Review

The Post-Translational Modifications of Human Salivary Peptides and Proteins Evidenced by Top-Down Platforms

Irene Messana [1], Barbara Manconi [2], Tiziana Cabras [2,*], Mozhgan Boroumand [3], Maria Teresa Sanna [2], Federica Iavarone [4,5], Alessandra Olianas [2], Claudia Desiderio [1], Diana Valeria Rossetti [1], Federica Vincenzoni [4,5], Cristina Contini [2], Giulia Guadalupi [2], Antonella Fiorita [5,6], Gavino Faa [7,8] and Massimo Castagnola [9]

[1] Istituto di Scienze e Tecnologie Chimiche "Giulio Natta", Consiglio Nazionale delle Ricerche, 00168 Rome, Italy; imessana53@gmail.com (I.M.); claudia.desiderio@cnr.it (C.D.); dianaross79@hotmail.com (D.V.R.)
[2] Department of Life and Environmental Sciences, University of Cagliari, 09124 Cagliari, Italy; bmanconi@unica.it (B.M.); sanna@unica.it (M.T.S.); olianas@unica.it (A.O.); cristina.contini93@unica.it (C.C.); giulia.guadalupi@unica.it (G.G.)
[3] National Institut on Aging, NIH, Baltimore, MD 21224, USA; mozhgan.boroumand@nih.gov
[4] Dipartimento di Scienze Biotecnologiche di Base, Cliniche Intensivologiche e Perioperatorie, Università Cattolica del Sacro Cuore, 00168 Rome, Italy; federica.iavarone@unicatt.it (F.I.); federica.vincenzoni@unicatt.it (F.V.)
[5] Fondazione Policlinico Universitario A. Gemelli Fondazione IRCCS, 00168 Rome, Italy; antonella.fiorita@policlinicogemelli.it
[6] Dipartimento di Scienze dell'Invecchiamento, Neurologiche, Ortopediche e della Testa e del Collo, Università Cattolica del Sacro Cuore, 00168 Rome, Italy
[7] Unit of Pathology, Department of Medical Sciences and Public Health, University of Cagliari, 09124 Cagliari, Italy; gavinofaa@gmail.com
[8] Department of Biology, College of Science and Technology, Temple University, Philadelphia, PA 19122, USA
[9] Proteomics Laboratory, European Center for Brain Research, (IRCCS) Santa Lucia Foundation, 00168 Rome, Italy; maxcastagnola@outlook.it
* Correspondence: tcabras@unica.it

Abstract: In this review, we extensively describe the main post-translational modifications that give rise to the multiple proteoforms characterized to date in the human salivary proteome and their potential role. Most of the data reported were obtained by our group in over twenty-five years of research carried out on human saliva mainly by applying a top-down strategy. In the beginning, we describe the products generated by proteolytic cleavages, which can occur before and after secretion. In this section, the most relevant families of salivary proteins are also described. Next, we report the current information concerning the human salivary phospho-proteome and the limited news available on sulfo-proteomes. Three sections are dedicated to the description of glycation and enzymatic glycosylation. Citrullination and N- and C-terminal post-translational modifications (PTMs) and miscellaneous other modifications are described in the last two sections. Results highlighting the variation in the level of some proteoforms in local or systemic pathologies are also reviewed throughout the sections of the manuscript to underline the impact and relevance of this information for the development of new diagnostic biomarkers useful in clinical practice.

Keywords: salivary proteins; top-down proteomics; post-translational modifications

1. Introduction

Human whole saliva is a hypotonic fluid lining the oral cavity and is composed of water (99%) and a complex mixture of organic and inorganic compounds resulting from salivary gland secretion, oral flora, the oropharynx, the upper airway, gastrointestinal reflux, gingival crevicular fluid, food deposits, and mucosal surface secretion containing blood-derived components [1]. Saliva is essential for the accomplishment of multiple physiological

functions, encompassing lubrication, buffering, the maintenance of tooth integrity, chewing, the initial digestion of some foods, swallowing, tissue hydration and lubrication, speech, and wound healing, and it exhibits antibacterial and antifungal activity [2]. The dramatic sequelae observed in patients suffering from Sjögren syndrome clearly demonstrate the relevance of saliva and its components, particularly salivary proteins, in the protection of the mouth. Recent proteomic inventories report more than 3000 proteoforms in human saliva [3]. Before, during, and after secretion, most salivary proteins undergo numerous post-translational modifications (PTMs), of which the roles have not yet been clearly elucidated. Over the last twenty-five years, our group has been able to characterize many PTMs of the salivary proteome, mainly by applying top-down proteomic platforms to saliva from healthy subjects of various ages and from patients affected by various pathologies. In this review, we report the PTMs most frequently observed, suggesting, when possible, some hypotheses on the possible role played in the protection of the mouth. We have conducted exhaustive research on these topics in the literature. Nonetheless, we apologize in advance for any possible omission.

2. Enzymatic Cleavages (Cryptides)

The most-common PTM detectable in human saliva is the proteolytic cleavage of proteins. Indeed, most salivary proteins are submitted to the action of endogenous and exogenous proteinases (the last mainly from oral flora), which leads to the formation of a myriad of fragments. Many of these fragments could be considered important members of the cryptide family, defined as bioactive peptides encrypted inside a bigger functional polypeptide and released by a proteolytic event, with distinct or related function, but not superimposable, to that of the parental polypeptides [4]. The action of proteases can occur before and after secretion, and in the following paragraphs, this topic will be reviewed by describing pre-secretory and post-secretory cleavages separately.

2.1. Pre-Secretory Cleavages

The secretory pathway for many proteins includes transit in the Golgi apparatus and storage in secretory granules preceding their release from the cell into the duct system and secretion into the mouth [5]. Most of the pre-secretory cleavages of proteins occur during the transport towards the granules of the trans-Golgi-network [6], which represents the major secretory-pathway sorting station. Glandular secretions, protein extracts from secretory granules isolated from the major salivary glands, and whole saliva were investigated by top-down proteomics to characterize and distinguish between events occurring prior to the storage in the secretory granules, those taking place between granule release and secretion into the mouth during the passage through the secretory duct, and those occurring into the mouth [6]. The workflow applied in that study is represented in Figure 1. Proteases involved in the most relevant cleavages belong to the convertase and exopeptidase families [6–9].

