shintani, S., Hamakawa, H., Ueyama, Y., Hatori, M., & Toyoshima, T. (2010). Identification of a truncated cystatin SA-I as a saliva biomarker for oral squamous cell carcinoma using the SELDI ProteinChip platform. *International Journal of Oral and Maxillofacial Surgery*, *39*(1), 68–74.

Shomers, J. P., Tabak, L. A., Levine, M. J., Mandel, I. D., & Hay, D. I. (1982). Properties of Cysteine-containing Phosphoproteins from Human Submandibular-sublingual Saliva. *Journal of Dental Research*, *61*(2), 397–399.

Singh, P. K., Jia, H. P., Wiles, K., Hesselberth, J., Liu, L., Conway, B. A., Greenberg, E. P., Valore, E. V., Welsh, M. J., Ganz, T., Tack, B. F., McCray, P. B. (1998). Production of beta-defensins by human airway epithelia. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(25), 14961–6.

Skerget, K., Taler-Vercic, A., Bavdek, A., Hodnik, V., Ceru, S., Tusek-Znidaric, M., Kumm, T., Pitsi, D., Pompe-Novak, M., Palumaa, P., Soriano, S., Kopitar-Jerala, N., Turk, V., Anderluh, G., & Zerovnik, E. (2010). Interaction between oligomers of stefin B and amyloid-beta in vitro and in cells. *The Journal of Biological Chemistry*, *285*(5), 3201–10.

Söderström, K. O., Rinne, R., Hopsu-Havu, V. K., Järvinen, M., & Rinne, A. (1994). Identification of acid cysteine proteinase inhibitor (cystatin A) in the human thymus. *The Anatomical Record*, *240*(1), 115–9.

Strupat, K., Rogniaux, H., Van Dorsselaer, A., Roth, J., & Vogl, T. (2000). Calcium-induced noncovalently linked tetramers of MRP8 and MRP14 are confirmed by electrospray ionization-mass analysis. *Journal of the American Society for Mass Spectrometry*, *11*(9), 780–8.

Stubbs, M. T.; Laber, B.; Bode, W.; Huber, R.; Jerala, R.; Lenarcic, B.; & Turk, V. (1990) The refined 2.4 A 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*(6), 1939–47.

Stubbs, M., Chan, J., Kwan, A., So, J., Barchynsky, U., Rassouli-Rahsti, M., Robinson, R., & Bennick, A. (1998). Encoding of human basic and glycosylated proline-rich proteins by the PRB gene complex and proteolytic processing of their precursor proteins. *Archives of Oral Biology*, *43*(10), 753–70.

Suzuki, T., Hashimoto, S., Toyoda, N., Nagai, S., Yamazaki, N., Dong, H. Y., Sakai, J., Yamashita, T., Nukiwa, T., & Matsushima, K. (2000). Comprehensive gene expression profile of LPS-stimulated human monocytes by SAGE. *Blood*, *96*(7), 2584–91.

Sztáray, J., Memboeuf, A., Drahos, L., & Vékey, K. (2011). Leucine enkephalin--a mass spectrometry standard. *Mass Spectrometry Reviews*, *30*(2), 298–320.

Tahiri, F., Le Naour, F., Huguet, S., Lai-Kuen, R., Samuel, D., Johanet, C., Saubamea, B., Tricottet, V., Duclos-Vallee, J. -C., & Ballot, E. (2008). Identification of plasma membrane autoantigens in autoimmune hepatitis type 1 using a proteomics tool. *Hepatology*, *47*(3), 937–48.

Taubert, R., Hardtke-Wolenski, M., Noyan, F., Wilms, A., Baumann, A. K., Schlue, J., Olek, S., Falk, C. S., Manns, M. P., & Jaeckel, E. (2014). Intrahepatic regulatory T cells in autoimmune hepatitis are associated with treatment response and depleted with current therapies. *Journal of Hepatology*, *61*(5), 1106–14.

Tegner, H. (1978). Quantitation of human granulocyte protease inhibitors in non-purulent bronchial lavage fluids. *Acta Oto-Laryngologica*, *85*(3–4), 282–9.

Terasaki, P. I., Park, M. S., Opelz, G., & Ting, A. (1976). Multiple sclerosis and high incidence of a B lymphocyte antigen. *Science*, *193*(4259), 1245–7.

Thompson, R. C., & Ohlsson, K. (1986). Isolation, properties, and complete amino acid sequence of human secretory leukocyte protease inhibitor, a potent inhibitor of leukocyte elastase. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(18), 6692–6.

