


Università degli Studi di Cagliari

PhD DEGREE

in

Life, Environmental and Drug Sciences
Biomedical Curriculum

XXXII Cycle

Bottom-up **proteomics: investigating the potential of the salivary proteome as a source of Schizophrenia biomarkers, and exploration of the protein composition of the oral mucosal pellicle**

BIO/10

PhD Student: Doctor Gianluigi Cabiddu
PhD Coordinator: Professor Simona Distinto
PhD Supervisor: Professor Tiziana Cabras

Final exam academic year 2018 – 2019

Thesis defence: January – February 2020

Contents

Abstract

The work presented in this thesis describes the *bottom-up* mass spectrometry-based techniques which have been applied to saliva samples in order to identify potential biomarkers of Schizophrenia and to characterize the protein composition of salivary film adhering to oral surface. Bidimensional electrophoretic analysis of 17 saliva samples from adult schizophrenic subjects and 17 healthy subjects, combined with mass spectrometry allowed the identification of different glycolytic enzymes, such as phosphoglycerate kinase 1, glyceraldehyde-3-phosphate dehydrogenase and alpha enolase, with a reduced concentration levels in schizophrenic patients. The altered levels of the glycolytic enzymes may lead to an imbalance in the pathways involved in energy production and in a massive oxidative stress that appears to be a central feature in schizophrenia. In addition, also annexin A2 showed a reduced level in schizophrenic patients and its possible role as a ROS scavenging molecule, due to the presence of a -Cys reactive residue in its structure, may be relevant in schizophrenia since this disease is often associated with a high oxidative stress. Protein S100 A12, often found increased in the serum of patients suffering from neurodegenerative or inflammatory diseases, displayed an increased level in our analysis of schizophrenic saliva samples, in accordance with the data previously obtained by the *top-down* proteomic analysis on the acid-soluble fraction of saliva. Interleukin-36-α, a potent activator of various chemokines and cytokines such as IL-6, often present at higher concentration in the serum of schizophrenic patients, showed increased levels in our analysis of schizophrenic saliva samples.

Oral mucosal pellicle is a thin layer of salivary proteins lining epithelial oral cells. This layer is involved in oral health by protection from bacterial colonization and maintaining lubrication and it also participates to taste perception.

In terms of protein composition, current data based on experiments *in vivo* highlight the presence of specific salivary components such as S-type Cystatins, Mucins (5B and 7), CAVI and secretory IgA. However, to our knowledge, the mucosal pellicle composition has not been thoroughly described using a proteomic approach. To perform such a characterization, we used a cell model based on TR146 cells that are suitable as a model of oral epithelium. This cell line was stably transfected in order to express MUC1, which can improve MUC5B adhesion on buccal cells. This model presents the advantage of forming the mucosal pellicle if exposed to human saliva. The aim of our work was therefore to isolate and characterize the protein composition of the mucosal pellicle formed *in vitro* on both TR146 and TR146/MUC1 cells, and to compare by nano-LC MS/MS the mucosal pellicle proteome with the salivary proteome. In this work we present a suitable method for the in vitro isolation of the mucosal pellicle by "shaving" it from the cells using trypsin, which also detaches the cells from the culture plate but does not induce cellular lysis, allowing the subsequent separation of cells and the pellicle. Our analyses on pellicles proteome confirmed the presence of specific salivary proteins in the pellicle like MUC7, but also revealed the adsorption of other proteins onto cells such as BPI fold-containing family B member 1, which was not previously reported in the pellicle composition. To conclude, this thesis presents the proteomic mass spectrometry analysis applied to the study of human saliva which demonstrates the versatility of mass spectrometry and has highlighted areas of clinical medicine and oral health where proteomics and a personalized biomedical approach could be further investigated.

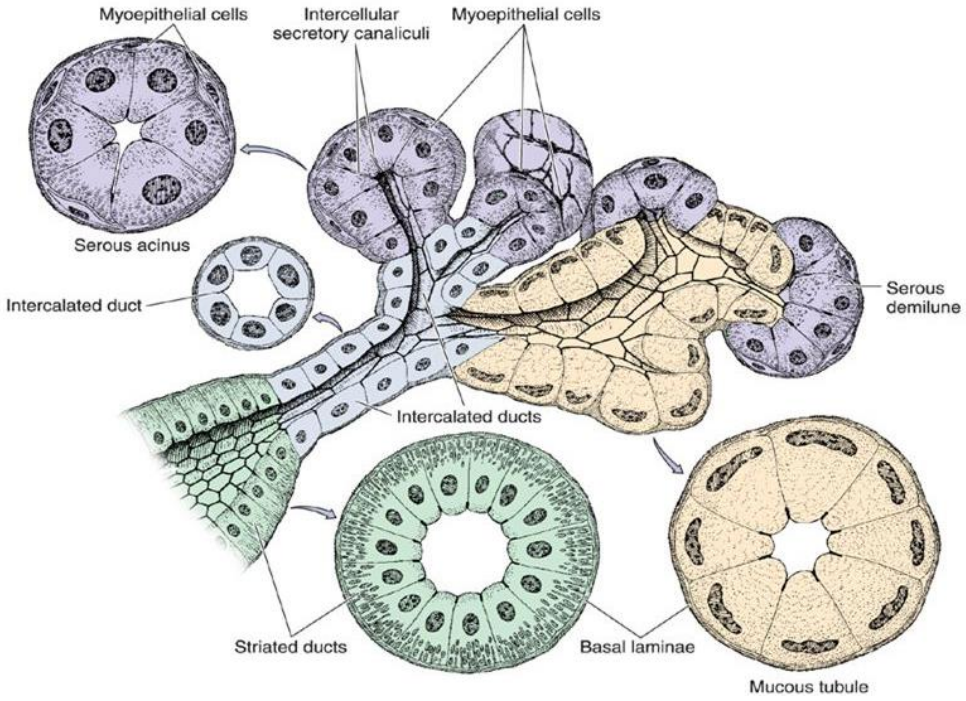
Introduction

1 Human saliva

Saliva can be defined as the mixed glandular secretion present in the oral cavity. It is constituted by the secretions of three paired major glands: parotid, submandibular and sublingual. Saliva also contains secretions of minor salivary glands found throughout the oral cavity (lower lip, tongue, palate, cheeks and upper pharynx). The type of salivary secretion is different based on the gland: parotid secretions are more serous or watery while submandibular and sublingual secretions are more viscous due to the high content of glycoproteins. Saliva formation occurs at the end pieces of glands (acini) where serous cells produce a watery seromucous secretion and mucous cells produce a viscous mucin-rich secretion. These secretions are consequence of the formation of interstitial fluid from blood in capillaries, which is then modified by the acini in the cells. This fluid is secreted into the lumen, from which it passes to the gland ducts where it is further modified. In the striated duct of the glands, ion exchange process changes this secretion from isotonic to hypotonic. The composition of saliva undergoes to a final modification in the excretory ducts before its secretion in the mouth [1] (Figure 1). We can refer at this fluid secreted by glands specifically as "saliva" while the often-used term "whole saliva" indicates the fluid containing saliva but also contribution of gingival crevicular fluid, mucosal transudation, bronchial and nasal secretion, serum and blood derivatives, desquamated epithelial cells, bacterial products and food residues [2]. However, it is important to distinguish between "free flowing saliva" and saliva adhering to mucosal surfaces. Free flowing saliva is swallowed and leaves a residual film on the mouth surface, whose thickness varies between 70 and 100 μm [3].

Saliva adhering to mucosal surface, i.e. the mucosal pellicle, is constituted by a small part of the salivary proteins which are selectively adsorbed onto the surface of mucosal epithelial cells, and has a thickness that can reach up to 100 nm [4]. Moreover, salivary residual film is mobile while the mucosal pellicle is strongly attached to oral cells.

Figure 1. *Structural organization of the salivary glands* [5].



Saliva is involved in oral health with various functions such as lubrication and mucosal homeostasis, antimicrobial activity, buffer capacity and teeth remineralisation, food processing and digestion, taste perception and phonation. Beside the presence of ions and water these functions are carried out mostly by proteins present in saliva which can be specific of the oral cavity or common to other organs and body fluids. Salivary proteins specific of the oral cavity are mainly secreted by glands (about 400) and constitute approximately 90% of saliva. They are: PRPs (acidic, basic and glycosylated), α-amylase, carbonic anhydrase 6, immunoglobulins (IgA and IgG), salivary mucins (MUC5B and MUC7), type II cystatins (S, SA, SN, C and D), histatins, statherins and P-B peptide. Non-secretory proteins (more than 2000) [6] derive from oral cells, blood exudates, gingival crevicular fluid and from the oral flora, and include: thymosins, defensins, type I cystatins (A and B), S100 (A7, A8, A9, A11, A12) proteins, polymeric immunoglobulin receptor and antileukoproteinase.

1.1 Proline-rich proteins (PRPs)

Proline-rich proteins (PRPs) constitute the majority of salivary proteins and are divided into acidic (aPRPs), basic (bPRPs) and glycosylated (gPRPs). Proteins of this family are characterized by high numbers of variants and PTMs like glycosylation, phosphorylation and proteolytic cleavage [7].

1.1.1 Acidic PRPs (aPRPs)

aPRPs are secreted mostly by parotid gland (70%) and in minor part by submandibular and sublingual glands (30%) [8]. There are five members of aPRPs: PIF-s, Db-s, Pa are codified by PRH1 gene and PRP1 and PRP2 are codified by PRH2. The acidic character of these proteins is due to the presence of several aspartic and glutamic acid residues in the first 30 amino acids at the N-terminus.

All five members of the group are phosphorylated at Ser_8 and Ser_{22} and present a pyro-glutamic residue at the N-terminus. In whole saliva it is possible to detect also minor amounts of mono, three and non-phosphorylated proteoforms [9]. All aPRPs, except for Pa, are post-translationally cleaved after Arg_{106} (Arg_{127} in Db-s) by a convertase activity generating PRP3, PRP4, PIF-f, Db-f and the C-terminal fragment PC peptide, common to all the proteoforms. Pa protein presents a cysteine residue instead of an arginine at position 103, making this protein resistant to convertase activity and as result it is usually present in saliva as a disulfide-dimer. aPRPs are inhibitors of calcium phosphate crystal growth. This activity resides in the 30 charged amino acids at the N-terminus containing the two phosphoserine residues, which are essential for the role of these proteins in mineral homeostasis [10].

1.1.2 Basic (bPRPs) and Glycosylated PRPs (gPRPs)

Basic PRPs are secreted by parotid glands and the genes codifying these proteins are PRB1, PRB2, PRB3, PRB4 and PBII (SMR3B). PRB1 and PRB3 contain four alleles named small (S), medium (M), large (L) and very large (VL), while PRB2 and PRB4 contain S, M and L alleles [11]. bPRPs are expressed as pre-proproteins that after signal peptide removal are completely cleaved before glandular secretion. PRB1 generates the following products: II-2 peptide (from S, M and 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 (L allele) derive IB-1, P-J, P-H, P-F peptides and IB-8a protein which due to a single nucleotide polymorphism Ser_{100}→Pro exists in two species [12]. The presence of Ser_{100} in the proteoform named IB-8a Con1+ (with Con standing for Concanavalin A-binding), generates a glycosylation sequon in which Asn68 in N-linked to different types of glycans [13].

On the contrary, the presence of Pro_{100} residue abolishes the recognition sequon leading to the formation of the non-glycosylated IB-8a Con1- proteoform. Also, PRB3 and PRB4 genes codify both for glycosylated proteins. PRB3, in particular the M allele, codifies for PRP3-M protein whose N- an O-glycosylation sites have been identified recently [14]. PRB4 gene codifies for PRP4 protein which is cleaved by an endoprotease into three chains: protein N1, glycosylated protein A and P-D peptide. Glycosylation gives these proteins lubrication properties similar to the other high molecular weight proteins present in saliva [15], which is a protective feature in oral cavity against abrasion and helps the speech. Moreover, gPRPs interact with dietary tannins leading to a lower astringency perception [16] and are also able to bind oral bacteria, in particular *Fusobactrium nucleatum* that colonizes the oral cavity early in life [17]. In fact, it has been demonstrated that gPRPs are the only PRPs expressed early in babies [18] probably as a first line of protection against bacterial infection in the mouth, facilitating the removal and clearance of bacteria.

1.2 α-amylase

α-amylase is one of the major salivary components secreted by both parotid and minor salivary glands, with a high number of PTMs. It is known that about 25% of α-amylase is found as glycosylated [19]. The main function of α-amylase is during digestion by hydrolyzing (1→4)-alpha-D-glucosidic linkages in polysaccharides such as starch. However, the impact of this protein on digestion is controversial, since it has a short active contact time with starch in the mouth, and once the food bolus is swallowed, salivary α-amylase is inactivated by the low gastric pH [20]. α-amylase is thought to be involved in oral perception of starches, pre-absorptive metabolic signaling, and plasma glucose responses to ingested starch, and it is correlated to development of metabolic syndrome [20].

1.3 Carbonic Anhydrase 6 (CAVI)

Carbonic Anhydrases are enzymes that catalyze the reaction $CO_2 + H_2O \rightleftharpoons HCO_3^- + H^+$. CAVI encoded by CA6 gene is secreted by parotid and submandibular glands [21]. The exact role of salivary CAVI is still unclear, but studies demonstrated that CAVI is associated to the enamel pellicle and protects it by converting salivary bicarbonate and microbe-delivered hydrogen ions to carbon dioxide and water [22]. Moreover, salivary gustin, a protein associated to taste buds, has been identified as CAVI in the work by Thatcher et al. [23].

1.4 Immunoglobulins

Immunoglobulins (Igs) are antigen recognition molecules encoded by genes located in three major loci: 14q32.33, 2p11.2 and 22q11.2 [24]. Igs are constituted by two identical heavy chains and two identical light chains (kappa or lambda), joined by disulfide bridges that bind a heavy chain to a light chain and the two heavy chains together. Both heavy and light chains have an N-terminal variable region which defines the antigen-binding site and a C-terminal constant region that has effector or signaling functions. Igs constitute the first defense against antigens in the mucosal layer on the epithelial surfaces and in the acquired enamel on the tooth, with secretory IgA and IgG predominant in saliva [25–28]. Secretory IgA deriving from plasma cells in salivary glands and mucosal cells, exists mainly in dimeric form even though in some cases it is found in polymeric form, while salivary IgG is monomeric and derives from serum and passes in saliva through gingival crevices [28]. In human saliva is possible to find also very small amount of other Igs like IgM, IgD and IgE [26].

1.5 Mucins

Mainly secreted by submandibular glands, mucins are highly glycosylated proteins present in the oral cavity. The most common are MUC5B, MUC7, MUC19 which are secreted, and membrane associated MUC4 and MUC1. MUC5B and MUC7 however, are the predominant in saliva [29]. MUC7 and MUC5B can bind, by their glycan moiety, salivary proteins with antimicrobial properties increasing their concentration and persistence in the oral cavity and modulating their antimicrobial activity. They are also able to bind to several bacterial strains through sialic acid residues or proteins expressed on bacterial surface [29].

1.6 Cystatins

Cystatins are a family of proteins whose main function in saliva is to inhibit cysteine proteases [30]. Type 1 cystatins expressed by CSTA and CSTB genes include Cystatin A and Cystatin B also known as stefins, while type 2 cystatins including Cystatin C, D, E, F, S, SN and SA are expressed by a gene family of eight to nine members. The third family of cystatins includes low-molecular-weight kininogen, high-molecular-weight kininogen and T-kininogens which are single-chain glycoproteins with multiple disulfide bonds found in plasma and secretions [30]. Cystatins A and B which are of intracellular origin, are expressed in several cells and tissues and are inhibitors of cathepsin L, S and H. They are single chain proteins of about 11 kDa that lack of signal peptide and have N-terminal acetylation [31]. In saliva Cystatin B is detected mainly as S-derivatives named Cys_3 S-glutathionylation, S-cysteinylation, S-S 2-mer [32] and S-CMC (Carboxymethyl) [33]. Type 2 Cystatins are extracellular proteins and include five S-types (S, S1, S2, SA and SN) and Cystatins C and D. They are proteins ranging from 13 to 14 kDa and present the signal peptide and disulfide bridges [34]. Among these cystatins only Cystatin S is phosphorylated on Ser_1 (S1) or on Ser_1 and Ser_3 (S2).

Cystatin C, expressed in almost the whole body, is a single chain protein containing four conserved cysteine residues that can form two disulfide bridges [31]. Cystatin C can form amyloid deposit in brain arteries of young people, leading to cerebral hemorrhage [35].

Cystatin D can be detected in two natural forms with Cys or Arg at position 26 of the chain due to a gene polymorphism [36]. Beside its classical role of protease inhibitor, Cystatin D can inhibit proliferation, invasion and migration of colon carcinoma [37], regulate the antigen presenting cell activity [38] and modulate gene expression [39]. Internalization process in antigen-presenting cells originated from parotid glands explains the presence of Cystatin D in other body fluids than saliva [38].

1.7 Histatins

Histatins are low molecular weight peptides whose source are parotid and submandibular glands [40]. All Histatins derive from Histatin 1 and Histatin 3 encoded by HIS1 and HIS2 genes [41]. Histatin 1 is phosphorylated at Ser_2 in the Ser_2-Asp_3-Glu_4 consensus sequence and partially sulfated on Tyr residues. In Histatin 3, Glu_4 is substituted by Ala_4, preventing the phosphorylation of Ser_2 [42]. Before secretion Histatin 3 is cleaved, generating Histatin 6, Histatin 5 and other fragments [43]. Histatins show antifungal and antibacterial properties, due to a domain of residues 12-25 in Histatin 3, which is present also in most of the histatin fragments [44]. The phosphate group present in Histatin 1 is not involved in antifungal activities but has a role in interaction with the teeth enamel minerals [45].

1.8 Statherin and P-B peptide

Statherin is a tyrosine-rich phospho-peptide secreted by parotid and submandibular glands, codified by STATH gene. Statherin can be present as di-phosphorylated on Ser_2 and Ser_3, mono-phosphorylated and non-phosphorylated [8].

Statherin can form a cyclo-structure through the formation of an intramolecular bridge between Lys_6 and Gln_{37} mediated by the action of transglutaminase 2 [46]. Statherin is involved in the formation of enamel pellicle on the teeth, especially the cyclo-statherin [46] and is also involved in the prevention of calcium phosphate precipitation, due to the presence of acidic residues (two phosphoserines, aspartic and glutamic acid) in the first six N-terminal amino acids [47,48].

P-B peptide, whose specific functions are still not known, is expressed by PROL3 gene that is very close to the statherin gene. P-B peptide is secreted by parotid, submandibular and sublingual glands [8].

1.9 α-defensins

α-defensins (HNP, human neutrophil peptides) are basic peptides rich in tyrosine and cysteine residues. DEFA 1, DEFA3 and DEFA4 genes codify respectively for α-defensin 1, 3 and 4 while α-defensin 2 originates from a proteolytic cleavage of the N-terminus of α-defensin 1 or 3 [49]. The oral source of these peptides in saliva is the gingival crevicular fluid [50] and in particular neutrophils. Defensins are involved in antimicrobial activities, regulation of cell volume, cytokine production, chemotaxis and inhibition of natural killer cells [51–53]. Moreover, α-defensin 4 also called corticostatin, is involved in inflammation by inhibiting the production of cortisol [54].

1.10 β-thymosins

β-thymosins are peptides codified by TMSB4X gene which produce thymosins β4 (Tβ4) and β4 oxidized and by TMSB10 gene generating thymosin β10 (Tβ10). As described for defensins also thymosins originate from the gingival crevicular fluid [55].

Their biological role is associated to actin polymerization, induction of metalloproteinases, chemotaxis, angiogenesis, inhibition of inflammation and bone marrow stem cell proliferation and are also related to cancer and metastasis [56–58].

1.11 S100 family

S100 family comprises 25 small proteins distinguished in three subfamilies: S100A (A7, A8, A9, A11, A12), S100B and S100P [59]. They have no intrinsic catalytic activity, but after binding calcium they can modulate the functions of other proteins. They can participate in intracellular and extracellular function such as calcium homeostasis, cytoskeletal rearrangement, contraction and motility, cell growth and differentiation, arachidonic acid transport, chemotaxis, apoptosis, promotion of wound repair, protection against microbial activity, control of ROS formation, inflammation and protein phosphorylation and secretion [60–65]. In saliva only S100A subfamily is detected [18].

S100A7 also called *psoriasin* is present in saliva in two isoforms of which the D_{27} variant is the most represented. After the loss of the initial methionine residue both S100A7 variants undergo to N-terminal acetylation.

S100A8 is the member of salivary S100A proteins mainly prone to extensive oxidation and collectively these different oxidized proteoforms are named hyper-oxidized S100A8 [66]. The oxidation may occur at Met_1, Met_{78}, Trp_{54} and Cys_{42} residues, the latter being oxidized also to sulfonic acid (S100A8-SO_3H). Oxidation may occur also at Trp_{54} (S100A8-$SO_3H/W_{54}ox$), at Trp_{54} and Met_{78} (S100A8-$SO_3H/W_{54}ox/M_{78}ox$). S100A8 has also been found *in vivo*, glutathionylated on Cys_{42} residue (S100A8-SSG) and through its Cys_{42} residue can form a dimer with S100A9 protein (S100A8/A9-SSdimer) [66].

S100A9 also called calgranulin B, has been characterized in four isoforms in saliva [18]. Two isoforms called long-type differ from each other for a phosphorylation on Thr_{112} and are N-terminal acetylated. The other two isoforms, called short-types, lack five N-terminal amino acid residues (MTCKM) and differ for the phosphorylation on Thr as well as long-types.

S100A11 (calgizzarin) is found in the cornified envelope of keratinocytes and is N-terminal acetylated after methionine loss [18]. S100A12 is found as an homodimer but depending on Ca^{2+} and Zn^{2+} concentrations, this protein can form oligomers [67].

1.12 Antileukoproteinase

Antileukoproteinase is a protein associated to innate immunity expressed by SLPI gene. This protein is produced by neutrophils, macrophages, β cells, renal tubules epithelial cells, acinar cells of parotid and submandibular glands, acinar cells of submucosal glands and epithelial cells lining mucosa of respiratory and gastrointestinal tract [68–71]. Antileukoproteinase is involved in inhibition of proteases, elastase, trypsin and chymotrypsin, chymase and tryptase [72,73]. Moreover, this protein is able to increase proliferation of adult NSC and to stimulate oligodendroglia differentiation resulting in the promotion of tissue repair [74].

1.13 Polymeric Immunoglobulin Receptor (pIgR)

pIgR is a type 1 transmembrane glycoprotein involved in mucosal immunity by transporting polymeric IgA across mucosal epithelial cells [75,76]. It is regulated by pro-inflammatory cytokines, hormones and microbial factors through activation of toll-like receptors 3 and 4 [76]. After a proteolytic cleavage in the glycosylated extracellular region of pIgR, the secretory component (SC) is released. SC has been detected in saliva and seems to be involved in the functions of the oral mucosal pellicle since it is found bound to oral cells [77,78].

2 Human saliva in proteomic research

Proteomics is a science that studies the quali/quantitative variation of the proteome which can be defined as the set of protein expressed by an organism, a tissue or a biofluid at a given time. Qualitative analysis aims to identify all the proteins and their PTMs present in a sample while quantitative analysis can be relative or absolute. Relative quantification allows to compare the concentrations of selected proteins or whole proteome between samples, while absolute quantification provides the exact concentration of a protein in each sample. Proteomes can be studied by means of antibodies and other affinity molecules (affinity proteomics) or by various proteomic platforms divided mainly in *top-down* or *bottom-up* platforms which rely mainly on mass spectrometry analysis [79]. *Top-down* strategy relies on detection of the intact proteins without any proteolytic cleavage. The advantage of this approach is that it provides information about proteins PTMs, but on the contrary, sample preparation for subsequent mass spectrometry analysis requires acidification which leads to precipitation of some proteins, moreover high molecular weight or glycosylated proteins are difficult to detect. On the other hand, *bottom-up* approach (Figure 2) focuses on selected proteins identification and characterization of their amino acid sequence by digesting proteins with enzymes. This approach needs pre-purification step such as electrophoresis or liquid chromatography. When *bottom-up* is performed on a mixture of proteins it is called shotgun proteomics. The drawback of *bottom-up* strategy is that some fragments of digested proteins can be too small for detection and PTMs information can be lost due to sample treatment. *Top-down* and *bottom-up* approaches can be successfully integrated during the proteomic analysis of a given sample resulting in the compensation of the drawbacks of each single approach.