2.1.1. Proline-Rich Proteins (PRPs)

Human PRPs constitute a very polymorphic family of proteins, missing amino acids in their sequence aromatic and characterized by a high content of proline (as well as glutamine and glycine). All PRPs are encoded by chromosome 12p13.2 and are divided into three classes based on their ionic properties, i.e., acidic proline-rich-proteins (aPRPs), encoded by PRH1 and PRH2 loci; basic proline-rich proteins (bPRPs), encoded by the PRB1, PRB2, and PRB4 loci located (considering the direction of translation) immediately before the aPRPs loci; and the glycosylated (basic) proline-rich proteins (gPRPs), encoded by the PRB3 locus, which is the fourth locus of the bPRPs cluster. It is relevant to underline that except for the bPRPs named P-D, the other three proteoforms of the PRB-4 locus, i.e., the IB-8a (Con 1+) of PRB-2 and all the products of the PRB-3 locus, are submitted to glycosylation, which will be described in more detail, if available, in the section devoted to this PTM.
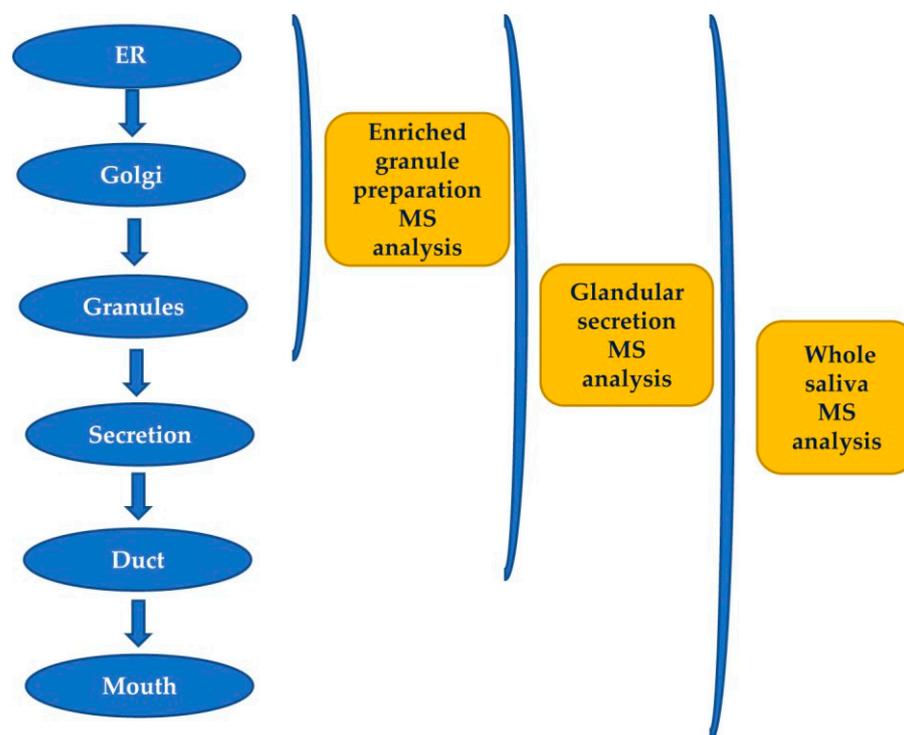
Figure 1. Workflow applied for studying the origin of post-translational modifications (PTMs) of salivary peptides and proteins secreted by the major salivary glands by a top-down proteomic approach.

Acidic Proline-Rich-Proteins (aPRPs)