Thorey, I. S., Roth, J., Regenbogen, J., Halle, J. P., Bittner, M., Vogl, T., Kaesler, S., Bugnon, P., Reitmaier, B., Durka, S., Graf, A., Wöckner, M., Rieger, N., Konstantinow, A., Wolf, E., Goppelt, A., & Werner, S. (2001). The Ca2+-binding proteins S100A8 and S100A9 are encoded by novel injury-regulated genes. *The Journal of Biological Chemistry*, *276*(38), 35818–25.

Tirumalai, R. S., Chan, K. C., Prieto, D. A., Issaq, H. J., Conrads, T. P., & Veenstra, T. D. (2003). Characterization of the low molecular weight human serum proteome. *Molecular & Cellular Proteomics*, *2*(10), 1096–103.

Troxler, R. F., Offner, G. D., Xu, T., Vanderspek, J. C., & Oppenheim, F. G. (1990). Structural Relationship Between Human Salivary Histatins. *Journal of Dental Research*, *69*(1), 2–6.

Tucci, M., Quatraro, C., & Silvestris, F. (2005). Sjögren's syndrome: an autoimmune disorder with otolaryngological involvement. *Acta Otorhinolaryngologica Italica*, *25*(3), 139–44.

Turk, V., & Bode, W. (1991). The cystatins: Protein inhibitors of cysteine proteinases. *FEBS Letters*, *285*(2), 213–219.

Turk, V., Stoka, V., & Turk, D. (2008). Cystatins: Biochemical and structural properties, and medical relevance. *BioScience, 13*(13), 5406–5420.

Turk, V.; Brzin, J.; Kotnik, M.; Lenarcic, B.; Popović, T.; Ritonja, A.; Trstenjak, M.; Begić-Odobasić, L.; & Machleidt, W. (1986). Human cysteine proteinases and their protein inhibitors stefins, cystatins and kininogens. *Biomed. Biochim. Acta, 45*(11–12), 1375–84.

Uy, R., & Wold, F. (1977). Posttranslational covalent modification of proteins. *Science, 198*(4320), 890–6.

Valore, E. V, & Ganz, T. (1992). Posttranslational processing of defensins in immature human myeloid cells. *Blood, 79*(6), 1538–44.

Vasilopoulos, Y.; Cork, M. J.; Teare, D.; Marinou, I.; Ward, S. J.; Duff, G. W.; & Tazi-Ahnini, R. (2007). A nonsynonymous substitution of cystatin A, a cysteine protease inhibitor of house dust mite protease, leads to decreased mRNA stability and shows a significant association with atopic dermatitis. *Allergy, 62*(5), 514–9.

VerBerkmoes, N. C., Bundy, J. L., Hauser, L., Asano, K. G., Razumovskaya, J., Larimer, F., Hettich, R. L., Stephenson, J. L. (2002). Integrating "top-down" and "bottom-up" mass spectrometric approaches for proteomic analysis of Shewanella oneidensis. *Journal of Proteome Research, 1*(3), 239–52.

Villanueva, J., Philip, J., Chaparro, C. A., Li, Y., Toledo-Crow, R., DeNoyer, L., Fleisher, M., Robbins, R. J., Tempst, P. (2005). Correcting common errors in identifying cancer-specific serum peptide signatures. *Journal of Proteom research, 4*(4), 1060-72.

Vitorino, R., De Morais Guedes, S., Ferreira, R., Lobo, M. J. C., Duarte, J., Ferrer-Correia, A. J., Tomer, K. B., Domingues, P. M., & Amado, F. M. L. (2006). Two-dimensional electrophoresis study of in vitro pellicle formation and dental caries susceptibility. *European Journal of Oral Sciences, 114*(2), 147–153.

Vitzthum, F., Behrens, F., Anderson, N. L., & Shaw, J. H. (2005). Proteomics: from basic research to diagnostic application. A review of requirements & needs. *Journal of Proteome Research, 4*(4), 1086–97.

Vogl, T., Pröpper, C., Hartmann, M., Strey, A., Strupat, K., van den Bos, C., Sorg, C., & Roth, J. (1999). S100A12 is expressed exclusively by granulocytes and acts independently from MRP8 and MRP14. *The Journal of Biological Chemistry, 274*(36), 25291–6.

Wang, Q., Yu, Q., Lin, Q., & Duan, Y. (2015). Emerging salivary biomarkers by mass spectrometry. *Clinica Chimica Acta; International Journal of Clinical Chemistry*, 438, 214–21.