The availability of mass spectrometry techniques has improved research in the salivary proteomic field and has produced qualitative and quantitative information on protein composition of saliva both in physiological and pathological conditions. Saliva is a very attractive body fluid for disease diagnosis since its collection is usually economical, safe, easy and it is considered an acceptable and non-invasive process for patients because it does not provoke any pain.

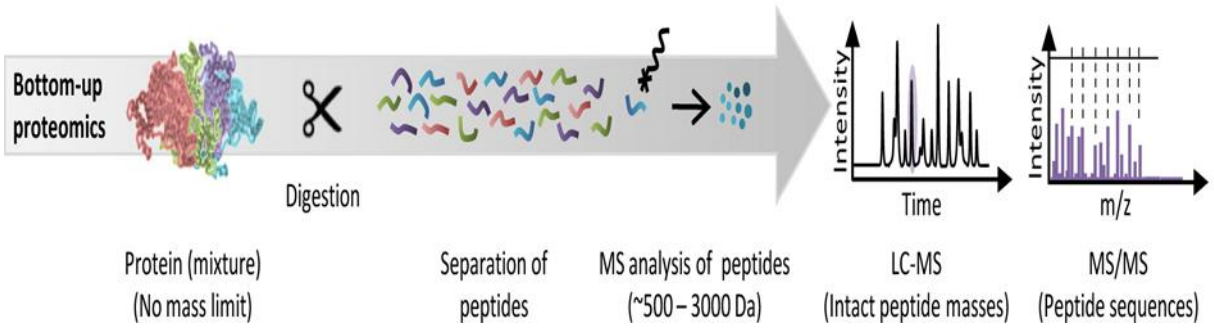
Recent proteomic platforms have analysed the human salivary proteome characterizing about 3000 differentially expressed proteins and peptides, most of them showing an age related variation in their level starting from 180 days of post-conceptional age to adulthood [80,81]. For example, bPRPs expression levels reach their maximum during adolescence, probably as result of growth and hormonal maturation that involves also salivary glands [82]. Due to its mixed origin, from glandular secretion and diffusion from blood, human saliva and its proteome can reflect the physio-pathological state of the whole organism not only confined to oral cavity, thus it can be potentially used a diagnostic tool for disease biomarkers discovery [83]. For instance, a mass spectrometry based *top-down* proteomic approach was recently used to analyse the acid-soluble fraction of saliva from subjects suffering of multiple sclerosis [84] leading to the identification of several differentially expressed proteins related to inflammation and disrupted immune response typical of this disease. As said above *top-down* and *bottom-up* approaches can be combined to have a better characterization of a sample proteome. An integrated *top-down* and *bottom-up* approach based on mass spectrometry has been applied for the proteomic study of saliva collected from subjects affected by Wilson disease [66]. This analysis highlighted different levels of expression of proteins related to inflammation and oxidative stress, which reflected the oxidative stress and inflammatory conditions characteristic of the pathology.

Although the use of gel-free technologies is rapidly growing, two-dimensional electrophoresis coupled to mass spectrometry (2DE-MS) remains the most popular and versatile procedure for proteome analysis. A validated *bottom-up* approach relies on 2DE as separative technique and MS for the identification of proteins in a sample. 2DE is a useful technique for the separation of a great number of proteins present in a complex sample. Since proteins are separated in specific spots on a gel, in-gel digestion of protein with specific endoproteases such as trypsin, give protein fingerprints which can be analysed further by MS. The protein spot can be then identified by comparing the mass spectrometric peptide map with that theoretically calculated in a database. This technique however, is not free from limitations because membrane proteins are usually poorly separated [85] and a high number of biological replicates is needed to achieve significant statistical differences. Nonetheless, several studies used *bottom-up* 2DE-MS based approaches to investigate the salivary proteome[86–88] revealing the presence of proteins, also with a high-molecular-weight, that cannot be detected by using a *top-down* approach alone. Proteomic analyses based on 2DE coupled to MS have been also used to highlight differences between groups of subjects under different pathophysiological conditions. For example a comparison study between smokers and non-smokers [89] revealed a different 2D profile with three proteins (lipocalin-1, thioredoxin and interleukin-1 receptor antagonist) up-regulated in the group of smokers subjects. Another work by the same authors [90] investigated the saliva of patients suffering of oral squamous carcinoma, evidencing a unique profile of salivary proteins different from healthy subjects profile. A 2DE analysis on saliva of diabetes type 2 patients [91] evidenced high levels of Cystatin SA related to periodontal disease severity occurring in diabetes.

A recent work investigated the acid-insoluble fraction of saliva from human preterm new-borns by using 2DE coupled to high resolution mass spectrometry [92]. This work integrated the results obtained on the acid-soluble fraction of saliva collected from the same subjects [18] since it allowed the identification of proteins not detectable in the soluble fraction due to sample treatment before *top-down* mass spectrometry. The results discussed suggest that in the near future human saliva will be a relevant diagnostic fluid for clinical diagnosis and prognosis.

Nowadays, many studies including the ones reported above, are focused on salivary biomarker search, but there are no validated biomarkers used in clinical practice, due to some limitations that can influence saliva sampling conditions like salivary flow rate, oral hygiene and diet. The present work is composed of two distinct parts. In the first part, by using a 2DE based *bottom-up* approach to investigate the salivary proteome of schizophrenic patients compared to salivary proteome of healthy subjects, the aim is to highlight differences in protein expression levels that can be of potential diagnostic use. The second part of the work focuses on exploration of the protein composition of the oral mucosal pellicle, which nowadays is still not fully characterized, by using a high-resolution MS based *bottom-up* approach.

Figure 2. *Example of a bottom-up approach workflow in which proteins present in a mixture are digested and the resulting peptides are analysed by MS* [93].



PART 1:

PROTEOMIC ANALYSIS OF THE ACID-INSOLUBLE FRACTION OF SALIVA FROM SCHIZOPHRENIC PATIENTS

Maybe each human being lives in a unique world, a private world different from those inhabited and experienced by all other humans… If reality differs from person to person, can we speak of reality singular, or shouldn't we really be talking about plural realities? And if there are plural realities, are some more true (more real) than others? What about the world of a schizophrenic? Maybe it's as real as our world. Maybe we cannot say that we are in touch with reality and he is not, but should instead say, his reality is so different from ours that he can't explain his to us, and we can't explain ours to him. The problem, then, is that if subjective worlds are experienced too differently, there occurs a breakdown in communication… And there is real illness.

Philip K. Dick

1 Schizophrenia

Schizophrenia (SCZ) is a chronic and debilitating mental disease affecting more than 21 million people worldwide and it develops between late adolescence and early adulthood. It affects men and women in equal manner, but men develop the illness earlier. The diagnosis is difficult and SCZ is often confused with other mental disorders such as depression or bipolarism. This illness presents various symptoms which are divided into “positives” or “negatives”. Symptoms are defined positives when they add behaviours that were not previously present in the patient. Examples of positive symptoms are hallucinations, delusions and other episodes that exaggerate and distort the normal behaviour of the patient. On the other hand, negative symptoms indicate a loss of some functions such as sense of pleasure and sociability. During this phase the patient tends to isolate himself and to have a poor personal care. This disorder is generally characterized by different phases. The first one or “prodromal” can last for days or months and determines a change in the patient that becomes apathetic and isolates himself from the community. Then, an acute phase follows, characterized by clear positive and negative symptoms that are uncontrolled without treatment. Later, a chronic phase takes place, during which positive symptoms are less present and negative symptoms remain. In this phase acute exacerbations can occur. In the end there is a residual phase with negative symptoms still present and a poor efficacy of the therapy. During this phase there could be also a remission of the illness.

SCZ aetiology is complex and there is no factor that is predominant over others. Genetic predisposition along environmental factors is decisive in the onset of the pathology. Environmental factors include also stressful events occurred during childhood. From a biological point of view there are different theories about the pathophysiology of SCZ.

One of them involves the dopaminergic neurotransmission. It has been shown that drugs effective against SCZ symptoms, decrease dopaminergic neurotransmission. Decreasing the dopaminergic transmission helps patients to manage their behaviour and to improve their perceptual abilities [94]. The morphology of the brain is also affected by this disease: it has been shown that the number of neurons is diminished and also some areas of the brain have a smaller volume respect to a healthy condition [94]. Recently researchers focused on the study of inflammatory process as a central feature in SCZ pathogenesis. Inflammatory molecules are often detected with increased levels in the serum or cerebrospinal fluid of SCZ patients. Inflammation seems also relevant in altering dopaminergic, serotonergic, noradrenergic and glutamatergic transmission. It can have also a neurotoxic effect resulting in the loss of brain volume described above. These assumptions are partially validated by the fact that anti-inflammatory drugs can help in alleviating SCZ symptoms [95].

Regarding SCZ treatment, antipsychotic drugs are divided into typical and atypical. Typical antipsychotic (i.e. chlorpromazine and haloperidol) are very effective against positive symptoms of SCZ due to the block of dopamine receptors D2. Despite their effectiveness these drugs have severe side effects such as extrapyramidal disorders. Atypical antipsychotic such as clozapine and risperidone are effective against negative symptoms because of the block of dopamine receptors D1 and D4, but they block also serotonin receptors 5-TH_2.

Their side effects can include moderate to severe gain weight, agranulocytosis and myocarditis. However the extrapyramidal disorders are less common because these drugs have lower affinity for D2 receptors [94]. Moreover, all the drugs used in SCZ treatment are only beneficial to a subset of symptoms, and only 40% of the patients responds to medication. Thus, it would be necessary to better characterize this pathology to improve treatment and life quality of people affected.

1.1 The proteome of Schizophrenia

Proteomic research in schizophrenia has started with the analysis of *post-mortem* brain tissues considering the nature of this disease. Later, research has been expanded towards other types of samples adding more data to those obtained by analyses on the brain alone. A review by Nascimento et al. [96], resumed all proteomic analyses done in various tissues such as brain tissue, cerebrospinal fluid, serum and saliva, highlighting the principal proteins and their pathways involved in SCZ.

Proteins related to neuronal transmission, synaptic plasticity and neurite outgrowth were found downregulated in SCZ patients. These differentially expressed proteins include NMDA receptors and glutamatergic signalling molecules, proving the imbalance in neuronal transmission and signal pathways in the brain of SCZ patients. Other proteins found downregulated in SCZ are related to calcium homeostasis such as calmodulin, calmodulin-like proteins and S100 family members. Calcium is a central metabolite for the dopaminergic involvement in SCZ, since its role in the function of D1 and D2 receptors for dopamine.

Another pathway dysregulated in SCZ regards the energy metabolism. Proteins involved in this pathway, especially glycolytic enzymes, were found up regulated or down regulated in SCZ patients. These findings are confirmed by observation of hyperglycaemia, insulin resistance and impaired glucose tolerance in patients. This imbalance in energy metabolism along with elevated calcium concentrations leads to a massive oxidative stress confirmed by the presence of differentially expressed proteins related to ROS scavenging and indeed oxidative stress seems to be a central feature in the pathophysiology of SCZ. In SCZ proteomics analyses also proteins related to cytoskeleton were found differentially expressed and their abnormal regulation may influence neuronal plasticity, vesicle trafficking and cytoskeletal arrangement.

Several proteins involved in inflammation and immune system were found differentially expressed in SCZ patients. Among these proteins the most relevant were interleukins, defensins and S100 A12 since inflammation is a central process in SCZ, proteins of this pathway can be possible biomarker candidates of the disease.

Recently Ivarone et al. [97] published a work focused on the analysis, by a *top-down* approach, of the acid soluble fraction obtained after acidification of saliva collected from SCZ and bipolar patients. In this work defensins, cystatin A and B and S100 A12 were found overexpressed in SCZ patients. All these proteins are related to inflammation and immune system regulation, confirming the results explained in the above-mentioned review. This work proved also that saliva can be use as diagnostic fluid since it can reflect physiological changes in all the organism and has the advantage of simple collection compared to other fluids such as blood.

2 Aim of the work

Thanks to the collaboration with the research group of the Catholic University in Rome, we were able to obtain the acid-insoluble fraction of saliva collected from the same patients enrolled in the study performed by Iavarone et al. [97]. Therefore, the aim of this work was to analyse by 2D electrophoresis coupled to high resolution mass spectrometry, the acid-insoluble fraction of saliva to better characterize the salivary proteome of SCZ by integrating the results obtained on the acid-soluble fraction.

3 Materials and methods

3.1 Sample collection, treatment and subjects enrolled

Resting whole saliva (from 0.2 to 1.0 mL) was collected from 17 SCZ adult patients enrolled in the study by Iavarone et al. [97], and 17 adult age/sex matched healthy subjects with a soft plastic aspirator. Samples were collected between 10.00 and 12.00 a.m. and the patients did

not drink or eat 30 min before the collection. The samples were mixed in 1:1 ratio with 0.2% 2,2,2 trifluoroacetic acid (v/v; TFA) in an ice bath. The samples were then centrifuged at 20000 g for 5 min, and the acid soluble and insoluble fractions obtained were separated and stored at -80°C until the analysis. The study protocol and written consent were approved by the Medical Ethics Committee of the Faculty of Medicine of the Catholic University of Rome. The informed consent procedures are consistent with the latest stipulations established by the Declaration of Helsinki.

3.2 Acid-insoluble fraction treatment

Each pellet containing the insoluble protein fraction collected after addition of TFA was pounded with a plastic pestle and emulsified with 1.5 mL of 10% TCA and 20 mM DTT in acetone, in a 1.5 mL microtube and incubated overnight at -20°C to allow protein precipitation. After centrifuging at 20000 g for 30 min at 4°C, pellets were submitted to three washing cycles with 1 mL of 20 mM DTT in cold acetone, dried under the hood for about 3 hours, and then stored at -80°C until the analysis. Before 2DE analysis, pellets were neutralized with the addition of 5 µL of 0.1 M NaOH and vortexed for 2 min. Then, pellets were resuspended in 230 µL of Destreak Rehydration Solution (GE Healthcare) and incubated for 3 hours at room temperature, with occasional vortexing. Samples were centrifuged at 20000 g for 15 min at 4°C in order to remove insoluble materials, such as mucins or cell debris and the clear supernatant was transferred in a clean microtube for determination of total protein concentration.

3.3 Protein quantification and passive rehydration

Total protein concentration of acid insoluble fraction of saliva obtained after Destreak Rehydration of the pellets, was determined using the 2D Quant Kit (GE Healthcare). Quantification followed the manufacturer's instructions ad was performed in duplicate using 15 µL of each sample. Then samples were applied on an IPG strip to allow overnight rehydration of the strip.

3.4 First dimension: IEF

2DE analysis was performed on 11.0 cm IPG strips containing a 3–10 nonlinear immobilized pH gradient (BioRad, Hercules, CA, USA). Onto each IPG strip were loaded 50 µg of total proteins for analytical gel analysis or 270 µg for preparative gel analysis. The electrical conditions applied consisted of four steps, for a total time of about 5 h, setting the maximum current applied to 50 µA. For the first step, voltage was set at 500 V for 90 min or until a total of 500 V/h was reached. The second and the third steps were carried out in gradient mode from 500 V to 1000 V in 1 h and from 1000 V to 6000 V in 2 h, respectively. For the last step, voltage was set at 6000 V for 30 min or until a total of 11000 V/h was reached.

3.5 Equilibration of IPG strips and second dimension: SDS-PAGE

Before running the second dimension, each IPG strip was equilibrated in 6 M urea, 30% (v/v) glycerol, 2% (w/v) SDS, 50 mM Tris-HCl pH 8.8, 0.01% (w/v) bromophenol blue in presence of 1% (w/v) DTT for 15 min. Then, a second equilibration step with the same buffer but in the presence of 2.5% (w/v) iodoacetamide (IAM) instead of DTT was performed for 15 min again. SDS-PAGE was performed using Tris-HCl precast Criterion gels (T= 10-20%) on a Criterion Cell (Bio-Rad). The run was performed applying 15 mA for 15 min and then 150 V until the exit of bromophenol blue from the lower end of the gel.

Molecular masses were determined by running low range standard protein markers (SDS-PAGE Molecular Weight Standards, Bio-Rad Laboratories). Analytical gels were stained with the fluorescent dye SYPRO Ruby (Molecular Probes, Eugene, OR, USA), while preparative gels were stained with colloidal Coomassie Brilliant Blue (EZBlue Gel Staining Reagent, Sigma-Aldrich, St. Louis, MO), following the manufacturer's instructions.

3.6 Image analysis

2DE gel images (both analytical and preparative) were acquired using the ChemiDoc MP imaging system (Bio-Rad). Images, after quality check (saturation, dimension, background), were analysed by SameSpots software version 4.6, (TotalLab, Newcastle upon Tyne, United Kingdom), processed for spot detection and normalization using the automated tools of the software, considering statistically significant the differences of at least 1,5 fold and ANOVA p values ≤0.05. All spots were manually reviewed to ensure proper detection and matching.

To evaluate the results from 2DE images obtained by SameSpots Multivariate Stats module, various statistical tests according to variances and distributions of the normalized spot volumes (calculated by SameSpots) were performed using Prism Graph Pad Software.

3.7 In-gel tryptic digestion

Differentially expressed protein spots were manually excised from preparative gels after EZ Blue gel staining and *in-gel* digestion was performed using Trypsin Singles Proteomics Grade Kit (Sigma-Aldrich) according to the manufacturer’s instruction. Tryptic digests were freeze-dried, solubilised in 0.1% formic acid (FA) and submitted to HPLC-ESI-high resolution MS/MS experiments for protein identification.

3.8 RP-HPLC-high resolution MS/MS characterization of tryptic peptides

HPLC-ESI-high resolution MS/MS experiments were carried out by an Ultimate 3000 RSLC nano system coupled to an Orbitrap ELITE apparatus (Thermo Fisher Scientific, Waltham, MA, USA). Zorbax 300 SB-C18 (3.5 μm particle diameter; column dimension 1 mm ×150 mm) (Agilent Technologies, Santa Clara, CA) was used as chromatographic column. The following eluents were used: (A) 0.1% (v/v) aqueous FA and (B) 0.1% (v/v) FA in ACN/water 80/20 v/v. The applied gradient was: 0–2 min 5% B, 2–40 min from 5 to 70% B (linear), 40–45 min from 70 to 99% B (linear), at a flow rate of 50 μL/min with a total time of 65 min. for run. MS spectra were collected with 120000 resolution and m/z range from 350 to 2000. In data-dependent acquisition mode the five most intense multiply charged ions were selected and fragmented in ion trap by using CID 35% normalized collision energy. Capillary temperature was 300°C and source voltage 4.0 kV.

3.9 MS data analysis

MS/MS data were elaborated by Proteome Discoverer software (version 1.4.1.14, Thermo Fisher Scientific), based on SEQUEST HT cluster as search engine against UniProtKB/Swiss-Prot Homo Sapiens database (released on 3rd of September 2014, homo sapiens 26132 entries). The search parameters were 10 ppm tolerance for precursor ions and 0.5 Da for product ions, 2 missed cleavage, carbamidomethylation of cysteine (+57.02 Da) as fixed modification, oxidation of methionine (+15.99 Da) and serine, threonine and tyrosine phosphorylation (+79.96 Da) as variable modification, minimum two identified peptides and on a decoy database search calculated false discovery rate under 5%. Protein characterisation was set with the identification of a minimum of two peptides per protein and two unique peptides by applying the high confidence filter.

4 Results

Figure 3 shows the 2DE PAGE of the acid insoluble fraction of saliva obtained from patients and healthy subjects. Proteins are separated in the gel during the first dimension through a non-linear pH gradient from 3 to 10, while in the second dimension proteins are separated according to their molecular weight in a range between 97.4 kDa and 14.4 kDa. A large number of spot present in the bidimensional map of saliva were already characterized by *bottom-up* experiments based on mass spectrometry performed by our research group on healthy subjects and subjects suffering from pathologies different from SCZ [92,98]. The upper part of the gel (97.4 – 45 kDa) contains high-molecular-weight proteins such as serum albumin, polymeric immunoglobulin receptor and amylase. In the middle part (45 kDa – 21.5 kDa) we found proteins like glycolytic enzymes, annexins and immunoglobulins. The lower part of the gel (21.5 kDa – 14.4 kDa) is characterized by the presence of small proteins like prolactin-inducible proteins, S100A protein family and cystatins.

Gel images obtained after 2DE analyses were processed with SameSpots software. This software allows the comparison between the gels analysed evidencing the differentially expressed protein spots under the conditions studied. The comparison consists of two principal steps: 1) gel alignment, when the software chooses a gel as the reference image, and automatically aligns all the gels onto the reference one. The alignment can be also checked manually to eventually correct errors. When the gels are aligned onto each other, the software proceeds to spot detection between the conditions (patients and healthy subjects); 2) spot detection and review, where the software calculates the normalized spot volume (as an intensity) of each spot present in all gels and then it calculates the mean normalized volume for a same spot detected in all the gels.

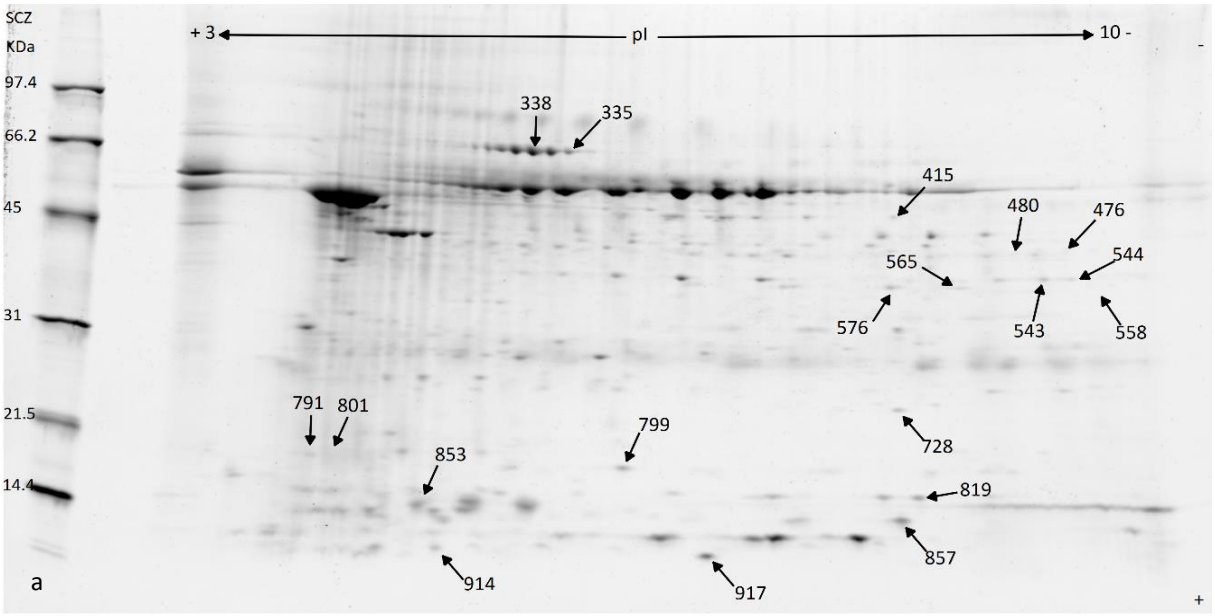
The mean normalized spot volume (intensity) is used as the parameter by which the software reports if a specific spot is differentially expressed between patients and healthy subjects. This step has to be checked manually, by setting the filters needed for that analysis (anova p value ≤ 0.05 and fold variation ≥ 1.5).

By aligning the 2DE PAGE of 17 SCZ adult patients and 17 adult healthy subjects a total of 19 protein spots were reported as differentially expressed. Among these, 12 spots were previously identified by 2DE-MS performed on the acid insoluble fraction of saliva from other pathologies investigated by our research group [92,98] (Table 1). Some of the identified proteins, are present in more than one spot, probably because of the presence of different PTMs such as glycosylation, phosphorylation, and/or endogenous proteolytic cleavages and they are: Serum albumin (spots 335 and 338), Glyceraldehyde-3-phosphate dehydrogenase (spots 543 and 544), Annexin A2 (spots 558, 565) and Prolactin-inducible protein (spots 791 and 801). 3 out of these 12 spots (labelled with asterisk in Table 1) showed an ambiguous protein identification in every analysis performed; these ambiguous attributions were most likely due to the very similar isoelectric point and molecular weight of the involved proteins and are: spot 543 (Glyceraldehyde-3-phosphate dehydrogenase and L-lactate dehydrogenase A chain) spot 558 (Keratin type I cytoskeletal 9, Keratin type II cytoskeletal 1, Keratin type I cytoskeletal 10, Malate dehydrogenase mitochondrial, Annexin A2) and spot 565 (Annexin A2 and Glyceraldehyde-3-phosphate dehydrogenase). The remaining other 4 spots were univocally identified, and they are: alpha-enolase (spot 415), cystatin SN (spot 819); S100 A9 short P (spot 853) and cystatin B (spot 857).