aPRPs exist as five principal isoforms called PRP-1, PRP-2 (coded by *PRH-2* locus) and Pif-s (parotid isoelectric-focusing isoform-slow), Pa (parotid acidic isoform), and Db-s (double band isoform-slow) (coded by *PRH-1* locus). The isoforms PRP-1, PRP-2, Pif-s, and Pa are all 150 residues long. Taking as reference PRP-2, the Pif-s and Pa isoforms have an asparagine (Asn_4) instead of Asp_4, and PRP-1 has Asn_{50} instead of the Asp_{50} present in the other isoforms (Figure 2). The Db-s isoform is 171 a.a. residues long due to the insertion of 21 residues repeated after position 81 and, as Pa, it has a Leu_{27} instead of $Ileu_{27}$. The PRP-1, PRP-2, and PIF-s isoforms are partially cleaved at the Arg_{106} by a convertase recognizing the . . .$R_{103}XXR_{106}$. . . consensus sequence, generating the PRP-3, PRP-4, and Pif-f (f stands for fast electrophoretic isoform) truncated isoforms and a terminal fragment of 44 amino-acid residues called the P-C (or IB-8b) peptide. The convertase consensus sequence of Db-s is shifted by the insertion, and it is cleaved at Arg_{127}, generating the Db-f isoform and always the P-C peptide. The consensus sequence of the convertase responsible for the cleavage is preserved because the Pa isoform, having a Cys_{103} residue instead of Arg_{103}, is not cleaved. Indeed, a truncated Pa isoform was never detected. Rather, it generates in the mouth a Pa-2-mer isoform due to the formation of a disulfide bridge between two Cys_{103} residues (Figure 2), which is the main Pa component detectable in whole saliva [9]. The cleavage of PRP-1, PRP-2, Pif-s, and Db-s is not complete, and in whole saliva, the entire and truncated isoforms are both detectable, approximatively in a proportion of 70/30%, respectively. All the entire isoforms are the PRP-1 type, while all the truncated are the PRP-3 type (obviously non-including the Pa proteoform). All these characteristics of aPRPs are summarized in Figure 2.
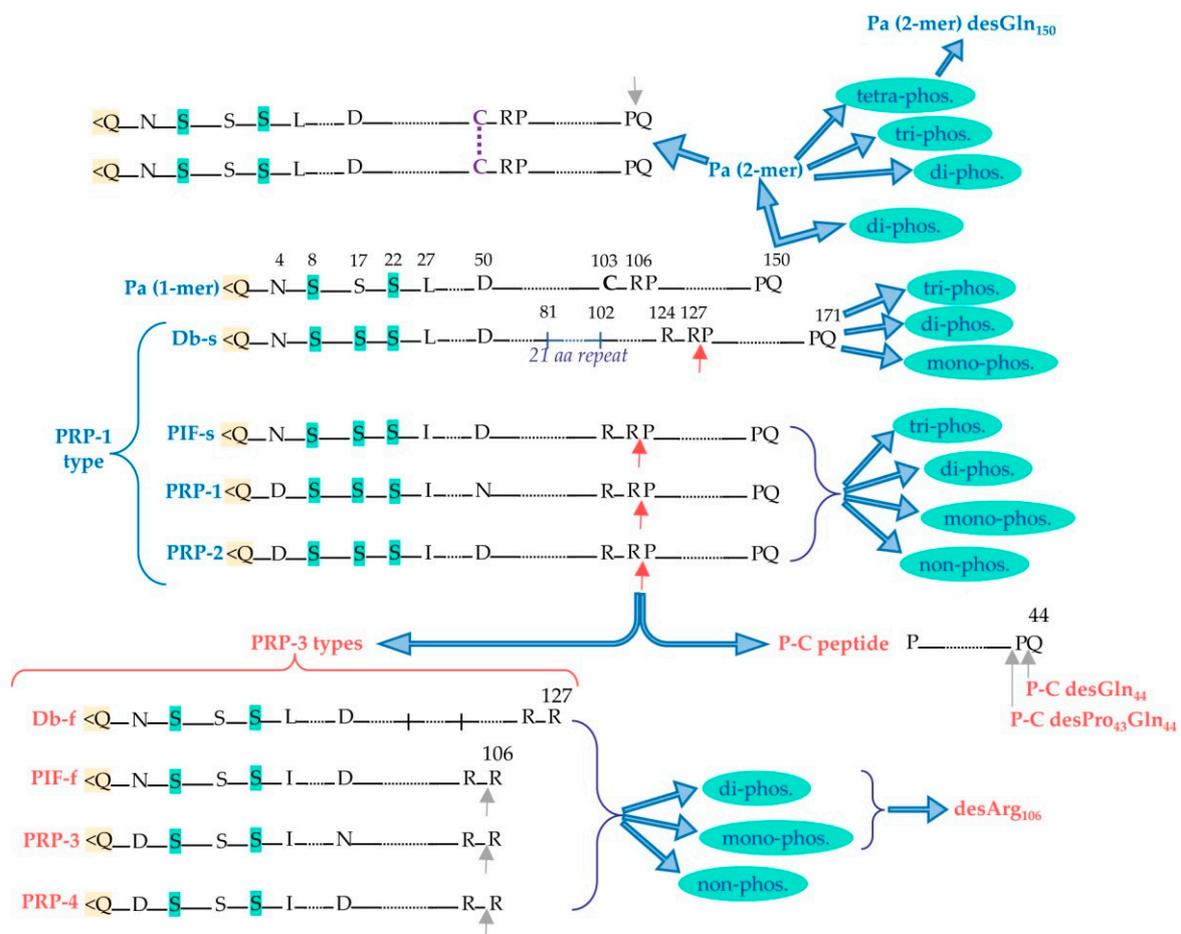
Figure 2. Principal acidic proline-rich protein (aPRP) proteoforms: in blue are the entire proteoforms and in red are the truncated ones. Red arrows indicate the cleavage site recognized by convertase activity, from which PRP-3-type truncated forms and the P-C peptide are generated from the entire PRP-1-type forms. Grey arrows indicate the possible loss of C-terminal residue by carboxypeptidase activity. Pyroglutamination N-terminal is evidenced in yellow, while the Ser residues that can undergo phosphorylation are evidenced in light green. The different phosphorylated proteoforms obtainable from the entire and truncated aPRPs are shown, as well as the disulfide dimeric form of Pa.

Basic-Proline-Rich-Proteins (bPRPs)

bPRPs are encoded by the polymorphic *PRB1*, *PRB2*, and *PRB4* loci, existing as three or four possible alleles: (S(mall), M(edium), and L(arge), or for the eventual fourth allele, V(ery)L(arge)). Unlike aPRPs, they are completely cleaved by convertases during granule storage [6]. The proteoforms generated from the pre-pro-proteins are listed below:

- Products of Locus *PRB-1*: II-2, P-E (or IB-9), P-Ko, IB-6, Ps-1, Ps-2.
- Products of Locus *PRB-2*: IB-1, P-J, P-H (or IB-4), P-F (orIB-8c), IB-8a(Con 1+), IB-8a(Con 1−).
- Products of Locus *PRB-4*: PGA, II-1, Cd-IIg, P-D (or IB-5).

The nomenclature reported for bPRPs is difficult and confusing since it derives from two different criteria. Kauffmann and colleagues [10–13] have purified eleven bPRPs and determined the sequence of ten of them. They named these proteoforms using the names of the different fractions obtained from an articulate multidimensional preparative chromatographic separation of whole saliva. In the same years, Saitoh and colleagues [14–16] and Isemura and colleagues [17,18] identified nine bPRPs, naming them from P-A to P-I, and sequenced seven of them. As reported in the list, the two nomenclatures have several overlaps. P-A and P-I are artifacts deriving from the proteolytic cleavage occurring during the purification of salivary

proteins. Although included in the class of basic PRPs, P-B and P-C peptides are not codified by *PRB1-PRB4* genes. Indeed, the P-B peptide is encoded by the *PROL3* gene (PBI) clustered on chromosome 4q13.317. It is related to statherin and, therefore, it cannot be considered to pertain to the bPRP family. As described in the previous section on aPRPs, the P-C peptide derives from the cleavage of four isoforms of aPRPs, namely PRP-1, PRP-2, Pif-s, and Db-s. In a top-down MS proteomic study, we were able to characterize a new bPRP deriving from the *PRB-2* locus, which was named P-J [19,20].