Wang, W., Zhou, H., Lin, H., Roy, S., Shaler, T. A., Hill, L. R., Norton, S., Kumar, P., Anderle, M., Becker, C. H. (2003). Quantification of proteins and metabolites by mass

spectrometry without isotopic labeling or spiked standards. *Analytical Chemistry, 75*(18), 4818–26.

Wang, X., Li, X., Xu, L., Zhan, Y., Yaish-Ohad, S., Erhardt, J. A., Barone, F., & Feuerstein, G. Z. (2003). Up-regulation of secretory leukocyte protease inhibitor (SLPI) in the brain after ischemic stroke: adenoviral expression of SLPI protects brain from ischemic injury. *Molecular Pharmacology, 64*(4), 833–40.

Watson, P. H., Leygue, E. R., & Murphy, L. C. (1998). Psoriasin (S100A7). *The International Journal of Biochemistry & Cell Biology, 30*(5), 567–71.

West, N. R., & Watson, P. H. (2010). S100A7 (psoriasin) is induced by the proinflammatory cytokines oncostatin-M and interleukin-6 in human breast cancer. *Oncogene, 29*(14), 2083–92.

Whitelegge, J. P.; Zabrouskov, V.; Halgand, F.; Souda, P.; Bassilian, S.; Yan, W.; Wolinsky, L.; Loo, J. A.; Wong, D. T. W.; & Faull, K. F. (2007). Protein-sequence polymorphisms and post-translational modifications in proteins from human saliva using top-down Fourier-transform ion cyclotron resonance mass spectrometry. *Int. J. Mass Spectrom. 268*(2–3), 190–7.

Wirths, O.; Breyhan, H.; Cynis, H.; Schilling, S.; Demuth, H. U.; & Bayer, T. A. (2009). Intraneuronal pyroglutamate-Abeta 3-42 triggers neurodegeneration and lethal neurological deficits in a transgenic mouse model. *Acta Neuropathol. 118*(4), 487–96.

Xiao, H., Zhang, L., Zhou, H., Lee, J. M., Garon, E. B., & Wong, D. T. W. (2012). Proteomic analysis of human saliva from lung cancer patients using two-dimensional difference gel electrophoresis and mass spectrometry. *Molecular & Cellular Proteomics, 11*(2), M111.012112.

Xu, T., Levitz, S. M., Diamond, R. D., & Oppenheim, F. G. (1991). Anticandidal activity of major human salivary histatins. *Infection and Immunity, 59*(8), 2549–54.

Yan and Bennick, 1995, Identification of histatins as tannin-binding proteins in human saliva. *Biochem J, 311*(Pt 1):341-347.

Yang, F., Tay, K. H., Dong, L., Thorne, R. F., Jiang, C. C., Yang, E., Tseng, H. Y., Liu, H., Christopherson, R., Hersey, P., & Zhang, X. D. (2010). Cystatin B inhibition of TRAIL-induced apoptosis is associated with the protection of FLIP(L) from degradation by the E3 ligase itch in human melanoma cells. *Cell Death and Differentiation, 17*(8), 1354–67.

Zavasnik-Bergant, T. (2008). Cystatin protease inhibitors and immune functions. *Frontiers in Bioscience: A Journal and Virtual Library*, 13, 4625–37.

Zhang, A., Sun, H., Wang, P., & Wang, X. (2013). Salivary proteomics in biomedical research. Clinica Chimica Acta, 415, 261–265.

Zhang, Z., & Marshall, A. G. (1998). A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra. *Journal of the American Society for Mass Spectrometry, 9*(3), 225–233.

Zimmer, D. B., Wright Sadosky, P., & Weber, D. J. (2003). Molecular mechanisms of S100-target protein interactions. *Microscopy Research and Technique, 60*(6), 552–559.

Publications on peer-reviewed international journals

J Sep Sci. 2016 May;39(10):1987-97. doi: 10.1002/jssc.201501306. Epub 2016 Apr 29.

N- and O-linked glycosylation site profiling of the human basic salivary proline-rich protein 3M.

Manconi B[1], Cabras T[1], Sanna M[1], Piras V[1], Liori B[1], Pisano E[2], Iavarone F[3], Vincenzoni F[3], Cordaro M[4], Faa G[2], Castagnola M[3,5], Messana I[5].

Author information

Abstract

In the present study, we show that the heterogeneous mixture of glycoforms of the basic salivary proline-rich protein 3M, encoded by PRB3-M locus, is a major component of the acidic soluble fraction of human whole saliva in the first years of life. Reversed-phase high-performance liquid chromatography with high-resolution electrospray ionization mass spectrometry analysis of the intact proteoforms before and after N-deglycosylation with Peptide-N-Glycosidase F and tandem mass spectrometry sequencing of peptides obtained after Endoproteinase GluC digestion allowed the structural characterization of the peptide backbone and identification of N- and O-glycosylation sites. The heterogeneous mixture of the proteoforms derives from the combination of 8 different neutral and sialylated glycans O-linked to Threonine 50, and 33 different glycans N-linked to Asparagine residues at positions 66, 87, 108, 129, 150, 171, 192, and 213.