The remaining 7 spots, which were not previously identified (Table 2 and Figure 4), were excised from the gel and digested with trypsin for protein identification with high resolution mass spectrometry. Proteome discoverer elaboration identified these proteins: Phosphoglycerate kinase 1 (spots 476 and 480); Annexin A2 (spot 576); Alpha-crystallin B chain (spot 728); Interleukin-36 alpha (spot 799); S100 A12 (spot 917). Unfortunately, we were not able to provide protein identification for spot 914. Figure 5 reports the identification of Annexin A2 in spot 576 obtained by Proteome discoverer elaboration. Panel A shows the protein sequence coverage obtained by matching the tryptic peptides generated during in gel digestion with trypsin of the spot with a database containing theoretical peptides derived from trypsin cleavage of Annexin A2. Panel B displays the MS/MS spectrum generated by b^+ and y^+ series ions of the tryptic peptide AEDGSVIDYELIDQDAR.

Considering only unambiguous identifications, proteins were reported to be over-expressed or under-expressed by SameSpots analysis as reported in both Tables 1 and 2. Increased levels of Prolactin-inducible protein, Interleukin-36 alpha and S100 A12 were found in the acid insoluble fraction of saliva from SCZ patients compared to healthy subjects. Serum albumin, Alpha enolase, Phosphoglycerate kinase 1, Glyceraldehyde-3-phosphate dehydrogenase, Annexin A2, Alpha-crytsallin B chain, Cystatin SN, Cystatin B and S100 A9 short P showed instead decreased levels. These results were further strengthened by performing the statistical analysis on the normalized spot volumes, obtained by SameSpots, by Prism Graph Pad Software, which confirmed the statistical significance for all the protein spots previously identified except for cystatin SN (Figure 6).

Figure 3. *Representative 2DE PAGE of acid-insoluble fraction of saliva from schizophrenic subjects (a) and healty controls (b). First dimension was performed on a 11 cm IPG strip containing an immobilized pH gradient 3–10 NL, second dimension on a precast Tris-HCl Criterion Gel (T = 10-20%). Gel was stained with Sypro Ruby fluorescent stain. Labels refer to the differentially expressed protein spots detected by SameSpots software.*



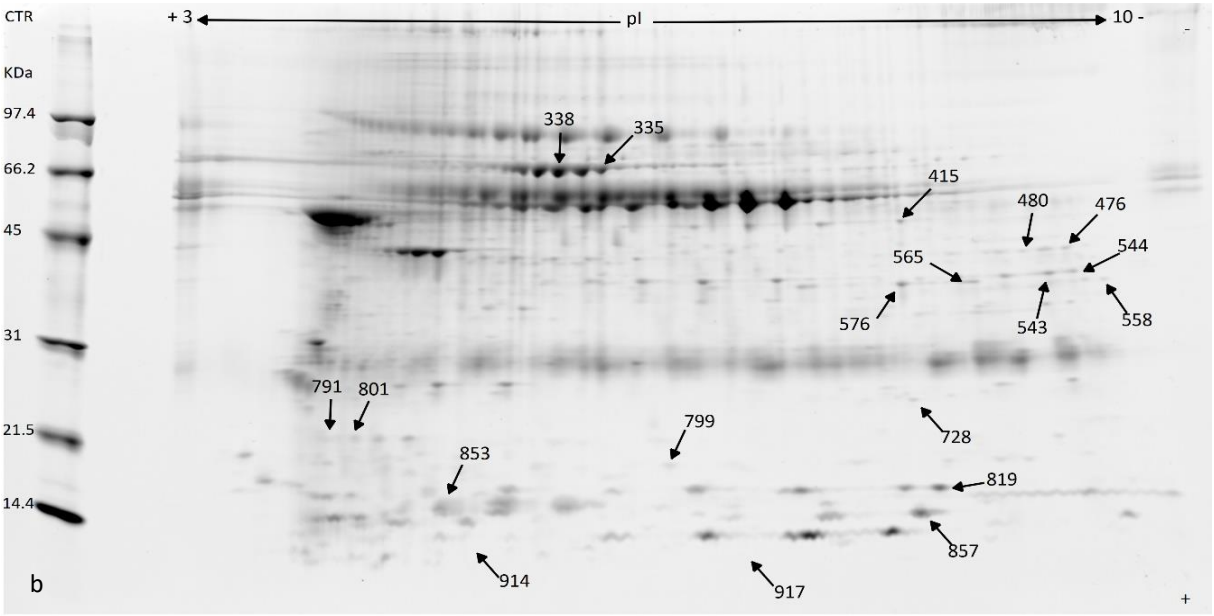


***Table 1**. List of differentially expressed protein spots found in the acid-insoluble fraction of saliva from schizophrenic patients, that were already identified in previous 2DE analysis performed by our research group. * indicates protein spots that have more than one protein identified. Statistical analysis between SCZ and CTR was performed by SameSpots software.*

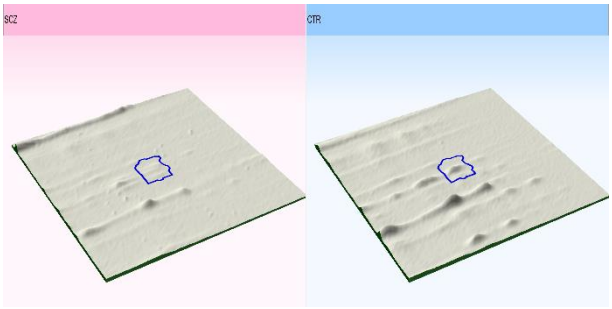
	Proteome Discoverer Data						Uniprot Data		SameSpot data		
spot	accession #	Description	Score	Coverage	# unique peptides	# peptides	MW (kDa)	pI	Anova (p)	Fold variation	Highest Mean
335	P02768	Serum albumin	241,8	42,2%	25	25	66,4	5,67	0,02	1,56	CTR
338	P02768	Serum albumin	245,1	42,5%	25	25	66,4	5,67	0,007	1,5	CTR
415	P06733	Alpha-enolase	661,3	50,2%	12	16	47,0	7,0	0,01	1,5	CTR
543*	P04406	Glyceraldehyde-3-phosphate dehydrogenase	32,5	21,2%	5	5	35,9	8,6	0,01	1,5	CTR
	P00338	L-lactate dehydrogenase A chain	20	15,6%	5	5	36,5	8,5			
544	P04406	Glyceraldehyde-3-phosphate dehydrogenase	15	12,8%	3	3	35,9	8,6	0,01	1,63	CTR
558*	P35527	Keratin, type I cytoskeletal 9	203,9	28,4%	7	7	62	5,14	0,01	1,5	CTR
	P04264	Keratin, type II cytoskeletal 1	127	26,4%	8	9	65,8	8,15			
	P13645	Keratin, type I cytoskeletal 10	92	27%	8	8	58,8	5,13			
	P40926	Malate dehydrogenase, mitochondrial	55,9	33,7%	7	7	32,9	8,54			
	P07355	Annexin A2	39,95	10%	3	3	38,5	7,56			
565*	P07355	Annexin A2	20,8	16,5%	5	5	38,4	7,6	2,35E-03	1,54	CTR
	P04406	Glyceraldehyde-3-phosphate dehydrogenase	8	13,7%	3	3	35,9	8,6			
791	P12273	Prolactin-inducible protein	461,1	63%	9	9	13,5	5,4	0,02	1,73	SCZ
801	P12273	Prolactin-inducible protein	1079,4	63%	10	10	13,5	5,4	0,003	1,90	SCZ
819	P01037	Cystatin SN	71.5	61%	6	8	14.3	6.92	0,04	1,74	CTR
853	P06702	S100 A9 Short P	160,3	79%	9	9	13,1	5,71	0,03	1,5	CTR
857	P04080	Cystatin-B	57,6	55%	5	5	11,1	7,56	7,52E-05	1,81	CTR

***Table 2**. List of differentially expressed protein spots that were identified after in gel trypsin digestion and MS/MS. Statistical analysis between SCZ and CTR was performed by SameSpots software.*

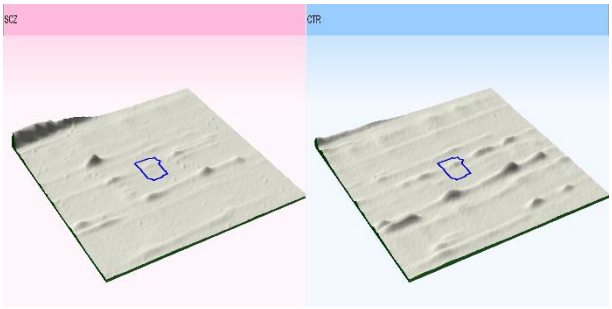
	Proteome Discoverer Data						Uniprot Data		SameSpot data		
spot number	accession #	Description	Score	Coverage	# unique peptides	# peptides	MW (kDa)	pI	Anova (p)	Fold variation	Highest Mean
476	P00558	Phosphoglycerate kinase 1	210,4	69,3%	25	25	44,4	8,3	0,003	1,73	CTR
480	P00558	Phosphoglycerate kinase 1	116	58,7%	19	19	44,4	8,3	0,0002	1,54	CTR
576	P07355	Annexin A2	279,1	71,3%	23	23	38,4	7,6	0,0002	1,63	CTR
728	P02511	Alpha-crystallin B chain	140,4	81%	13	13	20,1	7,33	0,001	1,77	CTR
799	Q9UHA7	Interleukin-36 alpha	167,6	64,5%	9	9	17,1	5,89	0,01	2,05	SCZ
914	-	NI	-	-	-	-	-	-	0,03	2,24	SCZ
917	P80511	Protein S100 A12	31,6	56,5%	3	3	10,4	5,8	8,14E-05	5,32	SCZ

Figure 4. *3D representation of the differentially expressed protein spots between SCZ and CTR groups, identified by MS/MS analysis after in gel trypsin digestion. For each protein the pink square refers to 2DE gels section of SCZ patients, while blue square refers to healthy subjects.*

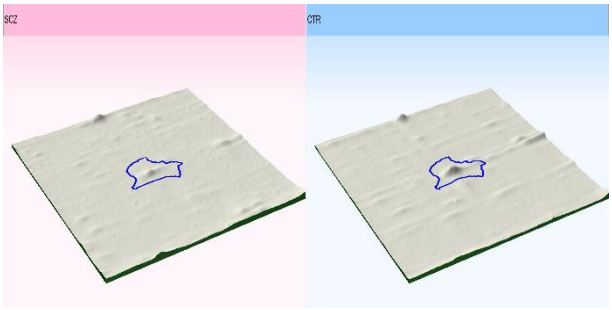
Phosphoglycerate kinase 1 (spot 476)



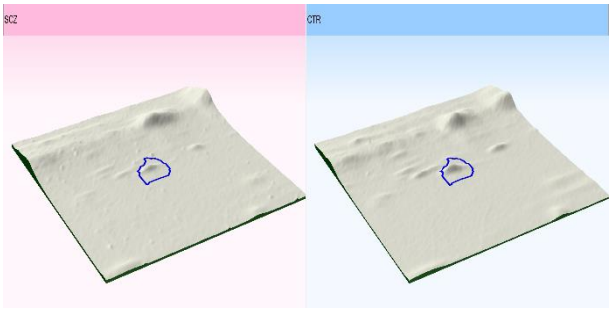
Phosphoglycerate kinase 1 (spot 480)



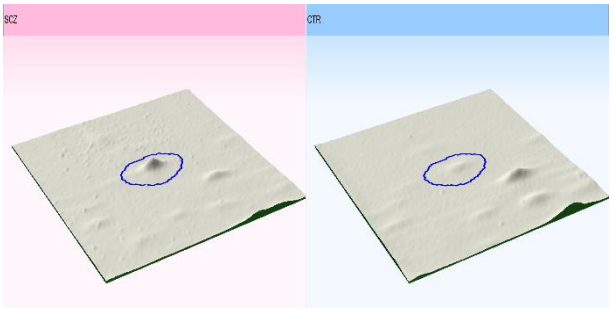
Annexin A2 (spot 576)



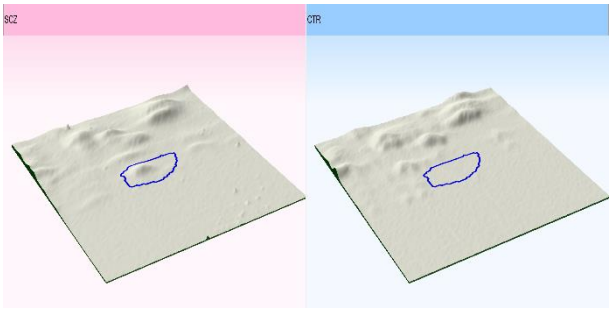
Alpha-crystallin B chain (spot 728)



Interleukin-36 alpha (spot 799)



NI (spot 914)



Protein S100 A12 (spot 917)

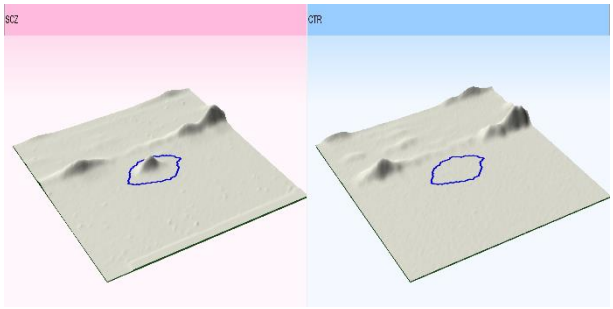
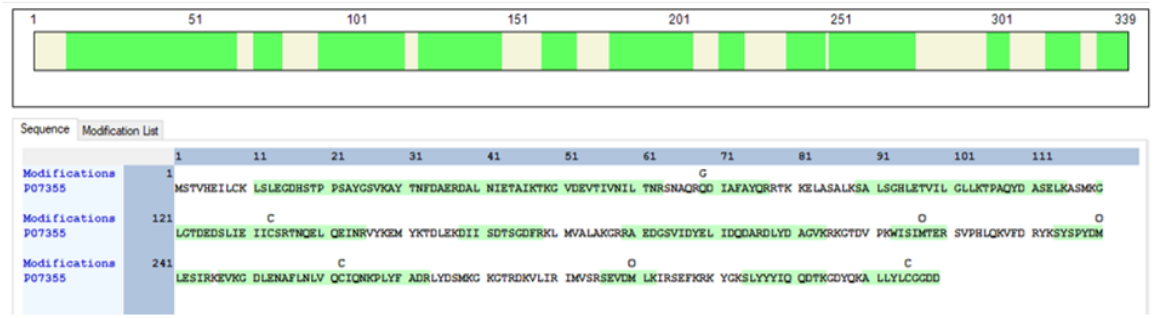
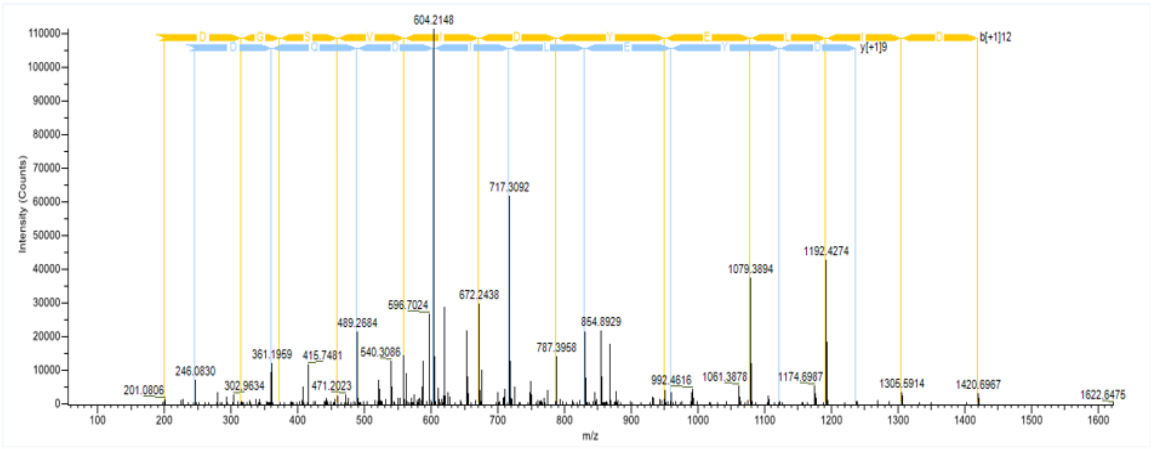


Figure 5. *Protein sequence coverage (A) and MS/MS spectrum of the tryptic peptide AEDGSVIDYELIDQDAR (B) obtained by Proteome discoverer elaboration for Annexin A2 in spot 576.*

A



B



A E⌋ D⌋ G⌋ S⌋ V⌋ I ⌡D⌡Y⌡E⌡L⌡I ⌡D⌡Q⌜D⌜A R

Figure 6. *Spot volume distributions of differentially expressed proteins between SCZ patients and healthy control subjects.*

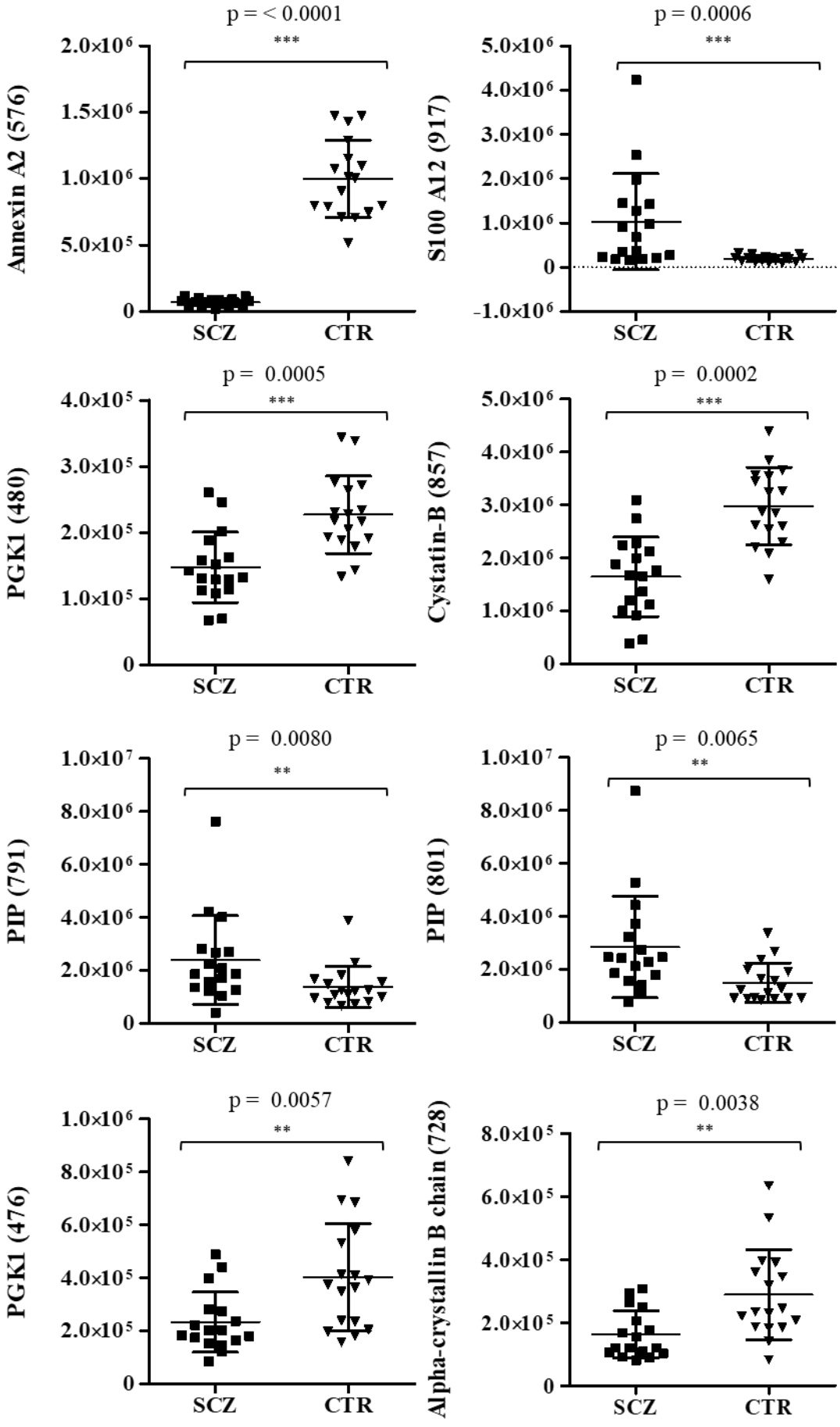
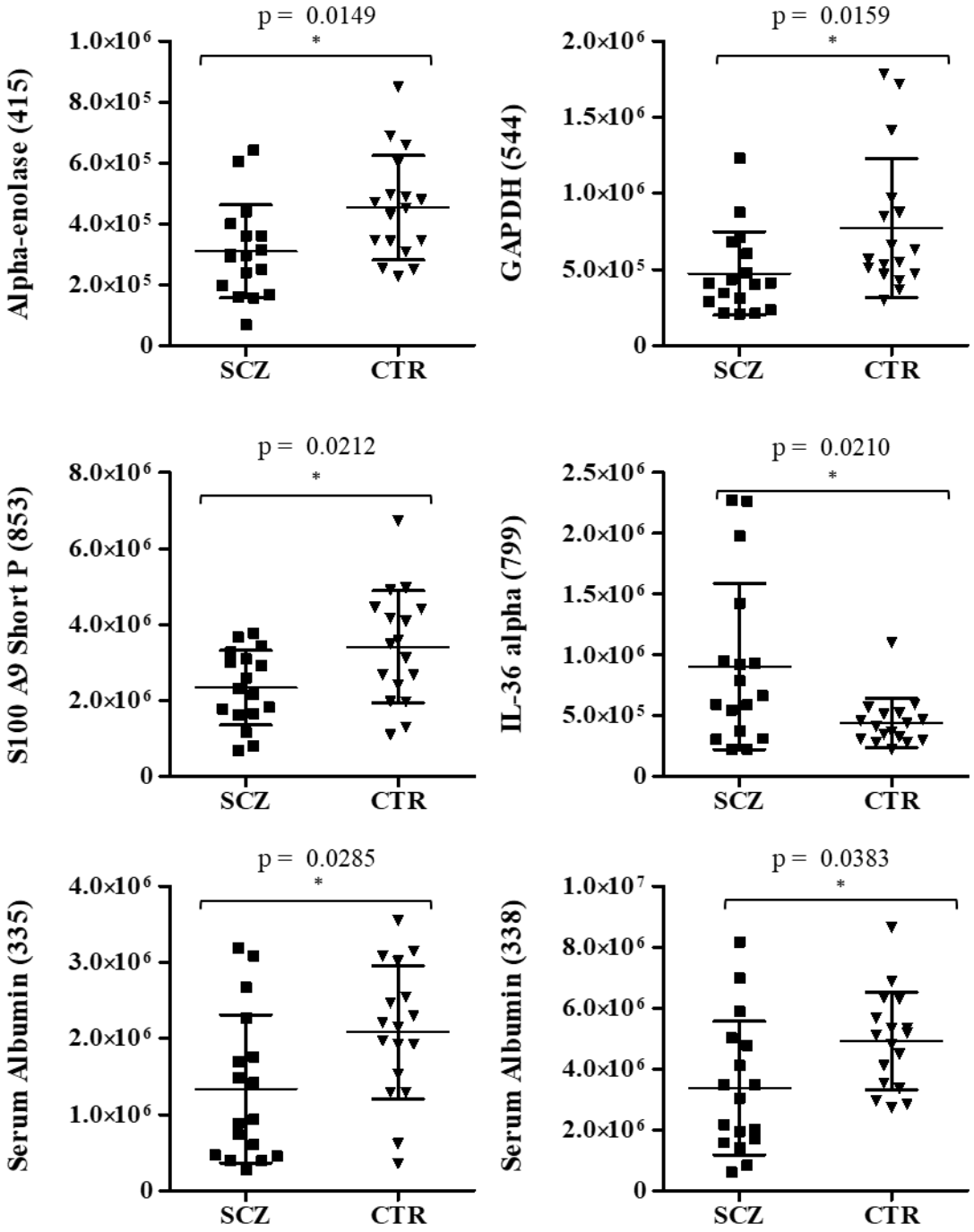


Figure 6. *Spot volume distributions of differentially expressed proteins between SCZ patients and healthy control subjects*.