The pre-secretory cleavage of bPRP pro-proteins usually occurs at the level of the consensus sequence KSRSXR↓, where X may be Pro, Ser, or Ala [6,21,22]. An extensive characterization of the different proteoforms and fragments of bPRPs has been recently published [23], where the interested reader can find the list of the bPRPs components detected until 2018. In this inventory, a new classification of bPRPs was proposed based on the similarities of their sequence, dividing them into three groups, reported in Figure 3a–c. The first group, which we named Group 1, includes P-E, P-Ko, IB-6, Ps-1, Ps-2, P-H, P-F, P-J, and P-D (Figure 3a). The sequence of all these bPRPs starts with the same SPPGKPQGPP motif, followed by sequences somewhat similar but showing small variations among the different components. The central part of the sequences shows similar repeats. Because P-E, IB-6, Ps-1, and Ps-2 sequences originate from DNA-length polymorphisms in exon 3 of the *PRB1* locus, they exhibit high similarity. The bPRP with a Mav of 10,433.5 Da, detected in whole saliva and in parotid secretory granules and named P-Ko, is encoded by *cP4*, a differentially spliced transcript of the *PRB1-L* allele.
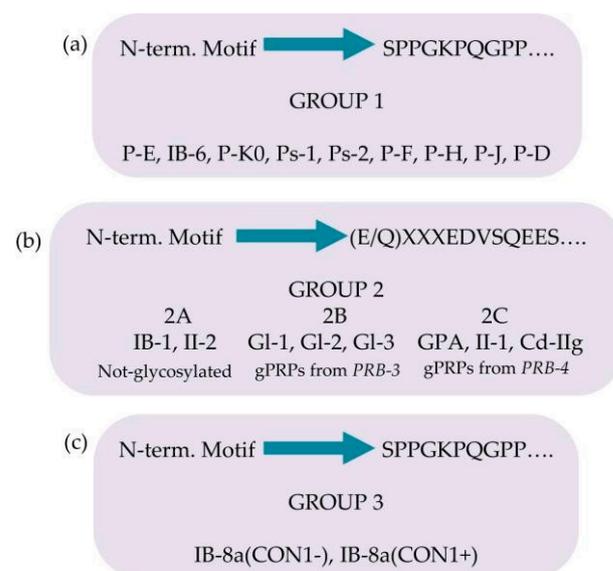
Figure 3. Classification of the basic proline-rich proteins (bPRPs) based on the similarity of the N-terminal motif and on the presence of glycosylations. Group 1 in panel (**a**), group 2 in panel (**b**), group 3 in panel (**c**).

Group 2 includes IB-1, II-2, and the glycosylated bPRPs codified by *PRB3* and *PRB4* genes, namely, Gl-1, Gl-2, Gl-3, GPA, II-1, and Cd-IIg (Figure 3b). Their sequences start with a similar motif (E/Q)XXXEDVSQEES, where XXX is LNE in IB-1, II-2, Gl-1, Gl-2, and Gl-3 and SSS in GPA, II-1, and Cd-IIg). The central part of the sequences comprises similar repeats, with differences from the repeats of the members belonging to Group 1. Based on structural differences and similarities, members of Group 2 can be divided into three subgroups: Group 2A, including IB-1 and II-2, without glycosylation sequons; Group 2B, including the Gl proteins codified by the alleles of *PRB3* locus; and Group 2C, including the glycosylated proteins codified by the alleles of *PRB4* locus. The Small Group (3) is a hybrid group, which includes the two proteoforms of IB-8a, Con1−, and Con1+ (Figure 3c). The initial sequence of these two proteins resembles that of Group 1, while the terminal

sequence is like the repeat responsible for the glycosylation of the bPRPs of Groups 2B and 2C [23].

2.1.2. Further Pre-Secretory Cleavages of PRPs

All the PRP families described above are substrates of a carboxypeptidase, which removes the C-terminal residue and commonly, obviously, is an arginine. However, the enzyme is not specific. It removes other residues, and sometimes a second C-terminal loss is detectable [22]. The C-terminal pre-secretory cleavages occurring on PRPs are reported in Table 1.

Table 1. Some terminal losses detected in salivary peptides/proteins.

Protein and Peptides	C-Term. Cleavage
aPRP Pa-dimer	…QSP↓Q
PRP3/PRP4/PIF-f	…RPP↓R
P-C (or IB-8b)	…QS↓P↓Q
bPRP II-2	…RSP↓R
bPRP IB-1	…RSP↓R
bPRP PF (or IB-8c) precursor	…RSA↓R
bPRP PE (or IB-9)	…RSP↓R

2.1.3. Role of PRPs

Salivary proline-rich proteins are highly conserved in mammalian saliva [24]. Nonetheless, significant structural differences are present in the mammal families, suggesting that they play a relevant role in oral protection, in the modulation of the activity of oral ions, in the colonization of oral microbiota and the gastrointestinal tract, and in the feeding habit [24]. However, while aPRPs and the gPRP, called Gl 3M, are detectable in the saliva of preterm newborns, the other bPRPs are not detectable in human saliva until puberty [20,25,26], suggesting either a role in the perception of the taste of foods or a function in secondary sexual maturation. As we are aware, this information is not available for other mammals. In humans, bPRPs are secreted only by parotid glands, and this regioselectivity is also puzzling. Some bPRPs exhibit the ability to bind harmful tannins [27], others can modulate the oral flora [28,29], and some others are involved in bitter taste perception [30]. Some bPRP fragments are involved in enamel pellicle formation [31] and others act as antagonists of progesterone-induced cytosolic Ca^{2+} mobilization [32].

The intrinsic propensity of some fragments to adopt a polyproline-II helix arrangement joined to PXXP motifs was suggestive of the interaction with the SH3 domain family [33]. Interestingly, interactions were highlighted with Fyn, Hck, and c-Src SH3 domains [34], which are included in the Src kinases family, suggesting that some basic bPRPs can be involved in the signal transduction pathways modulated by these kinases. Only a small amount of data on correlations between genes of bPRPs and diseases linked to their allelic variants has been reported so far. In fact, for some of the alleles (PRB1VL, PRB2S, M, VL, and PRB3VL) the genetic sequence is not reported. Moreover, for the small and large alleles of PRB1, the genetic sequence is incomplete [35] because the reference genome (NCBI Gene ID: 5542) encodes the medium allele. Regarding the primary structure of bPRP alleles, in the UniProtKB database (accession number P04280), the full amino-acid sequence of the large variant, deduced through experimental evidence at the protein level, is deposited.