KEYWORDS: Mass Spectrometry; N-Deglycosylation; Saliva; Site-specific glycosylation

PMID: 26991339 DOI: 10.1002/jssc.201501306

Arch Oral Biol. 2017 May;77:68-74. doi: 10.1016/j.archoralbio.2017.01.021. Epub 2017 Jan 31.

Top-down HPLC-ESI-MS proteomic analysis of saliva of edentulous subjects evidenced high levels of cystatin A, cystatin B and SPRR3.

Manconi B[1], Liori B[2], Cabras T[2], Iavarone F[3], Manni A[4], Messana I[5], Castagnola M[5], Olianas A[2].

Author information

Abstract

OBJECTIVE: This study aims to analyze the salivary peptidome/proteome of edentulous subject with respect to dentate control subjects.

DESIGN: Unstimulated whole saliva, collected from 11 edentulous subjects (age 60-76 years) and 11 dentate age-matched control subjects, was immediately treated with 0.2% aqueous trifluoroacetic acid and the acidic soluble fraction analyzed by High Performace Liquid Chromatography-Mass Spectrometry. The relative abundance of the salivary peptides/proteins was determined by measuring the area of the High Performace Liquid Chromatography-Mass Spectrometry eXtracted Ion Current peaks which is linearly proportional to peptide/protein concentration under identical experimental conditions. Levels of salivary peptides/proteins in the two groups were compared by the nonparametric Mann-Whitney test to evidence statistically significant differences.

RESULTS: Levels of cystatin A, S-glutathionylated, S-cystenylated, S-S dimer derivatives of cystatin B and S-glutathionylated derivative of SPRR3, were found significantly higher in edentulous subjects with respect to dentate controls. The major peptides and proteins typically deriving from salivary glands did not show any statistically significant differences.

CONCLUSIONS: Cystatin A, S-glutathionylated, S-cystenylated, S-S dimer derivatives of cystatin B and S-glutathionylated derivative of SPRR3, which are mainly of intracellular origin and represent the major constituents of the cornified cell envelope are a clue of inflammation of mucosal epithelia.

KEYWORDS: Cystatins; Edentulous subjects; Keratinocytes; Mass-Spectrometry; Proteomics; Saliva

PMID: 28178587 DOI: 10.1016/j.archoralbio.2017.01.021

J Proteome Res. 2017 Nov 3;16(11):4196-4207. doi: 10.1021/acs.jproteome.7b00567.

Salivary Cystatins: Exploring New Post-Translational Modifications and Polymorphisms by Top-Down High-Resolution Mass Spectrometry.

Manconi B[1], Liori B[1], Cabras T[1], Vincenzoni F[2], Iavarone F[2], Castagnola M[2,3], Messana I[2,3], Olianas A[1].

Author information

Abstract

Cystatins are a complex family of cysteine peptidase inhibitors. In the present study, various proteoforms of cystatin A, cystatin B, cystatin S, cystatin SN, and cystatin SA were detected in the acid-soluble fraction of human saliva and characterized by a top-down HPLC-ESI-MS approach. Proteoforms of cystatin D were also detected and characterized by an integrated top-down and bottom-up strategy. The proteoforms derive from coding sequence polymorphisms and post-translational modifications, in particular, phosphorylation, N-terminal processing, and oxidation. This study increases the current knowledge of salivary cystatin proteoforms and provides the basis to evaluate possible qualitative/quantitative variations of these proteoforms in different pathological states and reveal new potential salivary biomarkers of disease. Data are available via ProteomeXchange with identifier PXD007170.

KEYWORDS: S-type cystatins; cystatin A; cystatin B; cystatin D; human saliva; post-translational modifications; top-down high-resolution mass spectrometry

PMID: 29019242 DOI: 10.1021/acs.jproteome.7b00567

Acknowledgment

I would like to thank Prof. Eleonora Cocco and Dr. Lorena Lorefice (Center of Multiple Sclerosis, Department of Medical Sciences and Public Health, University of Cagliari) and Prof. Luchino Chessa and Dr. Simona Onali (Center for the Study of Liver Diseases, Department of Medical Sciences "M. Aresu", University of Cagliari) for the samples collection.