5 Discussion

Results obtained from the analysis of the acid-insoluble fraction of saliva allowed us to identify differentially expressed proteins belonging to biological pathways that were already reported to be impaired in SCZ by previous proteomic studies. Energy metabolism seems to be one of the pathways most involved in SCZ pathophysiology, with a specific impairment of glycolytic process [99–101]. Impairment of energy metabolism leads to mitochondrial dysfunction and oxidative stress which is a central feature of SCZ [99]. In our analysis we found several glycolytic enzymes, named glyceraldehyde-3-phosphate dehydrogenase (GAPDH), phosphoglycerate kinase 1 (PGK1) and alpha-enolase (ENO1) with decreased levels in saliva from SCZ patients with respect to controls. GAPDH is a key enzyme in glycolysis pathway that catalyses the step in which D-glyceraldehyde 3-phosphate (G3P) is converted into 3-phospho-D-glyceroyl phosphate. This protein exists in the cytoplasm as tetramer composed by four identical subunits of about 37 kDa each one containing a thiol group fundamental for catalytic activity [102]. GAPDH is also involved in inflammation participating in the GAIT (gamma interferon-activated inhibitor of translation) complex [103] and in antimicrobial humoral immune response mediated by an antimicrobial peptide constituted by its amino acids residues 2-32 [104]. Beside its role in glycolysis and immune response, GAPDH is involved in modulation of organization and assembly of cytoskeleton [105] since it was one of the first glycolytic enzymes known to interact with tubulin and actin, facilitating microtubule bundling and actin polymerization [106]. Cytoskeletal abnormalities are often associated to SCZ pathophysiology because alterations in neurons lead to loose of synaptic connectivity and the ability to transmit incoming axonal information to the somatodendritic domain [107]. The decreased level of GAPDH we observed may be related to cytoskeletal tensegrity architecture which plays a pivotal role in the aetiology of SCZ condition.

On the other hand, reduced level of salivary GAPDH may be related to oxidative stress that can promote the formation of high molecular weight disulphide linked GAPDH aggregates, through a process called nucleocytoplasmic coagulation. This cascade of oxidations may augment GAPDH misfolding, leading to intermolecular disulphide cross-linking and aggregation [108]. Likewise, we found decreased level of cystatin B, another protein involved in immune response and oxidative stress, in fact it has been demonstrated that this protein can protect neurons against ROS damage [109], but also it can interact with SOD1 to regulate the redox state of the cell [110]. In contrast with our study on the acidic-insoluble fraction of saliva, previous study, performed on the acid-soluble fraction of saliva from schizophrenic patients, highlighted increased levels of the S-glutathionylated and S-cysteinylated proteoforms of cystatin B [97]. However, since the reducing conditions used in 2DE experiments do not allow to discriminate between the differently oxidized proteoforms of cystatin B, it is impossible to draw a parallel between the two results.

PGK 1 is another glycolytic enzyme that catalyses the reversible transfer of a phosphate group from 1,3-bisphosphoglycerate (1,3-BPG) to ADP producing 3-phosphoglycerate (3-PG) and ATP. PGK1 exists as a 415-residue monomer containing two nearly equal-sized domains that correspond to the N- and C-termini of the protein. The two domains of the protein are separated by a cleft and linked by two alpha-helices. Among the metabolic pathways dysregulated in SCZ it is comprised also the oxidative phosphorylation [111], and PGK1 is reported to be a positive regulator of this pathway, therefore its diminished level can reflect in the dysfunction of the respiratory chain. ENO1 is a glycolytic enzyme that catalyses the conversion of 2-phosphoglycerate (2-PG) to phosphoenolpyruvate (PEP). ENO1 (47 kDa) is one of three enolase isoforms, along with gamma and beta enolase (ENO2 and ENO3).

Each isoform is a protein subunit that can hetero- or homodimerize to form αα, αβ, αγ, ββ, and γγ dimers. ENO1 is defined as a moonlighting protein, since it displays other functions beside the canonical one [112]. ENO1, along with GAPDH, is prone to reactive oxygen species (ROS) damages [111] and the high oxidative stress and ROS production peculiar of SCZ, may be responsible of the decreased levels of ENO1 in SCZ saliva.

The high oxidative stress condition typical of SCZ may be related also to the decreased levels of Annexin A2 which belongs to the annexin protein family. Proteins of this family bind calcium and have different roles based primarily on their structure. They possess a N-terminal domain which is different for every annexin member, thus determining the different functions, and a C-terminal domain which instead is conserved among members of the family [113]. This protein is considered a pleiotropic protein, since it is involved in gene regulation, cellular transformation, regulating membrane dynamics, cytoskeletal re-arrangement, and fibrinolysis. Annexin A2 is also a marker for different types of cancer, especially regarding metastatic process [114]. It has been shown that antibodies against Annexin A2 can reduce tumour progression and metastasis [115]. Beside these functions Annexin A2 has been proposed as a ROS scavenging molecule, in fact it possesses four cysteine residues (Cys-8, Cys-132, Cys-261, and Cys-334) among which Cys-8 is resulted to be the redox sensitive one [115]. Under physiological condition the proposed redox mechanism involving Annexin A2 includes the oxidation of Cys-8 and the concomitant reduction of H_2O_2 to H_2O. Oxidised annexin A2 is then reduced by the Thioredoxin redox system and can participate in multiple redox cycles [116]. On the contrary during the high oxidative stress condition typical of SCZ this function of Annexin A2, as a scavenger of ROS, can be impaired since Annexin A2 accumulates in cell nucleus during oxidative damage to protect cellular DNA from damage [116].

Another pathway involved in SCZ process is the immune and inflammatory system. We found several proteins differentially expressed in our analysis linked to this system, named Interleukin 36 alpha, S100 A12 and PIP that were all increased in patients whereas alpha-crystallin B chain and S100 A9 were found with decreased levels.

Interleukin 36 cytokines are part of the IL-1 family and include IL-36 alpha, beta and gamma. N-terminal processing catalysed by neutrophil proteases increases the biological activities of these cytokines [117]. Little is known about IL-36 alpha biological roles, but it seems involved in inflammatory process in the skin during pathologies such as psoriasis [118]. Interleukin 36 alpha is also a potent activator of other cytokines such as IL-8 and IL-6 which are always found increased in serum of schizophrenic patients and are considered markers of inflammation process in SCZ patients [119].

S100 family consist in 25 acidic proteins ranging between 9 and 13 KDa. These proteins are expressed by vertebrates and found in several tissues and fluids such as plasma, urine and saliva. Their principal function is the binding of calcium and regulation of its homeostasis, but they can also participate in extracellular signalling, inflammation and anti-microbial activity. The C-terminal domain is different between the various proteins of the family and determines the functions of each protein. S100 proteins are found primarily as homo or heterodimers [120]. Protein S100 A12 is expressed by neutrophils and its levels are often found increased in patients suffering of neurodegenerative, inflammatory and neoplastic disorders. It is considered as a marker of localized inflammation, and in the work of Iavarone et al. [97] exploring the soluble fraction of saliva from schizophrenic patients, protein S100 A12 was found to be increased exactly like in the insoluble fraction analysed in this work. Moreover, in our 2DE analysis, Protein S100 A12 spot (917) is visible only in SCZ patients and has the highest

fold variation (5,32) amongst the proteins resulted differentially expressed between the groups of subjects.

Protein S100 A9, as well as S100 A12, is associated to inflammatory states in the organism but otherwise we detected it under-expressed in patients. In the 2DE gel it is possible to distinguish four spots of S100 A9: two of them, defined as long type, are acetylated at the N-terminus and differ from each other for the presence of a phosphate on the Thr_{112} residue. The "short" isoforms are acetylated at the N-terminus but lack the first five amino acids and the phosphorylation is on the Thr_{108} residue [92]. In our analysis we identified protein S100 A9 short type phosphorylated, since the four 2DE spots of S100 A9 were characterized in the work of Arba et al. [92] by staining the 2DE gel with Pro-Q Diamond for specific detection of phosphoproteins along with MS/MS spectra analysis of the digested spots.

Prolactin-inducible protein (PIP) is expressed in several exocrine tissues such as glands and it is over-expressed by prolactin and androgens but under-expressed by estrogens [121]. The exact role of this protein remains unclear, but it is associated to fertility, immunoregulation due to its ability to bind immunoglobulin G (IgG) and tumoral progression [122]. Another biological role of this protein is to prevent bacterial colonization, that is why is expressed in proximity of the main sites of bacterial entry [123]. Early studies evidenced a prolactin elevation during therapy with antipsychotic [124] and more recently serum prolactin levels have been associated with antipsychotic response to risperidone in patients with SCZ [125]. The degree of prolactin elevation depends upon the binding affinity of antipsychotics to the dopamine receptors and indeed dopamine receptors blockade by antipsychotics and hyperprolactinemia has been well reported in various studies. Based on these studies, prolactin elevation may be considered as a central marker for antipsychotic function on dopamine receptors [126].

The increased levels of PIP we observed in saliva of SCZ may be associated with the different antipsychotic therapy used to treat the disease. However, this hypothesis should be validated with clinical data on serum prolactin levels and therapy used during saliva collection.

Alpha-crystallins are protein complex composed of αA and αB subunits that are highly expressed in the lens. Alpha-crystallin A is preferentially restricted to the eye lens whereas Alpha-crystallin B is expressed widely in many tissues and organs and belong to the Heat Shock proteins family. As molecular chaperones, Alpha-crystallins prevent aberrant protein interactions, but they can also remodel and protect the cytoskeleton, inhibit apoptosis, and enhance the resistance of cells to stress [127]. Due to its chaperone-like activity, Alpha-crystallin B can bind different proteins preventing their aggregation and precipitation. This function has been observed in several diseases such as Alzheimer's [128], Parkinson's [129] and Multiple Sclerosis [130]. In the work performed by Garbett et al. [131], the authors investigated the effect of maternal immune activation (MIA) upon an infection in mice, as a risk for onset of schizophrenia in the offspring. Several genes were found upregulated during MIA, including crystallins genes whit the expression level positively correlated with the severity of MIA itself. In contrast with the current literature our results highlighted decreased level of Alpha-crystallin B in saliva from SCZ, however it is worth noting that Alpha-crystallin B may regulate the dynamic assembly of the cytoskeleton [132], and on this regard its decreased level may be related to the cytoskeletal abnormalities often associated to SCZ pathophysiology.

6 Conclusions

The use of 2DE coupled to high resolution MS allowed us to provide a bidimensional map of the acid insoluble fraction of saliva from SCZ patients. Data obtained integrate the results provided by the analysis of acid soluble fraction of saliva from the same patients, giving a complete picture of the salivary proteome of SCZ. We were able to detect differentially expressed proteins between patients and control subjects, such as glycolytic enzymes, S100 A12, PIP and IL-36 alpha. Our results are in agreement with those obtained by proteomic studies mainly focused on brain or plasma which highlighted an imbalance in energy metabolism, immune system and inflammation.

These findings confirm that saliva reflects the pathophysiological state of the body and thus, can be used as a potent diagnostic tool since it is easier to collect compared to other biological samples.

PART 2:

PROTEOMIC CHARACTERIZATION OF THE MUCOSAL PELLICLE FORMED IN VITRO ON A CELLULAR MODEL OF ORAL EPITHELIUM

Introduction

1 The oral mucosal pellicle

The mucosal pellicle is a thin layer of salivary proteins lining epithelial oral cells. Most data available on the composition of the mucosal pellicle comes from *in vivo* studies, in which the pellicle is collected along with epithelial cells by rinsing the mouth with water or by scraping the mucosa. In terms of composition early study on the pellicle allowed to identify salivary mucins 5B and 7, amylase, Cystatin SN and PRPs [133] as more represented proteins, while more recently has been highlighted the presence of other proteins such as Immunoglobulin A (IgA), Carbonic Anhydrase VI (CAVI), Cystatin S and Statherins on sampled oral cells [78]. This study showed also that some proteins are retained specifically on the cell surface like IgA, whose concentration on cell surface is greater than in saliva, while other proteins like amylase are not retained on cell surface [78]. *In vitro* studies of the pellicles are often performed on artificial supports such as the polydimethylsiloxane (PDMS) that can mimic the oral mucosa. Such studies allowed to propose a model for the mucosal pellicle structure: a first anchorage layer constituted by hydrophobic regions of mucins and small salivary proteins, and an external layer constituted by glycosylated regions of salivary mucins [134] (Figure 7). The thickness of the pellicle formed on PDMS can reach a maximum of 120 nm while *in vivo* experiments showed that the thickness of the pellicle can reach up to 100 nm [4] (Figure 8). The structure of the mucosal pellicle, especially the external layer thickness depends of the ionic strength of the medium. Since the glycosylated regions of mucins are negatively charged due to the presence of sialic acid residues and sulphates [135], if the ionic strength lowers too much, negatively charged regions of the mucins collapses over the positively charged regions determining the gelification of these mucins.

Figure 7. *Structural model proposed for the mucosal pellicle (figure: courtesy of Francis Canon, INRA-CSGA, Dijon, France).*

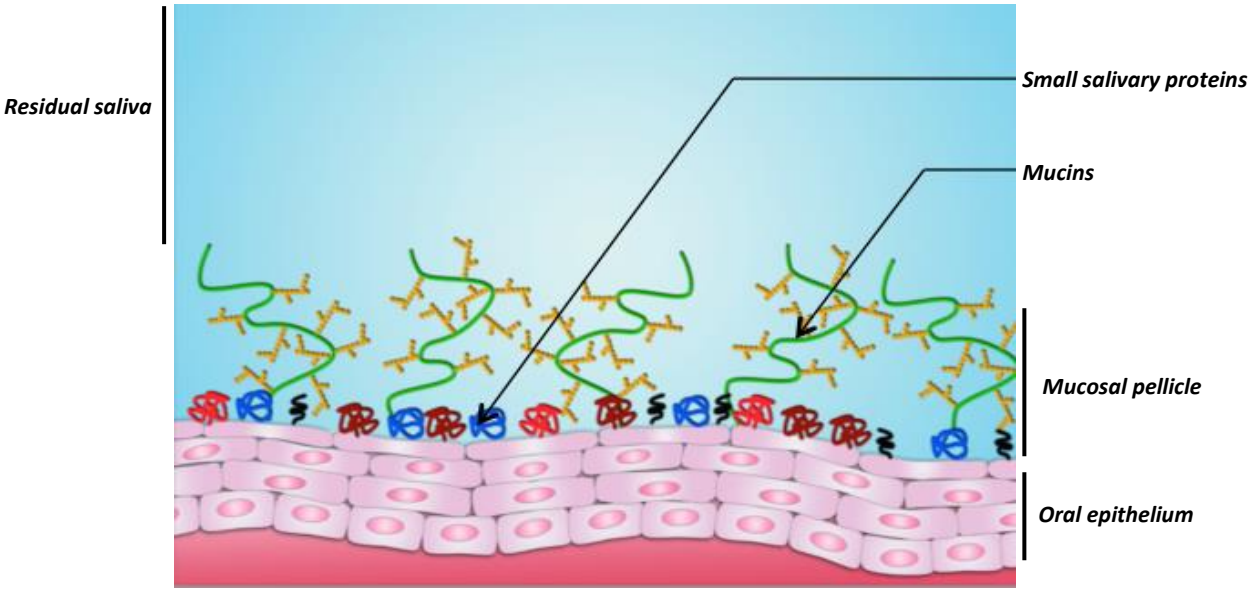
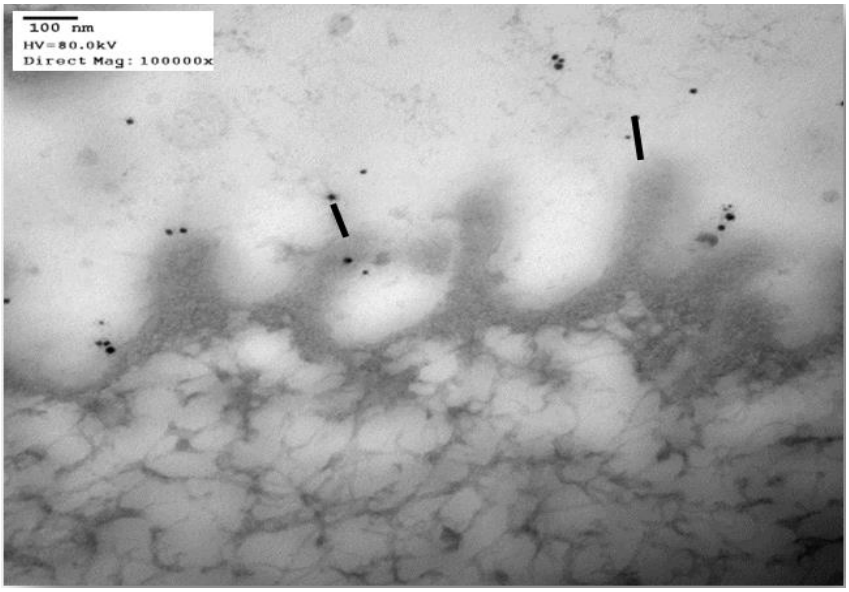


Figure 8. *Morphological observation of oral epithelium comprising the mucosal pellicle.Bars indicate the thickness of the pellicle* [4].



The mucosal pellicle has three major functions in oral cavity: lubrication, biological barrier and hypothetically taste perception. Lubrication of the oral cavity by the mucosal pellicle is determined by the mucin content of the pellicle. The two major mucins of the pellicle MUC5B and MUC7 are extensively glycosylated and are responsible of two rheological properties of saliva: viscosity and stretchability [136]. Negatively charged regions of mucins retain water molecules and can also form a hydrated gel. Electrostatic and steric repulsions between glycosylated regions of mucins contribute to the lubrication properties of the pellicle. The biological barrier properties of the pellicle are also determined by the mucus layer. This layer works as a selective filter due to the dimensions of the pores formed by mucins net, and also it can retain molecules not only by size filtering but also thanks to electrostatic and hydrophobic interactions [137]. The net formed by mucins can also block the access and the fixation of microorganisms on the oral mucosa. It has been demonstrated that IgA are concentrated by mucins in the mucosal pellicle, suggesting the role of reservoir for immune proteins of the pellicle and protection of the mucosa from infections [138]. The mucosal pellicle could also participate in sensory perception of flavour compounds. It is for example suggested that the interactions between proteins of the pellicle and aroma compounds and the subsequent progressive releas of aromas in the oral cavity can give rise to the phenomena of aromatic persistence. It seems also involved in perception of astringency [139]. For example, it has been recently demonstrated that tannins can cause alterations of the mucosal pellicle structure by aggregation of mucins [16]. The mechanism that drives the mucosal pellicle formation is not fully understood. Analysis of epithelial cells incubated with saliva showed the formation of complexes between cells and salivary proteins [138,140]. This process is mediated by transglutaminases (TGM) [141] which has been detected in oral cells lysates [142].

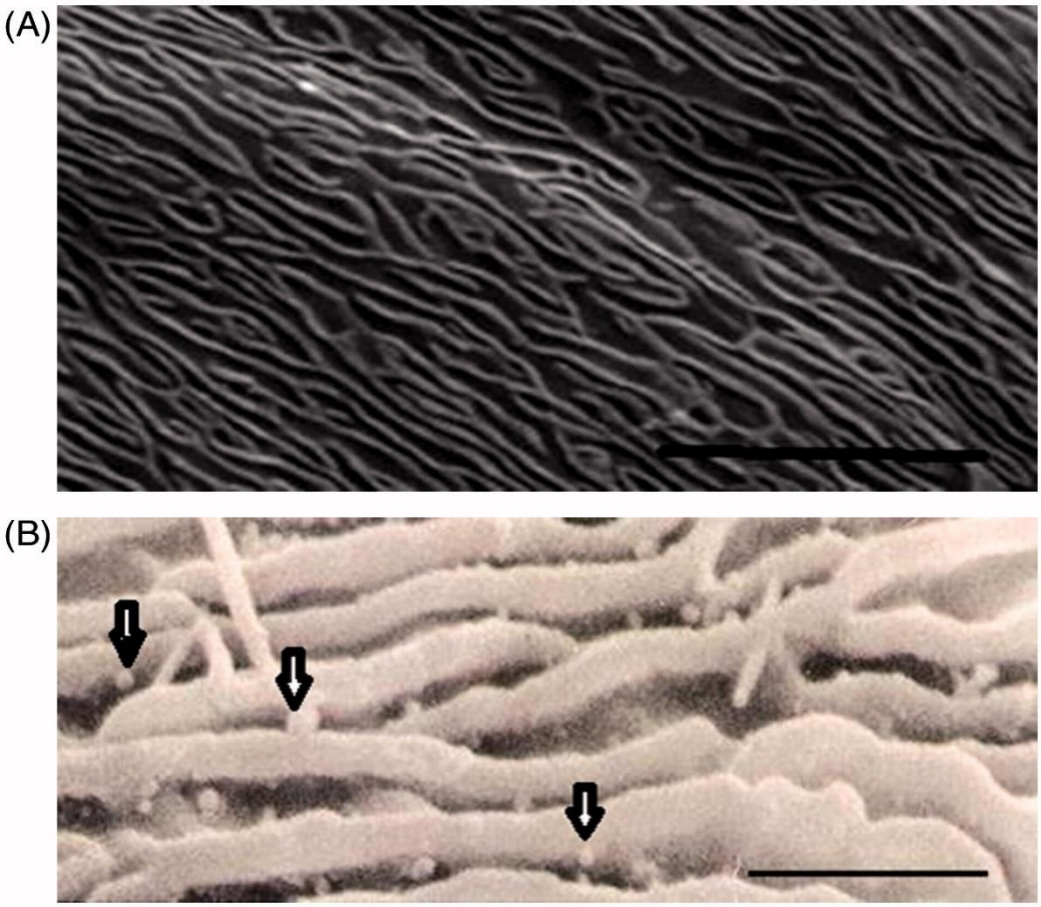
TGM can indirectly improve the adhesion of MUC7 on artificial supports due to interactions between MUC7 and salivary proteins retained on the surface by transglutamination reactions. This effect mediated by TGM is not shown in MUC5B adhesion on artificial surfaces [142]. It seems that the oral mucosa has a hydrophobic nature, therefore hydrophobic interaction with pellicle components can occur. MUC5B is retained only on hydrophobic surfaces confirming the pellicle model in which the anchorage layer is partly constituted by the hydrophobic regions of mucins. Another important aspect is that there are protein-protein interactions between pellicle and cell surface. For example adhesion of MUC5B on the cells is improved by the presence of mucin 1 (MUC1) [140]. MUC1 is a membrane mucin present in epithelia and has various isoforms. The interactions between MUC5B and MUC1 are fundamental because they provide the scaffold for the pellicle constitution [143].

1.1 Oral mucosa

The oral mucosa is an epithelial tissue of the oral cavity consisting of two layers: stratified squamous epithelium and lamina propria. According to the localization in mouth the squamous epithelium can be keratinized or non-keratinized. Non-keratinized squamous epithelium is present in the soft palate, inner lips, inner cheeks, the floor of the mouth, and ventral surface of the tongue. Keratinized squamous epithelium is present in the gingiva and hard palate. The dorsal surface of the tongue presents a specialized mucosa having the characteristics of the other two types of mucosa and being provided with taste buds [144]. On the surface of the oral mucosa, epithelial cells present folds of their membranes called microplicae (MPL) that have the same dimensions in all oral cavity regions but differ is their organization. Keratinized mucosa has an alveolar appearance while the surface of the cells covering the tongue, between the taste buds, presents different motives, parallel and in branches while microplicae on cells of lining mucosa present only branches with a greater

density than the other oral regions [145]. As shown in figure 9, microplicae seem to be discontinuous on oral surface and membrane associated proteins like MUC 1 were detected between them [146].