2.1.4. Statherin and P-B Peptides

Statherin is encoded by the *STATH* gene localized on chromosome 4q13.3 [36]. It is a 43-amino-acid-residues multifunctional phospho-peptide characterized by an anomalous high content of tyrosine, proline and glutamine. The roles of statherin in the oral cavity are various. Due to its great affinity for calcium phosfate minerals, such as hydroxyapatite, it

maintains the supersaturated state of calcium in human saliva, thus contributing to the mineral dynamics of the tooth surface and stabilizing the acquired enamel [37]. Furthermore, statherin is involved in bacterial colonization [38] and acts as a boundary lubricant on the enamel surface [39]. P-B is the product of the specific *PROL3* gene located on chromosome 4q13.3, very close to the *STATH* gene [40]. Differently from bPRPs, which do not have aromatic amino acids in their sequences, and similarly to statherin, it has three tyrosine residues in its sequence. Furthermore, like statherin, the P-B peptide is secreted mainly by submandibular/sublingual (SM/SL) glands and also by parotids [41]. It is important to recall that P-B is also called, at the Swiss-Prot site (code P02814), "submaxillary gland androgen-regulated protein 3B precursor", but we were not able to find any reference justifying the origin of this name. The specific role of the P-B peptide in saliva is still far from being clarified [42,43]. Nonetheless, its sequence resembles the ionic complement of statherin. Indeed, statherin has a small N-terminal negative tail followed by a long neutral domain with many proline residues, while the P-B peptide has a small N-terminal positive tail followed by a long neutral poly-proline domain. The two peptides could interact for the formation of the acquired enamel pellicle of the tooth. Different isoforms of statherin have been detected in human saliva; the variant called SV1 (Statherin variant 1) corresponds to statherin missing the C-terminal Phe residue (Statherin desPhe$_{43}$) [41,44] (Figure 4a). The variant SV2, showing the deletion of the 6–15 internal residues with respect to statherin (Statherin des6-15), originates from alternative splicing that excludes an exon of 30 nucleotides [45]. SV3 corresponds to the SV2 variant lacking the C-terminal Phe residue. SV-1 lacking the C terminal Phe$_{43}$ and statherin desThr$_{42}$Phe$_{43}$ (Figure 4a) was detected in parotid and submandibular secretory granules, suggesting that C-terminal removal occurs during the maturation of the secretory granules [6]. On the contrary, the N-terminal removal that generates statherin desAsp$_1$ (Figure 4a) is probably an event occurring after granule secretion because this derivative was not detected in granules [6]. Statherin desAsp$_1$ was also detected in human saliva by Vitorino and colleagues [46].

Figure 4. Pattern of proteoforms, mainly proteolytic fragments, originating from statherin (**a**), and the P-B peptide (**b**), characterized in human saliva by a top-down approach based on reversed-phase high pressure liquid chromatography-electrospray ionization tandem mass (RP-HPLC-ESI-MS/MS) analysis. The arrows indicate the main cleavage sites. The sequences recognized by furin-like convertase are evidenced in bold red, and the fragments derived by these cleavages are in shown in red.

Similarly, a small amount of P-B des1–5 was found in whole saliva as well as in parotid and submandibular-sublingual saliva but not in secretory granules, suggesting that this N-terminal cleavage also occurs after secretion [6]. Some P-B and statherin naturally occurring fragments are generated from a cleavage operated by a furin-like convertase and have been characterized by a top-down approach in our previous studies [6,41]. Statherin undergoes convertase cleavages at the consensus sequence . . .$R9R_{10}IGR_{13}$. . ., generating statherin des1-9, des1-10, and des1-13 (Figure 4a), while P-B des1-5 is a convertase fragment of P-B peptide generated at the consensus sequence . . .R_2GPR_5. . . (Figure 4b). The P-B peptide undergoes further cleavages at the level of various chymotryptic-like sequences, generating several fragments detectable by a top-down proteomic approach [20] (Figure 4b).

2.1.5. Histatins (Hst)

Histatins are a family of small peptides deriving their name from the high number of histidine residues in their sequence. It is widely accepted that all the members of this family arise from two parent peptides, named histatin 1 and histatin 3, with very similar sequences and encoded by two genes (HIS1 and HIS2) located on chromosome 4q13 [47]. Despite the very high sequence similarity, these two peptides follow completely different PTM pathways. Histatin 3 (Hst3), differently from histatin 1 (Hst1), is submitted to a sequential cleavage, generating Hst6 (Hst3 1/25) at first, then Hst5 (Hst 3 1/24) subsequently, with powerful anti-fungal activity [48], and then other fragments [49]. Their different susceptibility to cleavage derives from the presence in Hst3 of the . . .RGYR↓. . . convertase consensus sequence, which is absent in Hst1 and thus is not cleaved. Hst1 today is considered a proangiogenic factor that may contribute to oral wound healing [50,51].

Some years ago, the group of Oppenheim [48] was able to establish the sequence of 12 histatins and named them from Hst1 to Hst12. The advent of the high-throughput MS apparatus applied to proteomic studies allowed for the determination of many other Hst fragments, the majority deriving from Hst3 [49,52]. For this reason, the nomenclature of histatins fragments has been recently modified and is reported on the Swiss Prot site with the code numbers P15515 (Hst1) and P15516 (Hst3).

2.1.6. Cystatins

Cystatins include type-1 cystatins (cystatins A and B), type-2 cystatins (C, D, S, SN, SA), and kininogens, or type-3 cystatins. Various biological activities have been demonstrated for these proteins, but a major role is linked to the inhibitory action exerted against cysteine proteinases. Thus, for this ability to modulate the proteolytic system, they are considered central in various diseases, including cancer [53–56]. Cystatins S, SN, SA, C, and D are encoded by loci *CST1–5* closely clustered on chromosome 20p11.21. It has been observed that S-type cystatins (S, SN, and SA), present at higher concentrations in SM/SL secretion than in parotid saliva [6], were absent or present at very low concentrations from both parotid and SM/SL secretory granules. This finding suggested that the secretion of S-type cystatins is not granule-mediated [6]. It is relevant to remark that cystatin A and B are leaderless [57]. Cystatin C was found sporadically in parotids, SM/SL, and whole saliva, but it was absent from secretory granules, whereas cystatin D was not found in any sample. Thus, it was suggested that cystatins C and D may have a different origin than the other salivary cystatins. The presence in human saliva of truncated proteoforms of cystatins has been suggested by Lupi and colleagues, who observed by HPLC-ESI-MS several masses possibly related to N-terminally truncated cystatins in the acidic soluble fraction of human saliva [58]. Indeed, a subsequent in-depth study performed both in the acidic supernatant of whole saliva and in RP-HPLC-enriched fractions, by an integrated top-down/bottom-up pipeline, characterized some truncated proteoforms of cystatins [57]. The study evidenced that not all cystatins undergo proteolytic modifications. The following truncated proteoforms for the widespread cystatin SN and its natural variant SN Pro_{11} → Leu were detected and characterized: cystatin SN des1–4, and SN des1−7, cystatin SN Pro_{11} → Leu des1–4, and SN Pro_{11} → Leu des1–7. A truncated form of cystatin SA