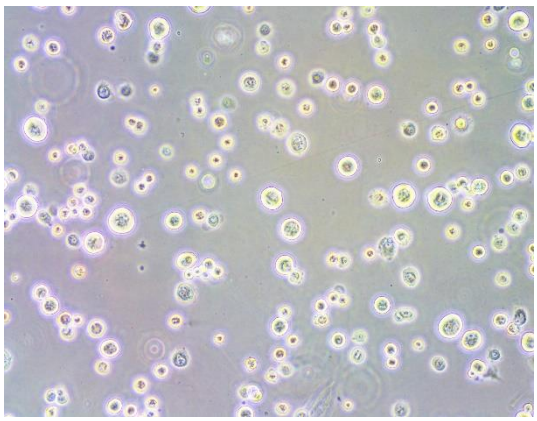
Figure 9. *A) microplicae on buccal cells surface (bar corresponds to 10 µm). B) magnification shows small molecules (pointed with arrows) protruding from membranes [146].*



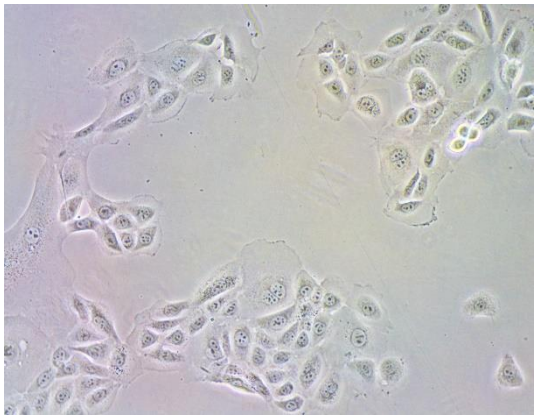
1.2 Studying the mucosal pellicle through a cellular model of oral epithelium

Research on salivary pellicles, as mentioned above, is performed using *in vivo* samples or artificial supports to mimic the oral mucosa. However, these approaches present some problems such as variability between subjects enrolled in the study for *in vivo* collection of samples and the fact that an artificial support lacks the structural characteristic of the real oral mucosa. During past years, cellular models have been used to study *in vitro* the permeability of intestinal and oral mucosa, or mucins-mucins interactions during pellicle formation [138,147,148]. Among the different cell lines that can be used, the TR146 cell line deriving from a squamous cellular carcinoma (Figure 10) was chosen for the present work. This cell line possesses the structural characteristics of oral epithelial cells (i.e. the presence of microplicae) but it is not mucus-secreting [148,149]. Since it is known that MUC1 is involved in pellicle constitution, this cell line has been stably transfected to express this protein (TR146/MUC1), resulting in a more accurate model of oral epithelium [140]. This model offers the advantages of a cellular model with less ethical and safety concerns compared to sampling on subjects, controlled conditions, reduction of inter-individual variability and enables to characterize the protein composition of the mucosal pellicle formed on a realistic support compared to synthetic surfaces. However, there are some limitations in this model like the cell's desquamation that is usually present *in vivo*, and which could influence the interactions between pellicle and salivary proteins. Moreover, this model used the truncated form of MUC1 which could show less retention of MUC5B [140].

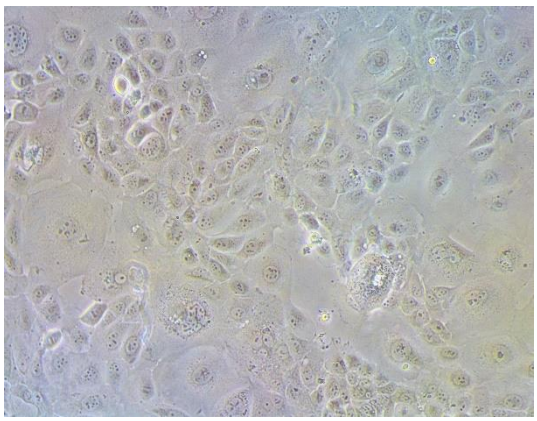
Figure 10. *TR146 cells during various steps of growth until reaching confluence in the culture flask.*



Day 1



Day 3



Day 7

2 Aim of the work

The aim of this work was first to establish a protocol suitable for sampling and if possible, isolating the mucosal pellicle formed onto the TR146 and TR146/MUC1 cell lines. Second, we aimed to characterize the proteome composition of the mucosal pellicle formed *in vitro* on both TR146 and TR146/MUC1 cells and to compare by a *bottom-up* mass spectrometry approach the mucosal pellicle proteome with the salivary proteome.

3 Materials and methods

3.1 Saliva Collection

Saliva was collected from 15 healthy volunteers who were in good oral health. The subjects did not drink, eat or smoke for at least two hours prior to saliva collection. Subjects donated saliva by spitting the oral fluid accumulating spontaneously in their mouth into plastic vessels in ice over 1 h period. Samples were then pooled and centrifuged at 14 000 g for 20 min at 4°C. Clarified saliva was aliquoted in 2 mL tubes and frozen at -80°C.

3.2 Cell culture and constitution of the mucosal pellicle.

TR146 epithelial cells were obtained from the European Collection of Cell Cultures (ECACC, Salisbury, Wiltshire, UK). The stably transfected cell line TR146/MUC1 [140] was also used. Cells were grown in T75 flasks, in DMEM/F12-Glutamax medium from Gibco (1:1, v:v), supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. 2.5 mg/mL of geneticin G418 were added to the medium only for TR146/MUC1 cell line. Cells were incubated at 37° C in a humidified atmosphere containing 7.5 % CO_2, subcultured at 7 days interval and medium was changed every 2 days. Cells were seeded at a density of 0.1 x 10^6 cells/mL, in order to reach confluence in 48 h, in a 24 or 12 well plate for constitution and isolation of mucosal pellicle.

Cells were observed with a light microscope Nikon Eclipse TS100 (Nikon, Tokyo, Japan) equipped with a 10x plan objective. At confluence, cells were incubated for 2 h with clarified saliva diluted in DMEM/F12-Glutamax, to constitute the mucosal pellicle.

3.3 Isolation of the mucosal pellicle.

After incubation with saliva, cells were washed twice with DPBS 1X (Gibco) in order to eliminate the non-adsorbed saliva. Washing solutions were collected and subjected to overnight precipitation with 10% Trichloroacetic Acid (TCA) and 20 mM Dithiothreitol (DTT) in Acetone. The precipitate was solubilized with extraction buffer (7 M Urea, 2 M Thiourea and 2% CHAPS) and stored at -20°C. To isolate the mucosal pellicle, the precipitate was solubilized with increasing concentration of SDS and Tween-20 or submitted to shaving protocol as following reported.

3.3.1 SDS and Tween-20 wash

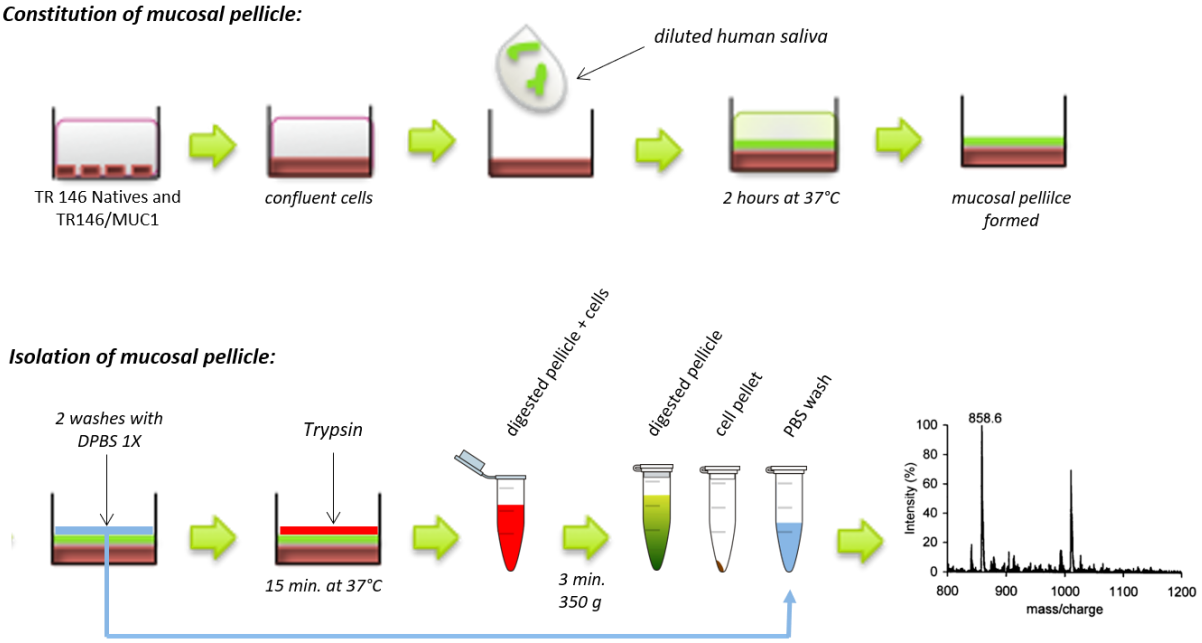
Cells were incubated with increasing concentrations of SDS (from 0.2% to 0.4%) or with increasing concentrations of Tween-20 (from 0.05% to 0.1%) for 10 min. at 40 rpm. After incubation, SDS and Tween-20 solutions were collected into an Eppendorf tube and stored at -20 °C for further analysis. Cells were washed again with DPBS 1X and homogenized with RIPA Buffer (Sigma-Aldrich) following the manufacturer's instructions and then stored at -20°C.

3.3.2 Shaving protocol

Cells were incubated for 30 sec at room temperature with 0.6 mL of Trypsin-EDTA 0.05% (Gibco). After removing 0.4 mL of Trypsin, cells were incubated at 37°C for 15 min with the remaining 0.2 mL of Trypsin. After incubation, 1 mL of DMEM/F12-Glutamax was added to the cells, and samples were transferred into Eppendorf tubes.

Samples were then centrifuged at 350 g in order to separate the cell pellet from the supernatant containing the digested mucosal pellicle. Supernatants were stored at -20°C for further analysis, while cell pellets were homogenized with RIPA Buffer (Sigma-Aldrich) following the manufacturer's instructions and then stored at -20°C. This experimental protocol enabled to produce four types of samples (each type in triplicate): saliva pool used for constituting the mucosal pellicle (S); washing solutions (W) i.e. proteins not strongly bound to cells; shaved-off pellicle (P) obtained by light trypsin digestion of the cell surface, and remaining cells extracts (C). Samples W, P and C were obtained for both types of cells and are named with reference to the cell line (e.g. W-TR146 and W-TR146/MUC1). In addition, control samples were also produced following the same protocol but omitting the incubation with saliva, resulting in so called W-Ctrl, P-Ctrl and C-Ctrl samples. Figure 11 shows a schematic representation of the experimental procedure used.

Figure 11. *Schematic representation of the experimental procedure utilized for constitution and isolation of the mucosal pellicle from the TR146 cell lines.*



3.4 Protein quantification

Total protein concentration in P samples obtained from the shaving protocol was determined using BSA as a standard and reading the absorbance at 214 nm with a Nanodrop ND-1000 spectrophotometer (Thermo Scientific). The other samples obtained by the two isolation protocols (washing with SDS or Tween-20) were quantified using a Bradford Assay.

3.5 Immunoblot

For immunoblot experiments different protein loads were used based on the isolation protocol: for SDS and Tween-20 washings 5 and 10 μg of saliva pool and 5 μg for pellicle samples were used; for shaving protocol 20 μg of proteins for pellicle and saliva pool samples were used. Samples were loaded onto a 0.45 Nitrocellulose membrane (Bio-Rad). For proteins detection the following primary antibodies produced in mouse were used: anti-MUC5B (F2, kind gift from Prof. Veerman, Free University of Amsterdam, 1:400), anti-SC (pIgR) (sc-56595, Santa Cruz Biotechnology, 1:200) and anti-Cystatin S/SA/SN (sc-73884, Santa Cruz Biotechnology, 1:800). The secondary antibody was goat-anti mouse IgG coupled to HRP (LifeTechnologies, 1:2000). Non-specific binding was prevented using 8% fat-free milk. HRP activation was performed using the Clarity Western ECL Substrate Kit (Bio-Rad). Blot images were acquired using the ChemiDoc imaging system (Bio-Rad).

3.6 Mass Spectrometry

3.6.1 In gel digestion

20 µg of each sample were loaded onto a 4-20% SDS-PAGE Gel, and let run at 150 V for about 1 cm length at 150 V. After fixation and staining with Coomassie Blue, gel bands were excised, reduced with TCEP 10mM and alkylated with iodoacetamide 55mM before tryptic digestion (200 ng of trypsin). After incubation overnight at 37°C, tryptic peptides were extracted from the gel bands using Acetonitrile (ACN) and evaporated with a SpeedVac. Peptides were resuspended with 20 µL of 2% ACN and 0.1% formic acid (FA).

3.6.2 Purification with C18 column of digested pellicle samples and MS analysis

200 µl of sample were desalted with C18 Microspin Columns (Harvard Apparatus). Peptides were eluted with 400 µL of a solution containing 60% ACN and 0.1% FA. After evaporation with SpeedVac, peptides were resuspended with 30 µL of 2% ACN and 0.1% FA.

3.6.3 MS analysis

Samples were analyzed with a nanoUPLC (nanoRSLC, ThermoFisher) system coupled with a mass spectrometer equipped with a nanospray source (Advion TriVersa Nanomate) and an ionic trap in tandem with an orbitrap (LTQ-Orbitrap elite, Thermo Scientific). 18 µL of sample were loaded on an enrichment pre-column (Acclaim PepMap C18 75µm*20mm; ThermoFisher) with solvent A (2% ACN, 0.1 FA) for 3 min. with a flow rate of 5µL/min. Elution of peptides was performed by increasing the concentration of solvent B (80% ACN, 0.1% FA) with a separation column (AcclaimPepMap 75µm*250mm, 2µm; ThermoFisher) kept at 33°C with a flow rate of 300 nL/min. A 2 h gradient was used increasing the quantity of solvent B from 2% to 44%. For quality check of the runs, 100 fmol of a BSA digest were used.

MS and MS/MS spectra were collected in positive mode with the resolution of 120000 (at 200 m/z). The acquisition range was from 400 to 1700 m/z. Tuning parameters: capillary temperature was 250°C, source voltage 1.5 kV, S-Lens RF level 60%. In data-dependent acquisition mode the 15 most abundant ions were selected and fragmented by using Higher-energy collisional dissociation (HCD), with 35% normalized collision energy, isolation width of 2 m/z, activation q of 0.25.

3.6.4 Proteins identifications

For identification of proteins the search algorithm Mascot (v2.5.1) was used starting from MS/MS data obtained. The search was performed with the software Proteome Discoverer (ThermoFisher, v2.1) in a database of human proteins (Uniprot, May 2017). The search parameters were 10 ppm tolerance for precursor ions and 0.6 Da for product ions, carbamidomethylation of cysteine (+57.02 Da) as fixed modification and oxidation of methionine (+15.99 Da) as variable modification. The results obtained for the triplicates of each sample were merged and only identifications with a high FDR confidence, a Mascot score >25 and at least 2 unique peptides were validated.

3.6.5 Protein quantification and statistics

MS raw data were processed using Maxquant (v1.5.3.30). Peak lists were created using default parameters and searched using the Andromeda search engine against human sequences from Uniprot. Data normalisation and protein abundance estimation were performed using the MaxLFQ (label free quantification) option implemented in Maxquant using a minimal ratio count of 2. Both modified and unmodified peptides were considered for quantification while shared peptides were excluded. Only proteins with two valid values per condition were kept for further analysis. Statistical analyses were then performed using the R open-source

software (https://cran.r-project.org). The differences in protein abundance in P-TR146 and P-TR146 MUC1 samples were tested using a Welch t-test. The Benjamini–Hochberg correction method was applied to control the False Discovery Rate (FDR).

4 Results

In order to isolate the mucosal pellicle from the TR146 cell lines three different protocols using washing with SDS and Tween-20 and shaving with trypsin were performed. Protein quantification assays evidenced a poor protein concentration (~ 0.04 mg/mL) in pellicle samples from washing with Tween-20 compared to SDS washing (~ 0.26 mg/mL) and shaving protocol (~ 1.6 mg/mL). Therefore, the samples from Tween-20 washing were not further analysed. Proteins from pellicle samples were firstly analysed by dot blot and western blot in order to evidence the presence of S-type Cystatins and SC component of pIgR in the pellicle. In fact, these proteins were already detected as pellicle components by using immunoblot techniques on *in vivo* collected samples [78,133].

As shown in figure 12, SDS washing protocol did not allowed detecting S-type Cystatins and SC by western blot or dot blot in pellicle samples even though in the saliva pool used for constituting the pellicle, these proteins are nicely detected. On the contrary the “shaving protocol” enables detecting both S-type Cystatins and SC in pellicle samples (Figure 13).

Figure 12. *A) Western blot perfromed to detect the presence of S-type cystatins in samples obtained after SDS washing protocol. 5 µg of total proteins were loaded onto a Stain Free TGX Gel 4-20% T (Bio-Rad) and after SDS-PAGE, proteins were transferred on a 0.45 µm nitrocellulose membrane (Bio-Rad). B) Dot blot performed to detect the presence of SC in samples obtained after SDS washing protocol. On a 0.45 µm nitrocellulose membrane, 10 and 5 µg of total proteins from the saliva pool were loaded while for pellicle samples 10 µg of proteins were loaded, followed by Red Ponceau staining. The following primary antibodies were used: anti-Cystatin S/SA/SN (sc-73884, Santa Cruz Biotechnology, 1:800), anti-SC (pIgR) (sc-56595, Santa Cruz Biotechnology, 1:200). The secondary antibody was goat-anti mouse IgG coupled to HRP (LifeTechnologies, 1:2000)*

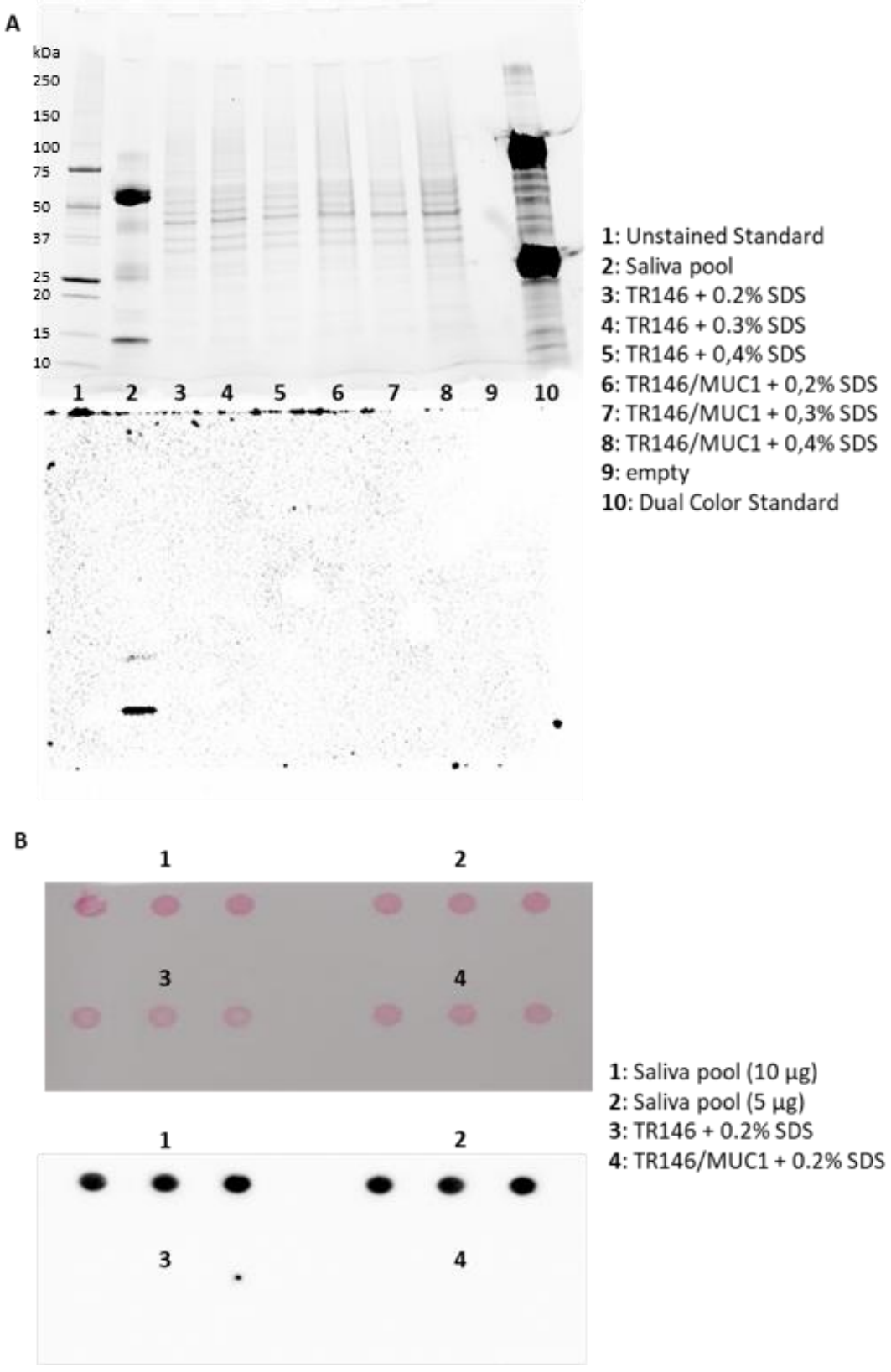
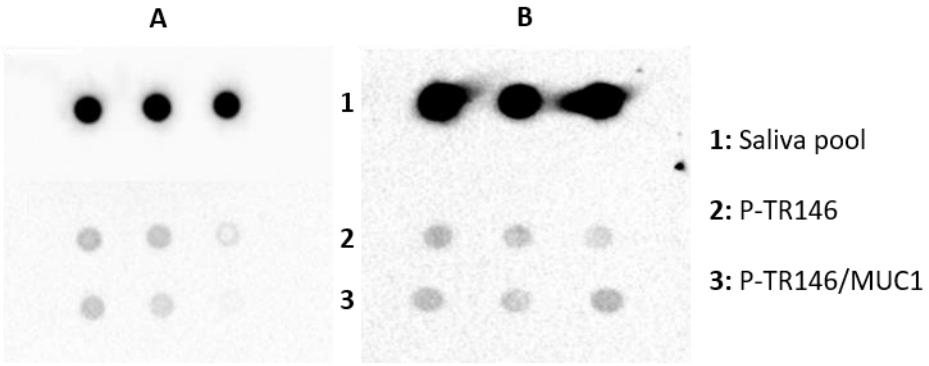


Figure 13. *Dot blots performed to detect the presence of SC (A) and S-type Cystatins (B) in samples obtained after the "shaving protocol". 20 µg of proteins for saliva pool and pellicle samples were loaded onto a 0.45 µm Nitrocellulose membrane (Bio-rad). For proteins detection the following primary antibodies produced in mouse were used: anti-SC (pIgR) (sc-56595, Santa Cruz Biotechnology, 1:200) and anti-Cystatin S/SA/SN (sc-73884, Santa Cruz Biotechnology, 1:800). The secondary antibody was goat-anti mouse IgG coupled to HRP (LifeTechnologies, 1:2000).*



Based on these results samples obtained with the “shaving protocol” were selected for mass spectrometry analysis. Samples S, W and C were all subjected to in gel digestion and subsequent analysed by high resolution mass spectrometry. Digested pellicle samples (P) were treated differently: they were purified with with C18 Microspin Columns and then analyzed by high resolution mass spectrometry. Mass spectrometry data from samples S, W, C and P were compared with those obtained from the pool of saliva used for constituting the mucosal pellicle in our experiments (Table 3). The “shaving protocol” was also performed for both cell lines on control samples (W-Ctrl, P-Ctrl and C-Ctrl) which were obtained with the same procedure but omitting the incubation of cells with saliva. These controls allowed us to verify if PBS washing could cause lysis of the TR146 cells: MS analyses performed on W-Ctrl TR146 and W-Ctrl TR146/MUC1 confirmed that only a few proteins were washed away from cell surface (1 and 20 proteins respectively) and therefore that cell integrity was preserved. Incubation with trypsin resulted as expected in the shaving-off of numerous proteins: 555 proteins identified in P-Ctrl TR146 and 445 proteins in P-Ctrl TR146/MUC1 samples. Term enrichment analysis (AmiGo 1.8 version) (category: cellular component) revealed that for both cell types, proteins could be part of several compartments such as cytoplasm, membranes, intracellular and extracellular organelles etc. For both cell types, however, the most significantly enriched was the category “GO:0070062 extracellular exosome” (37.8% of proteins, $p= 8*10^{-150}$ for P-Ctrl-TR146 and 38.6% of proteins, $p=5 \times 10^{-194}$ for P-Ctrl-TR146/MUC1). Of special interest is also the category “GO: 0071944 cell periphery” representing 25.1% and 23.1% of proteins for P-Ctrl-TR146 and P-Ctrl-TR146/MUC1, respectively. Analyses performed on C-CTRL TR146 and C-CTRL146 MUC1 identified surprisingly few proteins (40 and 27, respectively): these showed enrichment in the cellular component categories “extracellular exosome” and “nucleus” but, compared to proteins in

pellicle samples, also proteins from “cystoskeleton”. The small number of proteins in C-CTRL may result from low efficacy of the RIPA extraction protocol. In any case, taken together, these analyses suggest that the shaving protocol may result in some cell degradation with the release of intracellular content in the so-called “pellicle” extract. Nevertheless, the specific presence of the “cell periphery” component confirms that the shaving protocol targets also the surface of the cell and is thus suitable for the characterization of the mucosal pellicle. Table 4 provides the total number of proteins identified by mass spectrometry in W (washing solutions) and P (pellicle) samples, and also the number of proteins identified in these samples that are common to saliva. W samples provide information on proteins which are readily removed from the cells. Since the washing step itself does not induce extensive cell lysis, as reported in the previous paragraph, these proteins must originate mainly from saliva. It should be noted that we opted for stringent identification criteria (Confidence: high, Mascot score > 25, at least two unique peptides), which probably explains that only 43 (for TR146) and 47 (for TR146/MUC1) of the proteins identified in W samples were also present in the list of proteins identified in saliva. Looking at numbers of proteins present in P samples, and as previously described in control samples, the shaving protocol with trypsin results in the release of a great number of cellular proteins or fragments. However, P samples also contain proteins present in saliva (23 for TR146, 21 for TR146/MUC1). Table 5 provides details on the proteins identified both in saliva and in either W or P samples. In this table, we further highlighted the proteins which most likely reach whole saliva through their secretion by salivary glands. For that purpose, we used the gene expression database Bgee (https://bgee.org) and for each single protein, we used two criteria of selection: 1) salivary glands (“parotid gland” or “minor salivary gland” or “saliva-secreting gland”) appeared in the first 10% of tissues most-expressing the selected protein; 2) expression was more abundant in salivary glands than in oral mucosa.

In contrast to proteins which may be present in pellicle samples incidentally because they are also present in epithelial cells, glandular proteins are of special interest since their presence in pellicle samples suggest that they are strongly bound to the cells since they are not removed entirely by the PBS washing step. Mass spectrometry results obtained by analyzing the pellicle samples (3 replicates for P-TR146, 3 replicates for P-TR146/MUC1), were also used for statistical analysis: Welch test was chosen for detection of differentially expressed proteins or family of proteins. 9 proteins presented p-value < 0.05, but they have not been confirmed by the false discovery rate (FDR) test at 5%, so the probability of having false discoveries among the results remains high. The principal component analysis (PCA) made from the experimental data allows us to separate the P-TR146 samples from the P-TR146/MUC1 samples. This observation is further strengthened with a Hierarchical Classification putting them into two separate clusters. As a conclusion, we observe with the PCA and hierarchical clustering that the two types of samples have slightly different proteomes, but with not strongly enough differentially expressed proteins to characterize this difference (Figure 14 and 15). We did not identify in P samples some proteins which are known to be constitutive of the pellicle like MUC5B, S-type Cystatins and Polymeric Immunoglobulin Receptor (its secretory component SC to be more specific) probably due to a low abundance respectively to other proteins. Dot blot was performed in order to assess the presence of such proteins in the samples. SC and S-type Cystatin detection was already shown in Figure 13. MUC5B, as shown in Figure 16, was detected both in pellicle on TR146 cell line and pellicle on TR146/MUC1.

Table 3. *List of proteins identified in the pool of saliva used for the constitution of the mucosal pellicle on TR146 cells.*

Proteins identified in saliva			
Protein FDR confidence	**Accession**	**Name**	**Score**
High	P63104	14-3-3 protein zeta/delta	146,3433333
High	P52209	6-phosphogluconate dehydrogenase, decarboxylating	229,4193766
High	P60709	Actin, cytoplasmic 1	91,92
High	Q01518	Adenylyl cyclase-associated protein 1	119,2938919
High	P01011	Alpha-1-antichymotrypsin	83,8352762
High	P01009	Alpha-1-antitrypsin	144,0323458
High	P01023	Alpha-2-macroglobulin	86,18333333
High	A8K2U0	Alpha-2-macroglobulin-like protein 1	36,10760115
High	P04745	Alpha-amylase 1	382,08
High	P06733	Alpha-enolase	230,1965398
High	P04083	Annexin A1	98,54896919
High	P03973	Antileukoproteinase	87,23856478
High	P01008	Antithrombin-III	35,97
High	P20160	Azurocidin	131,27
High	Q96DR5	BPI fold-containing family A member 2	266,7048378
High	Q8TDL5	BPI fold-containing family B member 1	223,6794092
High	Q8N4F0	BPI fold-containing family B member 2	148,9279402
High	P23280	Carbonic anhydrase 6	212,5697484
High	P04040	Catalase	89,09333333
High	P07339	Cathepsin D	65,81
High	P00450	Ceruloplasmin	25,96
High	P23528	Cofilin-1	115,9
High	P01024	Complement C3	292,527489
High	P01040	Cystatin-A	67,69
High	P04080	Cystatin-B	156,4033333
High	P01034	Cystatin-C	207,3081312
High	P28325	Cystatin-D	196,05
High	P01036	Cystatin-S	372,4885143
High	P09228	Cystatin-SA	371,9751173
High	P01037	Cystatin-SN	317,3369231
High	P54108	Cysteine-rich secretory protein 3	74,04
High	Q9UGM3	Deleted in malignant brain tumors 1 protein	102,56
High	Q02487	Desmocollin-2	28,22
High	P32926	Desmoglein-3	38,52
High	Q01469	Fatty acid-binding protein, epidermal	68,92
High	P02671	Fibrinogen alpha chain	36,54
High	P02675	Fibrinogen beta chain	101,6238405
High	P02679	Fibrinogen gamma chain	102,1
High	P04075	Fructose-bisphosphate aldolase A	161,88
High	Q08380	Galectin-3-binding protein	209,557206

High	P06396	Gelsolin	170,9850497
High	P06744	Glucose-6-phosphate isomerase	136,8742857
High	P04406	Glyceraldehyde-3-phosphate dehydrogenase	240,2898114
High	P00738	Haptoglobin	142,2766667
High	P0DMV8	Heat shock 70 kDa protein 1B	200,6130769
High	P69905	Hemoglobin subunit alpha	57,41
High	P02042	Hemoglobin subunit delta	147,7996348
High	P02790	Hemopexin	51,69
High	Q9Y6R7	IgGFc-binding protein	254,7
High	P01876	Immunoglobulin heavy constant alpha 1	408,17
High	P01877	Immunoglobulin heavy constant alpha 2	305,34
High	P01857	Immunoglobulin heavy constant gamma 1	168,8
High	P01859	Immunoglobulin heavy constant gamma 2	142,49
High	P01871	Immunoglobulin heavy constant mu	341,6027185
High	A0A0B4J1V0	Immunoglobulin heavy variable 3-15	358,9103959
High	P01591	Immunoglobulin J chain	169,2304639
High	P01834	Immunoglobulin kappa constant	64,82
High	P01619	Immunoglobulin kappa variable 3-20	47,07
High	P0DOY2	Immunoglobulin lambda constant 2	247,9284984
High	P0DOY3	Immunoglobulin lambda constant 3	114,280401
High	P01701	Immunoglobulin lambda variable 1-51	161,9905653
High	B9A064	Immunoglobulin lambda-like polypeptide 5	96,85
High	P06870	Kallikrein-1	82,64
High	P13645	Keratin, type I cytoskeletal 10	475,1444714
High	P02533	Keratin, type I cytoskeletal 14	134,4628173
High	Q9C075	Keratin, type I cytoskeletal 23	42,39
High	P35527	Keratin, type I cytoskeletal 9	185,5507679
High	P04264	Keratin, type II cytoskeletal 1	290,1566986
High	P35908	Keratin, type II cytoskeletal 2 epidermal	168,0965953
High	P02538	Keratin, type II cytoskeletal 6A	37,39
High	P22079	Lactoperoxidase	479,4567617
High	P02788	Lactotransferrin	1246,756336
High	P30740	Leukocyte elastase inhibitor	318,8874374
High	P09960	Leukotriene A-4 hydrolase	108,2244315
High	P31025	Lipocalin-1	280,9261816
High	P00338	L-lactate dehydrogenase A chain	159,8009305
High	P61626	Lysozyme C	208,3774922
High	P14780	Matrix metalloproteinase-9	176,005
High	P01033	Metalloproteinase inhibitor 1	63,75
High	P26038	Moesin	173,53
High	P08571	Monocyte differentiation antigen CD14	27,54849354
High	Q9HC84	Mucin-5B	271,93
High	Q8TAX7	Mucin-7	43,67
High	P05164	Myeloperoxidase	201,7628571
High	Q8NF91	Nesprin-1	32,82

High	P59665	Neutrophil defensin 1	101,3861761
High	P08246	Neutrophil elastase	105,6942916
High	P80188	Neutrophil gelatinase-associated lipocalin	139,09
High	Q9BVI4	Nucleolar complex protein 4 homolog	29,03
High	P04746	Pancreatic alpha-amylase	1737,93344
High	O75594	Peptidoglycan recognition protein 1	73,03
High	P23284	Peptidyl-prolyl cis-trans isomerase B	30,35
High	P00558	Phosphoglycerate kinase 1	139,2319678
High	P18669	Phosphoglycerate mutase 1	69,18
High	P55058	Phospholipid transfer protein	67,22
High	P13796	Plastin-2	250,3102732
High	P01833	Polymeric immunoglobulin receptor	913,9254512
High	P07737	Profilin-1	175,7596556
High	P12273	Prolactin-inducible protein	139,4102617
High	P07237	Protein disulfide-isomerase	98,32
High	Q6P5S2	Protein LEG1 homolog	281,1102513
High	P05109	Protein S100-A8	131,7166667
High	P06702	Protein S100-A9	54,32
High	P00491	Purine nucleoside phosphorylase	49,04
High	P15153	Ras-related C3 botulinum toxin substrate 2	85,57
High	P52566	Rho GDP-dissociation inhibitor 2	81,18322039
High	Q13315	Serine-protein kinase ATM	29,03
High	P29508	Serpin B3	88,86934799
High	P02768	Serum albumin	1401,968503
High	Q14515	SPARC-like protein 1	59,36
High	O00391	Sulfhydryl oxidase 1	63,06
High	P37837	Transaldolase	107,0649037
High	P20061	Transcobalamin-1	68,68779682
High	P29401	Transketolase	180,9484928
High	P60174	Triosephosphate isomerase	278,7607204
High	P02774	Vitamin D-binding protein	42,66666667
High	P25311	Zinc-alpha-2-glycoprotein	444,7235101
High	Q96DA0	Zymogen granule protein 16 homolog B	287,6428617

Table 4. *Number of proteins identified in washing solutions (W) and in trypsin shaving-off solutions or "pellicle samples" (P) for both cell types TR146 and TR146/MUC1. It is indicated also the number of proteins identified in the samples that are common to saliva.*

	W-TR146	P-TR146	W-TR146/MUC1	P-TR146/MUC1
Proteins identified (total)	99	619	152	587
Proteins in common with saliva	42	23	46	21

Table 5. *Detection in washing solutions (W) and in trypsin shaving-off solutions or "pellicle samples" (P) of proteins identified in the pool of saliva used to create a pellicle layer on TR146 or TR146/MUC1 cells. Detection / non-detection of the proteins are indicated by the + and – signs, respectively. In this table are reported only the proteins that have been detected in at least one of the four types of samples (W-TR146, P-TR146, W-TR146/MUC1, P-TR146/MUC1), i.e. 57 proteins out of the initial 118 proteins identified in saliva. Proteins which have likely a glandular origin are shaded in grey.*

Proteins identified in saliva		W TR146	P TR146	W TR146/MUC1	P TR146/MUC1
Accession	**Name**				
Detected in W only					
P60709	Actin, cytoplasmic 1	+	-	+	-
Q01518	Adenylyl cyclase-associated protein 1	-	-	+	-
P04745	Alpha-amylase 1	+	-	+	-
P20160	Azurocidin	+	-	+	-
P07339	Cathepsin D	-	-	+	-
P01036	Cystatin-S	+	-	-	-
P09228	Cystatin-SA	-	-	+	-
P01037	Cystatin-SN	-	-	+	-
P06744	Glucose-6-phosphate isomerase	+	-	-	-
P01876	Immunoglobulin heavy constant alpha 1	+	-	-	-
P01871	Immunoglobulin heavy constant mu	+	-	-	-
A0A0B4J1V0	Immunoglobulin heavy variable 3-15	+	-	+	-
P01591	Immunoglobulin J chain	+	-	+	-
P0DOY2	Immunoglobulin lambda constant 2	-	-	+	-
P01701	Immunoglobulin lambda variable 1-51	+	-	-	-
P35908	Keratin, type II cytoskeletal 2 epidermal	-	-	+	-
P22079	Lactoperoxidase	+	-	+	-
P02788	Lactotransferrin	+	-	+	-
P30740	Leukocyte elastase inhibitor	+	-	+	-
P31025	Lipocalin-1	-	-	+	-
P05164	Myeloperoxidase	+	-	+	-
P01833	Polymeric immunoglobulin receptor	+	-	+	-
Q6P5S2	Protein LEG1 homolog	+	-	+	-
P52566	Rho GDP-dissociation inhibitor 2	-	-	+	-
P29508	Serpin B3	+	-	+	-
P02768	Serum albumin	+	-	+	-
P60174	Triosephosphate isomerase	+	-	+	-
P25311	Zinc-alpha-2-glycoprotein	+	-	+	-
Q96DA0	Zymogen granule protein 16 homolog B	+	-	+	-
Detected in W and P					
P06733	Alpha-enolase	+	+	+	+
P04083	Annexin A1	+	+	+	+
P63104	14-3-3 protein zeta/delta	+	-	+	+
Q96DR5	BPI fold-containing family A member 2	+	+	+	+
P23528	Cofilin-1	+	+	+	+
P04075	Fructose-bisphosphate aldolase	+	+	+	+
P04406	Glyceraldehyde-3-phosphate dehydrogenase	+	+	+	+
P0DMV9	Heat shock 70 kDa protein 1B	+	+	+	-

P13645	Keratin, type I cytoskeletal 10	+	+	+	+
P02533	Keratin, type I cytoskeletal 14	+	+	+	+
P35527	Keratin, type I cytoskeletal 9	+	+	+	-
P04264	Keratin, type II cytoskeletal 1	+	+	+	-
P02538	Keratin, type II cytoskeletal 6A	+	-	+	+
P00338	L-lactate dehydrogenase A chain	+	+	+	+
P26038	Moesin	+	+	+	+
P00558	Phosphoglycerate kinase 1	+	+	+	+
P18669	Phosphoglycerate mutase 1	+	-	+	+
P13796	Plastin-2	+	+	+	+
P07737	Profilin-1	+	+	+	-
P07237	Protein disulfide-isomerase	+	+	+	-
P37837	Transaldolase	-	+	+	+
P29401	Transketolase	+	+	+	-
Detected in P only					
Q8TDL5	BPI fold-containing family B member 1	-	+	-	+
P04080	Cystatin-B	-	+	-	-
Q8TAX7	Mucin-7	-	+	-	+
P23284	Peptidyl-prolyl cis-trans isomerase B	-	+	-	+
P00491	Purine nucleoside phosphorylase	-	-	-	+

Figure 14. *Principal component analysis performed on the mass spectrometry data obtained for pellicle samples. "DP" stands for digested pellicle, so DP from 1 to 3 are the replicas of P-TR146 samples while DP from 4 to 6 are the ones for P-TR146/MUC1 samples.*

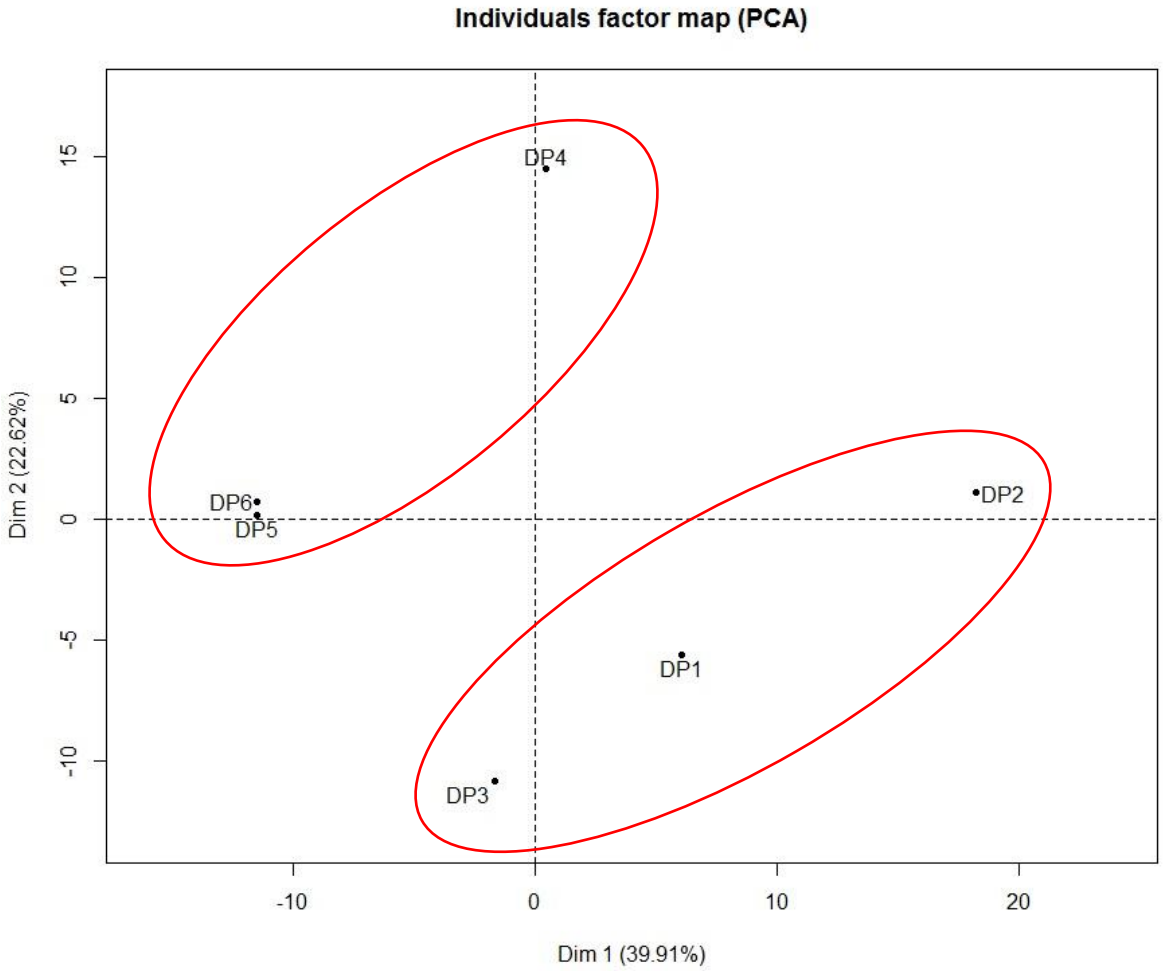


Figure 15. *Hierarchical classification performed on the replicas of pellicle samples obtained for P-TR146 and P-TR146/MUC1. The classification divided the replicas into two main clusters suggesting that there is a difference in proteome between the two cell lines.*

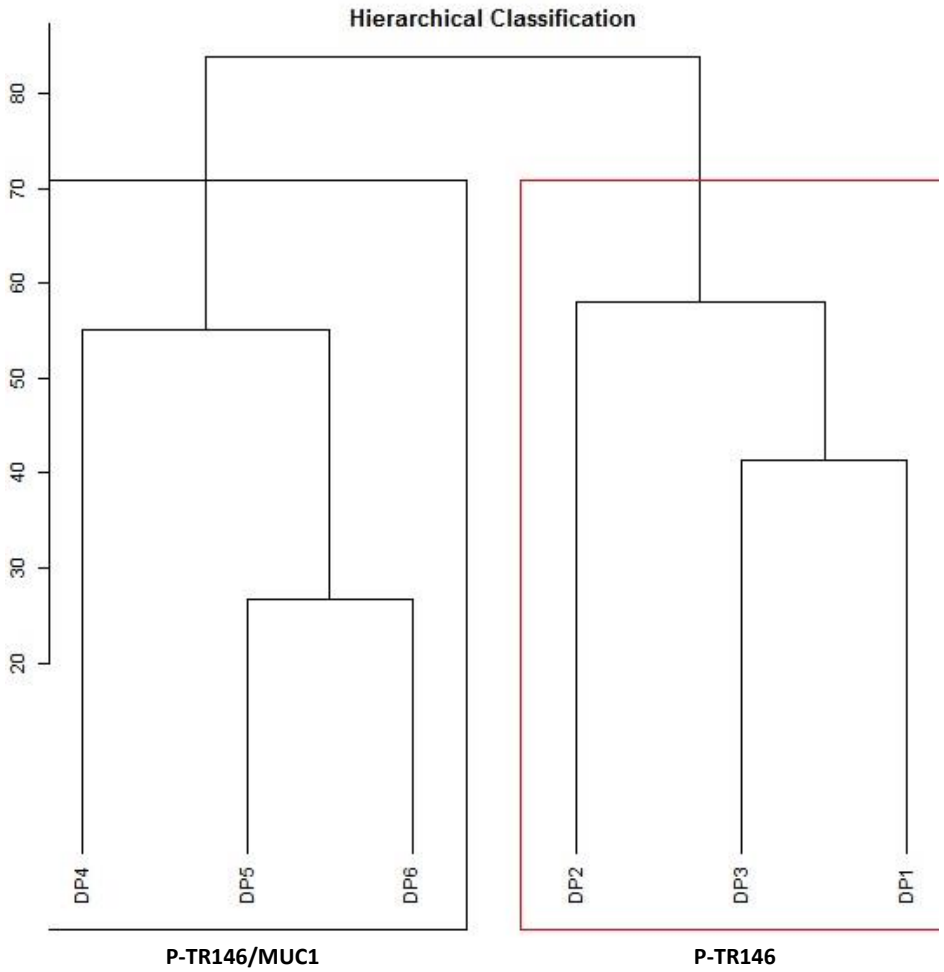
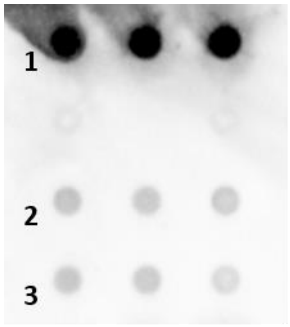


***Figure 16**. Dot blots performed to detect the presence of MUC 5B which was not detected by mass spectrometry analysis. 20 µg of proteins were loaded onto a 0.45 µm Nitrocellulose membrane (Bio-rad). For proteins detection the following primary antibodies produced in mouse were used: anti-MUC5B (F2, kind gift from Prof Veerman, Free university of Amsterdam, 1:400). The secondary antibody was goat-anti mouse IgG coupled to HRP (LifeTechnologies, 1:2000).*



1: Saliva pool

2: P-TR146

3: P-TR146/MUC1

5 Discussion

Isolation of the mucosal pellicle from epithelial cells is a difficult task to perform. Starting from literature search we opted at the beginning to follow the work performed *in vivo* by Gibbins et al. [78] in which buccal cells were collected and washed with different solution to see which salivary proteins were retained on them. Our first protocol consisted in using 24 well plates (then changed with 12 well plates, to have a larger sample volume) in which the pellicle was constituted on TR146 or TR146/MUC1 cells, and then washing the cells with increasing concentration of SDS and also Tween-20 which is a more delicate detergent, to tentatively separate the pellicle from oral cells. Tween-20 was not able to separate the pellicle from the cells, while SDS was able to separate the pellicle but caused also extensive cellular lysis. Then we tried to directly digest the pellicle formed onto the cells by using trypsin. This protocol has been chosen to be the best for our purpose since it can separate the pellicle with a minor cellular lysis compared to SDS. Results summarized in Table 5, obtained on the characterization of the mucosal pellicle led to evidence the presence of proteins found in saliva that could be adsorbed in the pellicle or not (due to their presence only in W and not in P samples). Of all these proteins, glandular proteins are of special interest since their external source enables to determine their capacity of adsorption on the cellular surface. 16 proteins marked as of "glandular origin" were detected only in W samples, therefore easily removed by PBS washing. These proteins are α-amylase 1, cathepsin D, cystatin S/SA/SN, various IgA parts, lactoperoxidase, lactoransferrin, pIgR, protein LEG1 homolog, zinc-α-2 glycoprotein and zymogen granule protein 16 homolog. By searching in the literature, all these proteins were reported in proteomic studies to be part of the acquired enamel pellicle (AEP) [150–152]. A proteomic study performed by Delius et al. [153] aimed to show which salivary proteins were enriched, depleted or proportionally adsorbed in AEP with respect to saliva.