lacking the first seven amino acids from the N-terminus (cystatin SA des1–7) was also detected. Three truncated forms of the variant cystatin D $Cys_{26} \rightarrow Arg$ have been detected in human saliva, but none for the variant $Arg_{26} \rightarrow Cys$. For easier reading, we will utilize the name cystatin D to refer to the variant $Cys_{26} \rightarrow Arg$. Cystatin D des1–4 and des1–8 have been characterized by a top-down Fourier-transform ion cyclotron resonance mass-spectrometry pipeline [59] and also detected in the study of Manconi and colleagues [57] A third truncated proteoform of cystatin D lacking the first five amino-acid residues with the N-terminal glutamine converted to pyroglutamic acid (pGlu-cystatin D des1–5) was also characterized in the latter study. The authors speculated that the high abundance of the latter truncated proteoform was due to the greater resistance from degradation by amino peptidases due to N-terminal pyroglutamination. Cystatin B in adult human saliva is commonly detectable as an intact proteoform, and the two fragments 1-53 and 54-98 were characterized in the saliva of human preterm newborns [60].

2.2. Post-Secretory Cleavages

Almost all the families of salivary proteins and peptides described above (and others) are substrates of several proteolytic enzymes present in the mouth deriving from oral flora [61,62]. Since they are myriad, reporting a complete list is impossible, and in the following, we will refer to excellent published lists. Concerning PRPs and particularly bPRPs, many recent mass-spectrometric studies allowed us to report many bPRPs fragments [23,43,46,52], with high affinity for the tooth enamel [52]. However, the observation of a recurrent XPQ C-terminal sequence in many fragments detected induced the group of Oppenheim and Helmerhorst to characterize a glutamine endopeptidase from Rhotia sp [63]. Using the synthetic substrates KPQ-pNA and GGQ-pNA, the overall K(m) values were determined to be $97 +/- 7.7$ and $611 +/- 28$ micromolar, respectively, confirming glutamine endoprotease activity in whole saliva and the influence of the amino acids in positions P(2) and P(3) on protease recognition. The pH optimum of KPQ-pNA hydrolysis was 7.0, and the activity was most effectively inhibited by antipain and 4-(2-aminoethyl) benzene sulfonyl fluoride. The enzyme is metal-ion-dependent and not inhibited by cysteine protease inhibitors. A systematic evaluation of enzyme activities in various exocrine and non-exocrine contributors to whole saliva revealed that the glutamine endoprotease derives from Rhotia and is localized in dental plaque [63,64].

A further protein submitted to fragmentation in the mouth is the polymeric immunoglobulin receptor (PIgR), a type-I transmembrane glycoprotein playing the main role in the adaptive immune response on mucosal surfaces [65,66]. It transports polymeric IgA across mucosal epithelial cells. A proteolytic cleavage occurring in the glycosylated extra-cellular portion of pIgR generates the secretory component (19–603 residues), which has also been detected in human saliva [67]. The cleavage occurs by the action of unknown proteases, probably released by activated neutrophils [66], and the highly conserved sequence 602–613 (PRLFAEEKAVAD) is believed to be the cleavage signal [65]. Two peptides deriving from PIgR are detectable in saliva by a top-down proteomic approach; they are named AVAD and ASVD [68]. The peptide named AVAD originates by the cleavage occurring in this region at the level of Lys_{609}, and the ASVD peptide derives from AVAD by the trypsin-like cleavage at Arg_{622}. AVAD and ASVD peptides do not derive from the secretory component and have a sequence partially overlapped by the transmembrane portion (639–661) of PIgR. Thus, they should originate by degradation of PIgR after its release from disrupted cell membranes.

2.3. Proteolytic Cleavages: Variations Related to Age and Pathologies

Interestingly, it was observed that several proteolytic cleavages change according to age and in several diseases. For instance, as previously reported in Section 2.1.6, Iavarone and colleagues [60] detected in the saliva of human preterm newborns sensible amounts of cystatin B fragments 1–53 and 54–98, suggesting in foetuses the presence of high-active specific proteolytic events that disappeared in adults. The cleavage involves a tyrosine

and phenylalanine couple (Tyr↓Phe) and can be ascribed to a chymotrypsin-like enzymatic activity with strict specificity, as suggested by the very precise consensus sequence necessary for the cleavage. The two fragments were detected at very high relative amounts with respect to intact cystatin B in very preterm newborns, for which the concentration decreased as a function of the post-conceptional age (PCA) and they were practically undetectable when the age corresponded to that of full-term newborns. A search on the Merops protease database (release 12.4, accessed on 9 January 2017, http://merops.sanger.ac.uk/) returned various possibilities other than chymotrypsin A, such as chymosin, cathepsin E, metalloproteinase 2, ADAMTS4, endothelin converting enzyme 1, and some peptidases of the chymase class (mast cell chymotrypsin-like proteinase). It is impossible to establish if cystatin B is a natural substrate of the enzyme, thereby implying a functional role for the fragments as potential cryptides, or if the fragments observed are byproducts of a proteinase, without any functional meaning, their activity is devoted to other specific foetal oral cleavage processes. Similarly, an increased activity of convertases and carboxypeptidases responsible for the cleavages of aPRPs, histatins, and statherin was observed in preterm newborns with a low PCA of approximately 190 days. The activity decreased according to the PCA, and it reached the levels observed in the adult around the normal term of delivery [69]. Interestingly, this behavior was observed for the release of the PRP3-type aPRPs and the P-C peptide from PRP-1-type aPRPs, as well as for the release of histatin 6 and histatin 5 from histatin 3 and for the release of the fragments of statherin missing C-terminal residues. This was not the case for statherin desAsp$_1$, which showed low levels or was not detectable in preterm newborns at low PCA but increased after birth, reaching values similar to those determined in at-term newborns [70]. Because statherin desAsp$_1$ is a fragment of statherin not detectable in granule preparations [6], this result is a further clue that the proteinases (convertases and carboxypeptidases) involved in the cleavages of these families of salivary proteins are confined in the Golgi apparatus.