On this regard our work evidenced that eight glandular washed-away proteins were in common with those reported by Delius, who described them as proportionally adsorbed (Lactoperoxidase and Zymogen granule protein 16 homolog B) or depleted (Alpha-amylase 1, Zinc-alpha-2-glycoprotein, Cystatin-SA, Polymeric immunoglobulin receptor, Immunoglobulin J chain and Immunoglobulin heavy constant alpha 1) in AEP. On the contrary, data on the presence of these proteins in the mucosal pellicle are poor: α-amylase was detected on scraped-off buccal cells washed in a buffer with the same ionic strength as saliva, which the authors interpreted as a clue that this protein was part of the mucosal pellicle [133]. In a recent study α-amylase displayed minimal binding to buccal cells and was rapidly removed by different types of washing solutions [78]. The same conclusions apply to the case of Cystatin S. Our results suggest that the 16 proteins originating from saliva detected only in W samples and marked as glandular are not highly abundant in the *in vitro* constituted mucosal pellicle.

Still considering glandular proteins, 5 of them were detected in P samples, suggesting a strong binding to the pellicle. These proteins are BPI fold-containing family A member 2, protein disulphide-isomerase, BPI fold-containing family B member 1, mucin 7 and peptidyl-prolyl-cis-trans isomerase B. The latter three proteins were detected only in P samples, suggesting selective adsorption on the pellicle. Moreover, protein disulphide-isomerase was identified in P-TR146 but not in P-TR146/MUC1. It should be noted that although mass spectrometry analysis did not report them, through immunoblots we were able to identify in P samples also Mucin 5B, pIgR and S-type cystatins. Their absence in mass spectrometry analysis could be because of their lower concentration compared to the other proteins identified.

The only proteins of glandular origin reported to bind strongly to epithelial cells [78] and thus to be part of the mucosal pellicle [141] are Mucin 5B (MUC5B) and Mucin 7 (MUC7). In the oral cavity several mucins can be detected: MUC5B, MUC7, MUC19, MU1 and MUC4 [154].

Among these, MUC5B and MUC7 are principal components of the mucosal pellicle [4,78,134] along with membrane-associated MUC1 which has a more structural function in the pellicle [140,143]. MUC5B is a gel forming mucin secreted by mucous cells of submandibular, sublingual, palatine and labial salivary glands [155,156]. MUC5B is organized in a structure comprising an N-terminus, central glycosylated region and a C-terminus [157–159]. The central glycosylated region contains tandem repeats of 29 amino acids rich in Ser and Thr residues [158]. MUC5B monomers tend to form dimers by disulfide bonds between the C-terminal regions, and then polymers by disulfide bonds at the N-terminus [160,161]. MUC5B is secreted in a granule and stabilized by divalent calcium ions that are exchanged with monovalent sodium ions, increasing osmotic pressure and leading to hydration and expansion of polymers with the formation of the gel [159,162]. MUC7 structure contains tandem repeats of 23 amino acids [163], and lacks a terminal cysteine rich domain, resulting in the incapacity of MUC7 to form polymers [164].

These structural differences between MUC5B and MUC7 make these two proteins perform the same functions differently. MUC5B and MUC7 can bind salivary proteins that exhibit antibacterial activity. In detail, based on an interaction screening with a library of submandibular gland proteins, MUC7 with its N-terminal domain can bind to aPRPs, bPRPs, statherins and histatin 1 [165]. MUC5B can bind the same proteins forming heterocomplexes, as it has been shown in the work of Iontcheva et al. [166]. MUC5B and MUC7 binding to antimicrobial proteins can result in transport of these proteins across the oral cavity, increase in their concentration and persistence on the mucosa and protection against their degradation. MUC5B and MUC7 can also bind to some bacterial strains and facilitate their removal from the oral cavity. Sialic acid residues on MUC7 glycans are responsible for binding some *Streptococci* strains *(S. gordonii* and *S. sanguinis*), in fact when sialic acid is removed with neuraminidase the binding of MUC7 to this bacteria is reduced [167,168].

MUC7 can also bind to *S. gordonii* through surface proteins expressed on bacterial surface: alpha-enolase, EF-G, EF-tu and oligopeptide-binding protein [169]. It has been demonstrated that MUC7 can bind also to *Escherichia Coli* and *Staphylococcus Aureus* [170,171]. MUC5B, unlike MUC7, binds a limited number of oral pathogens. *Haemophilus parainfluenzae* can bind to the polypeptide chain of MUC5B [172], while *Helycobater pylori* binds MUC5B through a neutrophil-activating protein that recognizes sulfated glycans on this mucin [173]. Another work has demonstrated that MUC5B prevents *Streptococcus mutans* colonization by keeping the bacterium in the planktonic state [174]. Thus, MUC5B protective effects can reside in the gel forming ability of this mucin, since the gel layer can repel bacteria from oral surfaces.

In this work, we report the presence in the pellicle samples of BPI fold-containing family (BPIF) A member 2 (SPLUNC2) and BPI fold-containing family B member 1 (LPLUNC1), which were not previously reported to be part of the pellicle. The first one has been identified in W and P samples of both cell lines while the other has been identified only in P samples of both cell lines, suggesting that this is more retained on the pellicle compared to the first one. BPIF proteins belong to the PLUNC (palate, lung and nose epithelium clone) protein family. This protein family is structurally and functionally related to BPI (bactericidal permeability increasing protein) and LPB (lipopolysaccharide binding protein) and consists of eight members divided in two groups based on their size [175]. The original nomenclature proposed for these proteins divided them into short or long proteins, thus having SPLUNC1, SPLUNC2, SPLUNC3, LPLUNC1, LPLUNC2, LPLUNC3, LPLUNC4 and LPLUNC6. The distinction between short and long proteins is based on the presence of BPI domain. BPI is a protein constituted by two domains with the same fold, that dock onto each other and are linked by a central β-sheet [176]. Therefore, a PLUNC protein is defined long if it presents the two BPI fold domains, whereas if it presents only one fold domain it is defined as short [175].

Nowadays, the nomenclature for these proteins has been changed: short proteins are grouped into family A while long proteins are part of family B and the name PLUNC has been replaced with BPIF to include the presence of the BPI fold-containing superfamily. So today we refer to the eight PLUNC proteins as BPIFA1 (SPLUNC1), BPIFA2 (SPLUNC2), BPIFA3 (SPLUNC3), BPIFB1 (LPLUNC1), BPIFB2 (LPLUNC2), BPIFB3 (LPLUNC3), BPIFB4 (LPLUNC4) and BPIFB6 (LPLUNC6) [177]. As already said, in this work we identified BPIFA2 and BPIFB1, with BPIFB1 detected only in P samples. BPIFA2 is found only in oral cavity and is expressed by serous cells of major and minor salivary glands, with predominant concentration in parotid gland [178], while BPIFB is expressed in upper airways, nasal epithelium and salivary glands [179]. The exact biological role of these two proteins is not currently known, even though their similarity with BPI and LPB suggest that they have an implication in host defence, especially in oral cavity and upper airways, by binding some bacterial strains and viruses [171,180,181].

Host defence is a crucial process for the organism, especially in mucosal regions that are the first barriers against pathogens or toxic substances. Therefore, in addition to BPIFA2 and BPIFB1, the presence of Cystatins in pellicle samples is not surprising. Cystatins are a family of proteins with the main function of inhibiting the cysteine proteinases [30]. It has been demonstrated that Cystatins have also antimicrobial function in human organisms, including the oral cavity [182,183]. In our work we detected Cystatin B (which is not of glandular origin) only in P-TR146, and S-type Cystatins (S, SA and SN) in P-TR146 and P-TR146/MUC1. Beside this important involvement in host defence, their activity as protease inhibitors in the mucosal pellicle may have implications also for taste perception. In fact, the presence of Cystatins in saliva could affect the integrity of the mucosal pellicle by protecting it from proteolysis.

If the pellicle is disrupted by proteolytic action, receptors for taste compounds could be more accessible giving hyper-sensitivity phenomena, while with the reduction of proteolysis mediated by Cystatins, the intact pellicle could mask these receptors [184].

Secretory component (SC) of Polymeric Immunoglobulin Receptor (pIgR) was detected by dot blot in P samples of both cell lines used in our work. SC is released from pIgR upon a proteolytic cleavage in the apical part of cell membrane, and it can be free or bound to Immunoglobulin A (IgA) which is the most abundant immunoglobulin in mucosa [185]. The main function of SC is to bind IgA (sIgA) and protect it from proteolysis, but it has been demonstrated that SC alone can act as a microbial scavenger [185]. SC has been previously detected in the mucosal pellicle, as well as IgA which seems concentrated in the pellicle respect to saliva. The concentration of IgA in the pellicle seems to be mediated either by SC and by mucin-mucin interactions [78].

The other two proteins identified in our work and marked as of "glandular origin" are peptidyl-prolyl cis-trans isomerase B and protein disulfide-isomerase. Peptidyl-prolyl cis-trans isomerase catalyzes the cis-trans isomerization of proline imidic peptide bonds in oligopeptides [186] whereas protein disulfide-isomerase catalyzes the formation, breakage and rearrangement of disulfide bonds. At the cell surface, it can act as a reductase that cleaves disulfide bonds of proteins attached to the cell [187]. Considering these structural functions, their presence could help in the formation or organization of the pellicle rather than in the biological processes in which the pellicle is involved. Finally, we identified Annexin A1 in W and P samples, that is not of glandular origin but could be secreted in saliva and has interesting functions that could be relevant in the mucosal pellicle. Annexin A1 is member of the annexin protein family. Proteins of this family bind calcium, and have a C-terminal domain which is conserved among members of the family [113].

The N-terminal domain is different for every annexin and determines the functions of these proteins. Annexin A1 has a N-terminal domain of 40 residues, of which the first 10-14 residues represent a binding site for S100 A11 protein [113]. This binding site is usually folded in a helix packed in the core of the protein. After calcium binding, the helix is freed from the protein core, and thus available for ligand binding. This conformational change seems to be also a requisite for the membrane aggregation activity of Annexin A1 [113]. Annexin A1 is abundant in neutrophils, especially in the cytoplasm, but upon cell activation it is translocated on cell surface [188]. Originally Annexin A1 was identified as the inhibitor of phospholipase A2 [189], suggesting is involvement in inflammatory processes mediated by prostaglandins. Recently, has been investigated the role of Annexin A1 in epithelial wound repair. Annexin A1 and its cleavage product Ac2-26 exert this function by binding to Formyl Peptide Receptors (FRPs), in particular to FPR1 and FPR2/ALX [190]. Annexin A1 induces homo-dimerization of FRP2/ALX and activation of p38 mitogen-activated protein kinase signalling pathway, while Ac2-26 induces hetero-dimerization of FRP1 and FRP2/ALX and activation of c-Jun N-terminal kinase signalling pathway [191]. Endogenous Annexin A1 is released as a component of extracellular vesicles (EVs) originated from epithelial intestinal cells, and exerts paracrine and autocrine effects on the epithelium to facilitate mucosal wound repair processes [192–195]. In our study on the mucosal pellicle, Annexin A1 was detected in W and P samples of both cell lines, suggesting that it is present in the pellicle but not strongly retained on it. Even though Annexin A1 wound repair functions have been observed in other tissues, this protein could have this function also in the mucosal pellicle, since it is a biological barrier often exposed to traumatic events that can lead to injuries and inflammation.

6 Conclusions

In this work, for the first time a trypsin “shaving protocol” has been optimized for the isolation and proteomic characterization of the oral mucosal pellicle constitute on the cell lines TR146 and TR146/MUC1. The mucosal pellicle characterization is a difficult task to perform, due to difficulties in its isolation from the buccal cells. Most studies are therefore performed *in vivo* or on artificial supports. However, the use of a cellular model could bring benefits and less ethical concerns in this kind of research. The aim of this work was in the first step to isolate and characterize the protein composition of the mucosal pellicle formed on both cell lines utilized, and secondly to compare pellicle proteome with the salivary proteome. The results obtained are overall similar to those obtained by other research groups, that performed *in vivo* studies, but also, we provided novel information on the pellicle composition. We were able to confirm the presence of proteins reported to be specific components of the mucosal pellicle like salivary MUC5B and 7, and other proteins like SC and Cystatins. Moreover, we reported the presence of BPIFA2 and BPIFB1, which were not previously reported as pellicle components. For the proteins identified we were also able to show if they are strongly retained or not on the pellicle, by analysing their presence in the PBS wash solution collected during the pellicle isolation procedure. Furthermore, statistical analyses performed on mass spectrometry data obtained revealed that there is a different pellicle proteome between TR146 and TR146/MUC1 cell lines.

7 Future perspectives

This thesis has further shown how *bottom-up* proteomics is a solid approach to study salivary proteome. Saliva has proved itself to be an extremely important fluid for its potential, through its protein composition, to be a source of biomarkers under different pathophysiological condition. Moreover, saliva constituting the mucosal pellicle has a central role in maintaining the homeostasis in the oral cavity, thus the comprehension of salivary proteins involved in mucosal pellicle formation is essential.

Regarding SCZ, *bottom-up* approach allowed to characterize the portion of salivary proteome lost due to sample treatment during *top-down* analysis of saliva. Although the results obtained are good and there are some differentially expressed proteins particularly interesting, it is only a starting point. Proteins differentially expressed needs to be validated by immunoblots in the future. On the other hand, the use of a *bottom-up* approach for exploration of the mucosal pellicle composition provides a good starting point for the characterization of the mucosal pellicle. However, it needs to be optimized by improving the isolation of the pellicle maybe using different proteolytic enzymes or other reagents to separate the pellicle layer without damaging the cells. It would be also interesting to better characterize the differences in the pellicle proteomes obtained for the two cell lines utilised.

Acknowledgements

This thesis has been produced during the PhD course in Life, Environmental and Drug Sciences, XXXII cycle, with the financial support from Sardinian Regional Government scholarship (P.O.R. Sardegna F.S.E. - Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2014-2020 - Axis III Education and training, Thematic goal 10, Investment Priority 10ii), Specific goal 10.5, and with further financial support from the European Erasmus+ program during the mobility period at the Centre des Sciences du Goût et de l'Alimentation, (CSGA) AgroSup Dijon, CNRS, INRAE, Université de Bourgogne Franche-Comté, Dijon, France.

Bibliography

1. Edgar. *Saliva and Oral Health ed*. (1996).

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

3. Collins, L. M. C. & Dawes, C. The Surface Area of the Adult Human Mouth and Thickness of the Salivary Film Covering the Teeth and Oral Mucosa. *J. Dent. Res.* **66**, 1300–1302 (1987).

4. Morzel, M., Siying, T., Brignot, H. & Lherminier, J. Immunocytological detection of salivary mucins (MUC5B) on the mucosal pellicle lining human epithelial buccal cells. *Microsc. Res. Tech.* **77**, 453–457 (2014).

5. Mescher, A. L. *Junqueira's Basic Histology: Text and Atlas, Fifteenth Edition*. (McGraw-Hill Education, 2018).

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

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

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

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

10. Hay, D. I., Carlson, E. R., Schluckebier, S. K., Moreno, E. C. & Schlesinger, D. H. Inhibition of calcium phosphate precipitation by human salivary acidic proline-rich proteins: Structure-activity relationships. *Calcif. Tissue Int.* **40**, 126–132 (1987).

11. Lyons, K. M., Stein, J. H. & Smithies, O. *Length Polymorphisms in Human Proline-Rich Protein Genes Generated by Intragenic Unequal Crossing Over*. (1988).

12. Azen, E. A., Amberger, E., Fisher, S., Prakobphol, A. & Niece, R. L. PRB1, PRB2, and PRB4 coded polymorphisms among human salivary concanavalin-A binding, II-1, and Po

proline-rich proteins. *Am. J. Hum. Genet.* **58**, 143 (1996).

13. Cabras, T. *et al.* HPLC-ESI-MS and MS/MS structural characterization of multifucosylated N-glycoforms of the basic proline-rich protein IB-8a CON1+ in human saliva. *J. Sep. Sci.* **35**, 1079–86 (2012).

14. Manconi, B. *et al.* N- and O-linked glycosylation site profiling of the human basic salivary proline-rich protein 3M. *J. Sep. Sci.* **39**, 1987–1997 (2016).

15. Levine, M. J. *et al.* Structural Aspects of Salivary Glycoproteins. *J. Dent. Res.* **66**, 436–441 (1987).

16. Ployon, S. *et al.* Mechanisms of astringency: Structural alteration of the oral mucosal pellicle by dietary tannins and protective effect of bPRPs. *Food Chem.* **253**, 79–87 (2018).

17. Könönen, E., Kanervo, A., Takala, A., Asikainen, S. & Jousimies-Somer, H. Establishment of oral anaerobes during the first year of life. *J. Dent. Res.* **78**, 1634–9 (1999).

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

19. Hirtz, C. *et al.* MS characterization of multiple forms of alpha-amylase in human saliva. *Proteomics* **5**, 4597–4607 (2005).

20. Peyrot des Gachons, C. & Breslin, P. A. S. Salivary Amylase: Digestion and Metabolic Syndrome. *Current Diabetes Reports* **16**, (2016).

21. Parkkila, S. *et al.* Salivary carbonic anhydrase protects gastroesophageal mucosa from acid injury. *Dig. Dis. Sci.* **42**, 1013–1019 (1997).

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

23. Thatcher, B. J., Doherty, A. E., Orvisky, E., Martin, B. M. & Henkin, R. I. Gustin from human parotid saliva is carbonic anhydrase VI. *Biochem. Biophys. Res. Commun.* **250**, 635–641 (1998).

24. Silver, L. Immunoglobulin Gene Superfamily. in *Encyclopedia of Genetics* 998–999

(Elsevier, 2001). doi:10.1006/rwgn.2001.0669

25. Fábián, T. K., Hermann, P., Beck, A., Fejérdy, P. & Fábián, G. Salivary defense proteins: Their network and role in innate and acquired oral immunity. *International Journal of Molecular Sciences* **13**, 4295–4320 (2012).

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

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

28. Brandtzaeg, P. Do salivary antibodies reliably reflect both mucosal and systemic immunity? in *Annals of the New York Academy of Sciences* **1098**, 288–311 (Blackwell Publishing Inc., 2007).

29. Frenkel, E. S. & Ribbeck, K. Salivary mucins in host defense and disease prevention. *J. Oral Microbiol.* **7**, 29759 (2015).

30. Dickinson, D. P. Salivary (SD-type) cystatins: over one billion years in the making--but to what purpose? *Crit. Rev. Oral Biol. Med.* **13**, 485–508 (2002).

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

32. Cabras, T. *et al.* RP-HPLC-ESI-MS evidenced that salivary cystatin B is detectable in adult human whole saliva mostly as S-modified derivatives: S-Glutathionyl, S-cysteinyl and S-S 2-mer. *J. Proteomics* **75**, 908–913 (2012).

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

34. Lupi, A. *et al.* Identification of the human salivary cystatin complex by the coupling of high-performance liquid chromatography and ion-trap mass spectrometry. *Proteomics* **3**, 461–467 (2003).

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

36. Balbín, M. *et al.* A sequence variation in the human cystatin D gene resulting in an amino acid (Cys/Arg) polymorphism at the protein level. *Hum. Genet.* **90**, 668–9 (1993).

37. Alvarez-Díaz, S. *et al.* Cystatin D is a candidate tumor suppressor gene induced by vitamin D in human colon cancer cells. *J. Clin. Invest.* **119**, 2343–58 (2009).

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

39. Ferrer-Mayorga, G. *et al.* Cystatin D locates in the nucleus at sites of active transcription and modulates gene and protein expression. *J. Biol. Chem.* **290**, 26533–48 (2015).

40. Ahmad, M., Piludu, M., Oppenheim, F. G., Helmerhorst, E. J. & Hand, A. R. Immunocytochemical Localization of Histatins in Human Salivary Glands. *J. Histochem. Cytochem.* **52**, 361–370 (2004).

41. OPPENHEIM, F. G., SALIH, E., SIQUEIRA, W. L., ZHANG, W. & HELMERHORST, E. J. Salivary Proteome and Its Genetic Polymorphisms. *Ann. N. Y. Acad. Sci.* **1098**, 22–50 (2007).

42. Oppenheim, F. G. *et al.* Histatins, a novel family of histidine-rich proteins in human parotid secretion. Isolation, characterization, primary structure, and fungistatic effects on Candida albicans. *J. Biol. Chem.* **263**, 7472–7 (1988).

43. Castagnola, M. *et al.* A cascade of 24 histatins (histatin 3 fragments) in human saliva. Suggestions for a pre-secretory sequential cleavage pathway. *J. Biol. Chem.* **279**, 41436–43 (2004).

44. Troxler, R. F., Offner, G. D., Xu, T., Vanderspek, J. C. & Oppenheim, F. G. Structural Relationship Between Human Salivary Histatins. *J. Dent. Res.* **69**, 2–6 (1990).

45. Driscoll, J. *et al.* Functional Comparison of Native and Recombinant Human Salivary Histatin 1. *J. Dent. Res.* **74**, 1837–1844 (1995).

46. Cabras, T. *et al.* HPLC-MS characterization of cyclo-statherin Q-37, a specific cyclization product of human salivary statherin generated by transglutaminase 2. *J. Sep. Sci.* **29**,

2600–8 (2006).

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

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

49. Valore, E. & Ganz, T. Posttranslational processing of defensins in immature human myeloid cells. *Blood* **79**, (1992).

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

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

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

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

54. Singh, P. K. *et al.* Production of beta-defensins by human airway epithelia. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14961–6 (1998).

55. Inzitari, R. *et al.* HPLC-ESI-MS analysis of oral human fluids reveals that gingival crevicular fluid is the main source of oral thymosins β_4 and β_{10}. *J. Sep. Sci.* **32**, 57–63 (2009).

56. Huff, T., Müller, C. S., Otto, A. M., Netzker, R. & Hannappel, E. beta-Thymosins, small acidic peptides with multiple functions. *Int. J. Biochem. Cell Biol.* **33**, 205–20 (2001).

57. Hannapel, E. beta-Thymosins. *Ann. N. Y. Acad. Sci.* **1112**, 21–37 (2007).

58. Hannappel, E. Thymosin beta4 and its posttranslational modifications. *Ann. N. Y. Acad. Sci.* **1194**, 27–35 (2010).

59. Marenholz, I., Heizmann, C. W. & Fritz, G. S100 proteins in mouse and man: from evolution to function and pathology (including an update of the nomenclature). *Biochem. Biophys. Res. Commun.* **322**, 1111–1122 (2004).

60. Ravasi, T. *et al.* Probing the S100 protein family through genomic and functional analysis. *Genomics* **84**, 10–22 (2004).

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

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

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

64. Thorey, I. S. *et al.* The Ca2+-binding proteins S100A8 and S100A9 are encoded by novel injury-regulated genes. *J. Biol. Chem.* **276**, 35818–25 (2001).

65. Donato, R. Intracellular and extracellular roles of S100 proteins. *Microsc. Res. Tech.* **60**, 540–551 (2003).

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

67. Moroz, O. V. *et al.* Both Ca2+and Zn2+are essential for S100A12 protein oligomerization and function. *BMC Biochem.* **10**, (2009).

68. Abe, T. *et al.* Expression of the secretory leukoprotease inhibitor gene in epithelial cells. *J. Clin. Invest.* **87**, 2207–15 (1991).