It has been reported that proteolytic fragmentations of salivary proteins may vary in some pathological states compared to physiological conditions. For instance, a study performed on a group of children and adolescents affected by type-1 diabetes revealed that several small peptides, most likely originated by post-secretory proteolytic cleavages occurring in the mouth, showed a higher concentration with respect to sex- and age-matched controls, suggesting increased activity of exogenous proteinases in the oral cavity of diabetics [70]. The peptides, characterized as fragments 1–14, 1–25, 5–25, 26–35, 26–44, and 36–44 of the P-C peptide, have been already detected in saliva from healthy subjects [63], and fragments 1–14 and 26–44 were also detected in the parotid saliva of healthy subjects by Hardt and colleagues [71]. Furthermore, an association of fragments 1–14 and 26–44 to high numbers of dental caries and the presence of fragments 1–14 in the saliva of subjects affected by Sjogren's syndrome have been demonstrated by Huq and colleagues [72]. A top-down proteomic study investigated the salivary proteome of 49 multiple sclerosis (MuSc) patients and 54 healthy controls, quantifying 119 salivary peptides/proteins [73]. Among the observed differences, the fragments 1–14, 26–44, and 36–44 of the P-C peptide, the SV1 fragment of statherin, and cystatin SN des1–4 showed higher levels in patients with respect to controls. In a study performed to characterize possible differences in the salivary proteome of subjects affected by Wilson's disease (WD) with respect to healthy controls, increased levels of the AVAD and ASVD fragments of PIgR (see above) were observed in patients, probably because of an increased disruption of cell membranes due to the high production of ROS typical of WD [68]. A higher level of the ASVD peptide was also determined in MuSc patients with respect to healthy controls [73]. Interestingly, a study performed on patients affected by systemic mastocytosis (SM) and grouped into SM with (SM+C) or without (SM-C) additional cutaneous lesions evidenced that the two SM forms were distinguished by the lower levels of PRP-3, PRP-3 desArg$_{106}$, statherin desPhe$_{43}$, P-B des1–5, and cystatin D des1–5 and des1–8 in SM-C patients with respect to SM+C [74]. It should be outlined that lower levels of cystatin D des1–5 and des1–8 have been also observed by comparing saliva of healthy elderly subjects with respect to adults [75].

The study was performed on patients affected by autoimmune hepatitis (AIH) and primary biliary cholangitis (PBC), two autoimmune liver diseases characterized by chronic hepatic inflammation and progressive liver fibrosis, to establish a panel of salivary proteins/peptides able to classify with good accuracy PBC patients vs HCs, AIH patients vs HCs, and PBC vs AIH patients. Among the other data, they revealed significantly different levels of PRP-3 desArg$_{106}$ by comparing the two patient groups [76].

Higher levels of statherin desPhe$_{43}$ and two others naturally occurring fragments (des1–9, and des1–13) were measured in the saliva of a group of Alzheimer's disease patients compared to controls [77].

3. Phosphorylation

Phosphorylation is probably (after proteolytic cleavage) the most-common PTM in human saliva. Generally, phosphorylation is a reversible PTM, catalyzed by more than 500 different human protein kinases, while de-phosphorylation is due to enzymes called phosfatases [78]. The detection of phosphorylated proteoforms in human saliva is strongly dependent on the proteomic pipeline utilized: bottom-up or top-down. Commonly, the bottom-up strategies are accomplished by enrichment capture in order to increase the number of phosphorylated fragments.

A study of the group of Oppenheim and Helmerhorst [79] utilized chemical derivatization using dithiothreitol (DTT) of the phospho-serine/threonine-containing peptides obtained after the trypsin digestion of whole-saliva samples. The DTT-phospho-peptides were enriched by covalent disulfide-thiol interchange chromatography and analysis by nanoflow liquid chromatography and electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS). The specificity of DTT chemical derivatization was evaluated separately under different base-catalyzed conditions with NaOH and Ba(OH)$_2$, blocking cysteine residues by iodoacetamide and enzymatic O-deglycosylation prior to the DTT reaction. Further analysis of whole-saliva samples that were subjected to either of these conditions provided supporting evidence for phosphoprotein identifications. The combined chemical strategies and mass-spectrometric analyses identified 65 phosphoproteins in whole saliva; of these, 28 were based on 2 or more peptide identification criteria with high confidence and 37 were based on a single phospho-peptide identification. Most of the identified proteins (∼80%) were previously unknown phosphoprotein components.

A phospho-proteomic study based on a bottom-up pipeline on human saliva also generated a large-scale catalog of phosphorylated protein fragments [80]. To circumvent the wide dynamic range of phosphoprotein abundance in whole saliva, the proteomic platform combined dynamic range compression using hexapeptide beads, strong cation exchange HPLC peptide fractionation, and immobilized metal-affinity chromatography prior to mass spectrometry. In total, 217 unique phospho-peptides sites were identified representing 85 distinct phosphoproteins at 2.3% global FDR. From these peptides, 129 distinct phosphorylation sites were identified, of which only 57 were previously known. Cellular localization analysis revealed salivary phosphoproteins had a distribution like all known salivary proteins but with less relative representation in "extracellular" and "plasma membrane" categories compared to salivary glycoproteins. Sequence alignment showed that phosphorylation is mainly linked to the action of the Golgi casein kinase called Fam20c (see below), but it also occurred at acidic-directed kinase, proline-directed, and basophilic motifs [80].

The top-down pipelines are more conservative, giving precise information on the salivary protein substrate of kinases and on their sites of modification. All the phosphorylated proteins pertaining to the secretory pathway are phosphorylated by a pleiotropic Golgi casein kinase which, until a few years ago, was elusive. An interesting study performed by Tagliabracci and colleagues [81,82] was able to establish that the enzyme is the kinase named Fam20C. The main consensus sequence of Fam20C is a serine with a +2 specific negative residue, either glutamic acid or phospho-serine (SXE/S(phos)). The salivary

proteins and peptides submitted to the action of Fam20C in agreement with this consensus sequence are reported in Table 2.

Table 2. Phosphorylation sites recognized by Fam20C of human-specific secretory proteins.