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

70. Farquhar, C. *et al.* Salivary secretory leukocyte protease inhibitor is associated with reduced transmission of human immunodeficiency virus type 1 through breast milk. *J. Infect. Dis.* **186**, 1173–6 (2002).

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

72. Gipson, T. S. *et al.* Regulatory effects of endogenous protease inhibitors in acute lung inflammatory injury. *J. Immunol.* **162**, 3653–62 (1999).

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

74. Mueller, A. M. *et al.* Novel role for SLPI in MOG-induced EAE revealed by spinal cord expression analysis. *J. Neuroinflammation* **5**, (2008).

75. Asano, M. & Komiyama, K. Polymeric immunoglobulin receptor. *Journal of oral science* **53**, 147–156 (2011).

76. Kaetzel, C. S. The polymeric immunoglobulin receptor: bridging innate and adaptive immune responses at mucosal surfaces. *Immunol. Rev.* **206**, 83–99 (2005).

77. Ramachandran, P. *et al.* Identification of N-Linked Glycoproteins in Human Saliva by Glycoprotein Capture and Mass Spectrometry. *J. Proteome Res.* **5**, 1493–1503 (2006).

78. Gibbins, H. L., Proctor, G. B., Yakubov, G. E., Wilson, S. & Carpenter, G. H. Concentration of salivary protective proteins within the bound oral mucosal pellicle. *Oral Dis.* **20**, 707–713 (2014).

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

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

81. Morzel, M. *et al.* Salivary protein profiles are linked to bitter taste acceptance in infants. *Eur. J. Pediatr.* **173**, 575–582 (2014).

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

83. Roi, A. *et al.* A New Approach for the Diagnosis of Systemic and Oral Diseases Based on Salivary Biomolecules. *Dis. Markers* **2019**, 8761860 (2019).

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

85. Chevalier, F. Highlights on the capacities of 'Gel-based' proteomics. *Proteome Science* **8**, (2010).

86. Walz, A. *et al.* Proteome analysis of glandular parotid and submandibular-sublingual saliva in comparison to whole human saliva by two-dimensional gel electrophoresis. *Proteomics* **6**, 1631–9 (2006).

87. Vitorino, R. *et al.* Identification of human whole saliva protein components using proteomics. *Proteomics* **4**, 1109–15 (2004).

88. Ghafouri, B., Tagesson, C. & Lindahl, M. Mapping of proteins in human saliva using two-dimensional gel electrophoresis and peptide mass fingerprinting. in *Proteomics* **3**, 1003–1015 (2003).

89. Jessie, K., Pang, W. W., Rahim, Z. H. A. & Hashim, O. H. Proteomic analysis of whole human saliva detects enhanced expression of interleukin-1 receptor antagonist, thioredoxin and lipocalin-1 in cigarette smokers compared to non-smokers. *Int. J. Mol. Sci.* **11**, 4488–4505 (2010).

90. Jessie, K. *et al.* Aberrant proteins in the saliva of patients with oral squamous cell carcinoma. *Electrophoresis* **34**, 2495–502 (2013).

91. Techatanawat, S. *et al.* Salivary and serum cystatin SA levels in patients with type 2 diabetes mellitus or diabetic nephropathy. *Arch. Oral Biol.* **104**, 67–75 (2019).

92. Arba, M. *et al.* Proteomic characterization of the acid-insoluble fraction of whole saliva from preterm human newborns. *J. Proteomics* **146**, 48–57 (2016).

93. Switzar, L., Giera, M. & Niessen, W. M. A. Protein digestion: An overview of the available techniques and recent developments. *Journal of Proteome Research* **12**, 1067–1077 (2013).

94. Freedman, R. Schizophrenia. *N. Engl. J. Med.* **349**, 1738–1749 (2003).

95. Müller, N., Weidinger, E., Leitner, B. & Schwarz, M. J. The role of inflammation in schizophrenia. *Front. Neurosci.* **9**, (2015).

96. Nascimento, J. M. & Martins-De-Souza, D. The proteome of schizophrenia. *npj Schizophr.* **1**, 1–11 (2015).

97. Iavarone, F. *et al.* Characterization of salivary proteins of schizophrenic and bipolar disorder patients by top-down proteomics. *J. Proteomics* **103**, 15–22 (2014).

98. Arba, M. *et al.* Comparative proteomics of acid-insoluble fraction of saliva in type 1 diabetes-celiac desease subjects, celiac desease or type 1 diabetes subjects. in *Italian Proteomics Association - XII Annual Conference (Lecce)* (2017).

99. Martins-de-Souza, D., Harris, L. W., Guest, P. C. & Bahn, S. The Role of Energy Metabolism Dysfunction and Oxidative Stress in Schizophrenia Revealed by Proteomics. *Antioxid. Redox Signal.* **15**, 2067–2079 (2010).

100. Martins-de-Souza, D. *et al.* Proteome analysis of the thalamus and cerebrospinal fluid reveals glycolysis dysfunction and potential biomarkers candidates for schizophrenia. *J. Psychiatr. Res.* **44**, 1176–1189 (2010).

101. Martins-de-Souza, D. *et al.* Proteome analysis of schizophrenia patients Wernicke's area reveals an energy metabolism dysregulation. *BMC Psychiatry* **9**, 1–8 (2009).

102. Tristan, C., Shahani, N., Sedlak, T. W. & Sawa, A. The diverse functions of GAPDH: Views from different subcellular compartments. *Cell. Signal.* **23**, 317–323 (2011).

103. Arif, A., Chatterjee, P., Moodt, R. A. & Fox, P. L. Heterotrimeric GAIT Complex Drives Transcript-Selective Translation Inhibition in Murine Macrophages. *Mol. Cell. Biol.* **32**, 24 (2012).

104. Wagener, J. *et al.* A peptide derived from the highly conserved protein GAPDH is involved in tissue protection by different antifungal strategies and epithelial immunomodulation HHS Public Access. *J Invest Dermatol* **133**, 144–153 (2013).

105. Tisdale, E. J. Glyceraldehyde-3-phosphate Dehydrogenase Is Phosphorylated by Protein Kinase C/ and Plays a Role in Microtubule Dynamics in the Early Secretory Pathway*. (2001). doi:10.1074/jbc.M109744200

106. Kumagai, H. & Sakai, H. A porcine brain protein (35 K protein) which bundles microtubules and its identification as glyceraldehyde 3-phosphate dehydrogenase. *J.*

Biochem. **93**, 1259–69 (1983).

107. Benitez-King, G., Ramírez-Rodríguez, G., Ortíz, L. & Meza, I. The neuronal cytoskeleton as a potential therapeutical target in neurodegenerative diseases and schizophrenia. *Curr. Drug Targets. CNS Neurol. Disord.* **3**, 515–33 (2004).

108. Samson, A. L. *et al.* Oxidation of an exposed methionine instigates the aggregation of glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **289**, 26922–26936 (2014).

109. Lehtinen, M. K. *et al.* Cystatin B deficiency sensitizes neurons to oxidative stress in progressive myoclonus epilepsy, EPM1. *J. Neurosci.* **29**, 5910–5 (2009).

110. Ulbrich, L. *et al.* Cystatin B and SOD1: Protein–Protein Interaction and Possible Relation to Neurodegeneration. *Cell. Mol. Neurobiol.* **34**, 205–213 (2014).

111. Prabakaran, S. *et al.* Mitochondrial dysfunction in schizophrenia: Evidence for compromised brain metabolism and oxidative stress. *Mol. Psychiatry* **9**, 684–697 (2004).

112. Jeffery, C. J. Moonlighting proteins: old proteins learning new tricks. *Trends Genet.* **19**, 415–417 (2003).

113. Gerke, V. & Moss, S. E. Annexins: From Structure to Function. (2002). doi:10.1152/physrev.00030.2001

114. Christensen, M. V., Høgdall, C. K., Umsen, K. M. J. & Høgdall, E. V. S. Annexin A2 and cancer: A systematic review. *International Journal of Oncology* **52**, 5–18 (2018).

115. Bharadwaj, A., Bydoun, M., Holloway, R. & Waisman, D. *Annexin A2 heterotetramer: Structure and function*. *International Journal of Molecular Sciences* **14**, (2013).

116. Madureira, P. A. & Waisman, D. M. Annexin A2: The importance of being redox sensitive. *Int. J. Mol. Sci.* **14**, 3568–3594 (2013).

117. Dietrich, D. *et al.* Interleukin-36 potently stimulates human M2 macrophages, Langerhans cells and keratinocytes to produce pro-inflammatory cytokines. *Cytokine* **84**, 88–98 (2016).

118. Frey, S. *et al.* The novel cytokine interleukin-36α is expressed in psoriatic and rheumatoid arthritis synovium. *Ann. Rheum. Dis.* **72**, 1569–1574 (2013).

119. Al-Hakeim, H. K., Al-Rammahi, D. A. & Al-Dujaili, A. H. IL-6, IL-18, sIL-2R, and TNFα proinflammatory markers in depression and schizophrenia patients who are free of overt inflammation. *J. Affect. Disord.* **182**, 106–114 (2015).

120. Heizmann, C. W. & Cox, J. A. New perspectives on S100 proteins: a multi-functional Ca(2+)-, Zn(2+)- and Cu(2+)-binding protein family. *Biometals* **11**, 383–97 (1998).

121. Simard, J. *et al.* Inhibitory effect of estrogens on GCDFP-15 mRNA levels and secretion in ZR-75-1 human breast cancer cells. *Mol. Endocrinol.* **3**, 694–702 (1989).

122. Hassan, M. I., Waheed, A., Yadav, S., Singh, T. P. & Ahmad, F. Prolactin inducible protein in cancer, fertility and immunoregulation: Structure, function and its clinical implications. *Cellular and Molecular Life Sciences* **66**, 447–459 (2009).

123. Umadat, V., Ihedioha, O., Shiu, R., Uzonna, J. & Myal, Y. The prolactin-inducible-protein (PIP): A regulatory molecule in adaptive and innate immunity. *Open J. Immunol.* **03**, 210–217 (2013).

124. Fitzgerald, P. & Dinan, T. G. Prolactin and dopamine: What is the connection? A Review Article. *Journal of Psychopharmacology* **22**, 12–19 (2008).

125. Charan, A., Shewade, D. G., Rajkumar, R. P. & Chandrasekaran, A. Relation between serum prolactin levels and antipsychotic response to risperidone in patients with schizophrenia. *Psychiatry Res.* **240**, 209–213 (2016).

126. Pérez-Iglesias, R. *et al.* Long-term effect of haloperidol, olanzapine, and risperidone on plasma prolactin levels in patients with first-episode psychosis. *J. Clin. Psychopharmacol.* **32**, 804–808 (2012).

127. Andley, U. P. Crystallins in the eye: Function and pathology. *Prog. Retin. Eye Res.* **26**, 78–98 (2007).

128. Renkawek, K., Voorter, C. E. M., Bosman, G. J. C. G. M., van Workum, F. P. A. & de Jong, W. W. Expression of αB-crystallin in Alzheimer's disease. *Acta Neuropathol.* **87**, 155–160 (1994).

129. Outeiro, T. F. *et al.* Small heat shock proteins protect against α-synuclein-induced toxicity and aggregation. *Biochem. Biophys. Res. Commun.* **351**, 631–638 (2006).

130. van Noort, J. M. *et al.* The small heat-shock protein αB-crystallin as candidate autoantigen in multiple sclerosis. *Nature* **375**, 798–801 (1995).

131. Garbett, K. A., Hsiao, E. Y., Kálmán, S., Patterson, P. H. & Mirnics, K. Effects of maternal immune activation on gene expression patterns in the fetal brain. *Transl. Psychiatry* **2**, (2012).

132. Boelens, W. C. Cell biological roles of αB-crystallin. *Prog. Biophys. Mol. Biol.* **115**, 3–10 (2014).

133. Bradway, S. D., Bergey, E. J., Jones, P. C. & Levine, M. J. Oral mucosal pellicle. Adsorption and transpeptidation of salivary components to buccal epithelial cells. *Biochem. J.* **261**, 887–896 (1989).

134. Cá, M., Elofsson, U. & Lindh, L. Salivary Mucin MUC5B Could Be an Important Component of in Vitro Pellicles of Human Saliva: An in Situ Ellipsometry and Atomic Force Microscopy Study. (2007). doi:10.1021/bm061055h

135. Tabak, L. A., Levine, M. J., Mandel, I. D. & Ellison, S. A. Role of salivary mucins in the protection of the oral cavity. *J. Oral Pathol.* **11**, 1–17 (1982).

136. Inoue, H. *et al.* Rheological Properties of Human Saliva and Salivary Mucins. *J. Oral Biosci.* **50**, 134–141 (2008).

137. Lieleg, O. & Ribbeck, K. Biological hydrogels as selective diffusion barriers. *Trends in Cell Biology* **21**, 543–551 (2011).

138. Gibbins, H. L., Proctor, G. B., Yakubov, G. E., Wilson, S. & Carpenter, G. H. SIgA binding to mucosal surfaces is mediated by mucin-mucin interactions. *PLoS One* **10**, e0119677 (2015).

139. Nayak, A. & Carpenter, G. H. A physiological model of tea-induced astringency. *Physiol. Behav.* **95**, 290–294 (2008).

140. Ployon, S. *et al.* The membrane-associated MUC1 improves adhesion of salivary MUC5B on buccal cells. Application to development of an in vitro cellular model of oral epithelium. *Arch. Oral Biol.* **61**, 149–155 (2016).

141. Bradway, S. D. *et al.* Formation of salivary-mucosal pellicle: the role of

transglutaminase. *Biochem. J.* **284**, 557–564 (1992).

142. Gibbins, H. L., Yakubov, G. E., Proctor, G. B., Wilson, S. & Carpenter, G. H. What interactions drive the salivary mucosal pellicle formation? *Colloids Surfaces B Biointerfaces* **120**, 184–192 (2014).

143. Wang, L. *et al.* MUC1 and MUC5B can form a protective mucin scaffold. in *2004 IADR/AADR/CADR General Session (Honolulu, Hawaii)* (2004).

144. Squier, C. A. The permeability of oral mucosa. *Crit. Rev. Oral Biol. Med.* **2**, 13–32 (1991).

145. Asikainen, P., Mikkonen, J. J., Ruotsalainen, T. J., Koistinen, A. P. & Kullaa, A. M. Microstructure of the superficial epithelial cells of the human oral mucosa. *Ultrastruct. Pathol.* **38**, 6–12 (2014).

146. Kullaa, A. M., Asikainen, P., Herrala, M., Ukkonen, H. & Mikkonen, J. J. W. Microstructure of oral epithelial cells as an underlying basis for salivary mucosal pellicle. *Ultrastruct. Pathol.* **38**, 382–6 (2014).

147. Boegh, M., Baldursdóttir, S. G., Müllertz, A. & Nielsen, H. M. Property profiling of biosimilar mucus in a novel mucus-containing in vitro model for assessment of intestinal drug absorption. *Eur. J. Pharm. Biopharm.* **87**, 227–35 (2014).

148. Teubl, B. J. *et al.* The oral cavity as a biological barrier system: design of an advanced buccal in vitro permeability model. *Eur. J. Pharm. Biopharm.* **84**, 386–93 (2013).

149. Jacobsen, J., van Deurs, B., Pedersen, M. & Rassing, M. R. TR146 cells grown on filters as a model for human buccal epithelium: I. Morphology, growth, barrier properties, and permeability. *Int. J. Pharm.* **125**, 165–184 (1995).

150. Ventura, T. M. da S. *et al.* The proteomic profile of the acquired enamel pellicle according to its location in the dental arches. *Arch. Oral Biol.* **79**, 20–29 (2017).

151. Lee, Y. H. *et al.* Proteomic evaluation of acquired enamel pellicle during in vivo formation. *PLoS One* **8**, e67919 (2013).

152. Siqueira, W. L., Zhang, W., Helmerhorst, E. J., Gygi, S. P. & Oppenheim, F. G. Identification of protein components in in vivo human acquired enamel pellicle using LC-ESI-MS/MS. *J. Proteome Res.* **6**, 2152–60 (2007).

153. Delius, J. *et al.* Label-free quantitative proteome analysis of the surface-bound salivary pellicle. *Colloids Surfaces B Biointerfaces* **152**, 68–76 (2017).

154. Linden, S. K., Sutton, P., Karlsson, N. G., Korolik, V. & McGuckin, M. A. Mucins in the mucosal barrier to infection. *Mucosal Immunology* **1**, 183–197 (2008).

155. Nielsen, P. A. *et al.* Identification of a major human high molecular weight salivary mucin (MG1) as tracheobronchial mucin MUC5B. *Glycobiology* **7**, 413–419 (1997).

156. Thornton, D. J. *et al.* Salivary mucin MG1 is comprised almost entirely of different glycosylated forms of the MUC5B gene product. *Glycobiology* **9**, 293–302 (1999).

157. Preciado, D. *et al.* MUC5B Is the predominant mucin glycoprotein in chronic otitis media fluid. *Pediatr. Res.* **68**, 231–6 (2010).

158. Desseyn, J. L., Aubert, J. P., Van Seuningen, I., Porchet, N. & Laine, A. Genomic organization of the 3' region of the human mucin gene MUC5B. *J. Biol. Chem.* **272**, 16873–16883 (1997).

159. Kesimer, M., Makhov, A. M., Griffith, J. D., Verdugo, P. & Sheehan, J. K. Unpacking a gel-forming mucin: A view of MUC5B organization after granular release. *American Journal of Physiology - Lung Cellular and Molecular Physiology* **298**, (2010).

160. Perez-Vilar, J., Eckhardt, A. E. & Hill, R. L. Porcine submaxillary mucin forms disulfide-bonded dimers between its carboxyl-terminal domains. *J. Biol. Chem.* **271**, 9845–9850 (1996).

161. Perez-Vilar, J., Eckhardt, A. E., DeLuca, A. & Hill, R. L. Porcine submaxillary mucin forms disulfide-linked multimers through its amino-terminal D-domains. *J. Biol. Chem.* **273**, 14442–14449 (1998).

162. Ridley, C. *et al.* Assembly of the respiratory Mucin MUC5B a new model for a gel-forming Mucin. *J. Biol. Chem.* **289**, 16409–16420 (2014).

163. Bobek, L. A., Tsai, H., Biesbrock, A. R. & Levine, M. J. Molecular cloning, sequence, and specificity of expression of the gene encoding the low molecular weight human salivary mucin (MUC7). *J. Biol. Chem.* **268**, 20563–9 (1993).

164. Gururaja, T. L. *et al.* Structural features of the human salivary mucin, MUC7. *Glycoconj.*

J. **15**, 457–67 (1998).

165. Bruno, L. S. *et al.* Two-hybrid analysis of human salivary mucin MUC7 interactions. *Biochim. Biophys. Acta - Mol. Cell Res.* **1746**, 65–72 (2005).

166. Iontcheva, I., Oppenheim, F. G. & Troxler, R. F. Human salivary mucin MG1 selectively forms heterotypic complexes with amylase, proline-rich proteins, statherin, and histatins. *J. Dent. Res.* **76**, 734–743 (1997).

167. Plummer, C. & Douglas, C. W. I. Relationship between the ability of oral streptococci to interact with platelet glycoprotein Ibα and with the salivary low-molecular-weight mucin, MG2. *FEMS Immunol. Med. Microbiol.* **48**, 390–399 (2006).

168. Stinson, M. W. *et al.* Adherence of Streptococcus sanguis to salivary mucin bound to glass. *J. Dent. Res.* **61**, 1390–3 (1982).

169. Kesimer, M., Kiliç, N., Mehrotra, R., Thornton, D. J. & Sheehan, J. K. Identification of salivary mucin MUC7 binding proteins from Streptococcus gordonii. *BMC Microbiol.* **9**, (2009).

170. Moshier, A., Reddy, M. S. & Scannapieco, F. A. Role of type 1 fimbriae in the adhesion of Escherichia coli to salivary mucin and secretory immunoglobulin A. *Curr. Microbiol.* **33**, 200–8 (1996).

171. Heo, S. M. *et al.* Host defense proteins derived from human saliva bind to Staphylococcus aureus. *Infect. Immun.* **81**, 1364–1373 (2013).

172. Veerman, E. C., Ligtenberg, A. J., Schenkels, L. C., Walgreen-Weterings, E. & Nieuw Amerongen, A. V. Binding of human high-molecular-weight salivary mucins (MG1) to Hemophilus parainfluenzae. *J. Dent. Res.* **74**, 351–7 (1995).

173. Namavar, F., Sparrius, M., Veerman, E. C., Appelmelk, B. J. & Vandenbroucke-Grauls, C. M. Neutrophil-activating protein mediates adhesion of Helicobacter pylori to sulfated carbohydrates on high-molecular-weight salivary mucin. *Infect. Immun.* **66**, 444–7 (1998).

174. Frenkel, E. S. & Ribbeck, K. Salivary mucins protect surfaces from colonization by cariogenic bacteria. *Appl. Environ. Microbiol.* **81**, 332–338 (2015).

175. Bingle, C. D. PLUNC: A novel family of candidate host defence proteins expressed in the upper airways and nasopharynx. *Hum. Mol. Genet.* **11**, 937–943 (2002).

176. Beamer, L. J., Carroll, S. F. & Eisenberg, D. Crystal structure of human BPI and two bound phospholipids at 2.4 Angstrom resolution. *Science (80-.).* **276**, 1857–1860 (1997).

177. Bingle, C. D., Seal, R. L. & Craven, C. J. Systematic nomenclature for the PLUNC/PSP/BSP30/SMGB proteins as a subfamily of the BPI fold-containing superfamily. in *Biochemical Society Transactions* **39**, 977–983 (2011).

178. Bingle, L. *et al.* Characterisation and expression of SPLUNC2, the human orthologue of rodent parotid secretory protein. *Histochem. Cell Biol.* **132**, 339–349 (2009).

179. Bingle, L. & Bingle, C. D. Distribution of human PLUNC/BPI fold-containing (BPIF) proteins. in *Biochemical Society Transactions* **39**, 1023–1027 (2011).

180. Prokopovic, V. *et al.* Isolation, biochemical characterization and anti-bacterial activity of BPIFA2 protein. *Arch. Oral Biol.* **59**, 302–309 (2014).

181. Shin, O. S. *et al.* LPLUNC1 modulates innate immune responses to vibrio cholerae. *J. Infect. Dis.* **204**, 1349–1357 (2011).

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

183. Blankenvoorde, M. F. *et al.* Cystatin and cystatin-derived peptides have antibacterial activity against the pathogen Porphyromonas gingivalis. *Biol. Chem.* **379**, 1371–5 (1998).

184. Dsamou, M. *et al.* Salivary Protein Profiles and Sensitivity to the Bitter Taste of Caffeine. *Chem. Senses* **37**, 87–95 (2012).

185. Phalipon, A. & Corthésy, B. Novel functions of the polymeric Ig receptor: Well beyond transport of immunoglobulins. *Trends in Immunology* **24**, 55–58 (2003).

186. Davis, T. L. *et al.* Structural and biochemical characterization of the human cyclophilin family of peptidyl-prolyl isomerases. *PLoS Biol.* **8**, (2010).

187. Wilkinson, B. & Gilbert, H. F. Protein disulfide isomerase. *Biochim. Biophys. Acta* -

Proteins Proteomics **1699**, 35–44 (2004).

188. Perretti, M. *et al.* Mobilizing lipocortin 1 in adherent human leukocytes downregulates their transmigration. *Nat. Med.* **2**, 1259–1262 (1996).

189. Flower, R. J. & Blackwell, G. J. Anti-inflammatory steroids induce biosynthesis of a phospholipase A2 inhibitor which prevents prostaglandin generation. *Nature* **278**, 456–9 (1979).

190. Leoni, G. *et al.* Human neutrophil formyl peptide receptor phosphorylation and the mucosal inflammatory response. *J. Leukoc. Biol.* **97**, 87–101 (2015).

191. Cooray, S. N. *et al.* Ligand-specific conformational change of the G-protein-coupled receptor ALX/FPR2 determines proresolving functional responses. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18232–18237 (2013).

192. Leoni, G. *et al.* Annexin A1-containing extracellular vesicles and polymeric nanoparticles promote epithelial wound repair. *J. Clin. Invest.* **125**, 1215–1227 (2015).

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

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

195. Babbin, B. A. *et al.* Annexin I regulates SKCO-15 cell invasion by signaling through formyl peptide receptors. *J. Biol. Chem.* **281**, 19588–99 (2006).