Protein/Peptide	Partial Sequence	Consensus	Refs.
Statherin (Ser$_2$, Ser$_3$)	DSSEEKF...	SXE	[41]
Histatin 1 (Ser$_2$)	DSHEKR...	SXE	[49]
Cystatin S (Ser$_3$)	SSSKEENR...	SXE	[57–59,83]
Cystatin S1 (Ser$_1$)	SSS$_{(phos)}$KEENR...	SXS$_{(Phos)}$	"
aPRP (Types 1 and 3) (Ser$_8$)	...DEDVSQEDV....	SXE	[9]
aPRP (Types 1 and 3) (Ser$_{22}$)	...GGDSEQFIDEER...	SX$_{(3/4)}$(E/D/S$_{(phos)}$)$_3$	"
aPRP (Types 1 and 3) (Ser$_{17}$)	...LVISDGG DS$_{(phos)}$EQFI...	SX$_{(3/4)}$(E/D/S$_{(phos)}$)$_3$	"
bPRP II-2 (Ser$_8$)	...NEDVSQEESPS...	SXE	[23]
bPRP IB-1 (Ser$_8$)	...NEDVSQEESPS...	SXE	"
gPRP Gl-2 or PRP-3M (Ser$_8$)	...NEDVSQEESPS...	SXE	"
gPRP Gl-1 or PRP-3L (Ser$_8$) [1]	...NEDVSQEESPS...	SXE	"
gPRP Gl-3 or PRP-3S (Ser$_8$) [1]	...NEDVSQEESPS...	SXE	"
gPRP Glycosyl. Protein A (Ser$_8$) [1]	...SEDVSQEESLFL...	SXE	"
gPRP II-1 (Ser$_8$) [1]	...SEDVSQEESLFL...	SXE	"
gPRP Cd-IIg (Ser$_8$) [1]	...SEDVSQEESLFL...	SXE	"

[1] Hypothetical (by similarity).

The phosphorylation of cystatin S, whose N-terminal sequence is $S_1SS_3KE_5$..., is interesting. Monophosphorylated cystatin S is called cystatin S1, and due to the presence of the glutamic acid residue in position +2, Ser$_3$ is necessarily the first site of phosphorylation. Di-phosphorylated cystatin S is called cystatin S2, and the second phosphorylation at Ser1 is strictly hierarchical, occurring only after that of Ser3. Cystatin S1 is the most abundant in adult human saliva and the approximate ratio between the relative percentages of the three components (S, S1, and S2) is 5:80:15, respectively [58,83]. A more complex situation concerns the phosphorylation of aPRPs, which are commonly di-phosphorylated, having Ser$_8$ and Ser$_{22}$ as the two main sites of phosphorylation [9,84]. However, by mass spectrometry, it is possible to detect also in small amounts non-, mono-, and three-phosphorylated aPRPs (on Ser$_8$, Ser$_{22}$, and Ser$_{17}$) [9] (Figure 2). The consensus sequence recognized for Ser$_8$ phosphorylation is the canonical S_8QE, with a negatively charged residue at position +2, while the phosphorylation of Ser$_{22}$ by Fam20c is ensured by the recognition of the secondary consensus sequence $S(X)_{3-4}$ (D/E/S(phos))$_3$ [84]. To emphasize that the phosphorylation of Ser$_{22}$ allows for the hierarchical phosphorylation of Ser$_{17}$, which is located in the sequence ...$S_{17}DGGDS_{22}EQFIDEE$... [9]. Hence, the mono-phosphorylated aPRPs can be phosphorylated either on Ser$_8$ or on Ser$_{22}$ (with a different proportion in favor of Ser$_8$) [9]. The di-phosphorylated components (the most abundant) are commonly phosphorylated on Ser$_8$ and Ser$_{22}$, but a very small percentage of aPRPs phosphorylated on Ser$_{17}$ and Ser$_{22}$ could be present. The hierarchical phosphorylation on Ser$_{17}$ is further substantiated by a study characterizing the aPRP-1 Roma–Boston Ser$_{22}$(phos) → Phe variant, which was never detected as di-phosphorylated proteoform, also using high-resolution HPLC-MS apparatus [85]. Statherin is mainly di-phosphorylated on Ser$_2$ and Ser$_3$ but with the N-terminal sequence being DS_2S_3EE..., the two phosphorylation sites are independent and the phosphorylated Ser of the mono-phosphorylated proteoform, always detectable in adult human saliva, can be either Ser$_2$ or Ser$_3$. The N-terminal sequence of Hst1 is DS_2HE..., and about 90% of the peptide is phosphorylated on Ser$_2$, and 10% of the non-phosphorylated peptide is detectable in adult human saliva [49]. Even though the Fam20C consensus sequence was not respected, a study by Halgand [86] and colleagues, carried out with a top-down pipeline, evidenced a second minor phosphorylation site on Ser$_{20}$.

The canonical consensus sequence of Fam20C is also responsible for the phosphorylation of several bPRPs. Indeed, the IB-1 and II-2 proteoforms are both phosphorylated on Ser$_8$. Among the gPRPs Gl-1, Gl-2, Gl-3, GPA, II-1, and Cd-IIg, which have an N-

terminal sequence motif similar to IB-1 and II-2, ((E/Q)XXXEDVS$_8$QEES..., where XXX is LNE in IB-1, II-2, Gl-1, Gl-2, and Gl-3 and SSS in GPA, II-1, and Cd-IIg), Gl-2 is phosphorylated on Ser$_8$ [23]. Phosphorylation is an almost complete event because <1% of the non-phosphorylated forms can be detected in parotid granules, parotids, and whole saliva and probably occurs after the cleavage of the proprotein [6]. It can be supposed, by sequence similarity, that Gl-1 and Gl-3 also undergo the same PTMs, although experimental evidence is missing. The presence of the ...S8QE consensus sequence of GPA, II-1, and Cd-IIg (Group 2C of bPRPs) also suggests the phosphorylation of Ser$_8$ for these bPRPs, even though these modifications have not been experimentally evidenced until now. A second potential, but not demonstrated, phosphorylation site at Ser$_3$ is present in the sequence of Group 2C bPRPs (<ESS$_3$SED...). The activity of other kinases can be revealed in human saliva by using a top-down proteomic approach, among them MAPK14, a kinase pertaining to the p38 mitogen-activated protein kinase pathway that can partly phosphorylate the protein S100A9 on the penultimate Thr residue of its sequence [87].

Variation of Phosphorylation as a Function of Age and for the Diagnosis of Different Disease

Some bottom-up studies suggested that the analysis of the phospho-proteome of salivary extracellular vesicles could offer a possibility for either the diagnosis of lung cancer [88] or to distinguish oral squamous cell carcinoma patients from healthy individuals [89]. Our studies were able to evidence that during the last months of foetal growth, the phosphorylation of secretory salivary proteins, and therefore the activity of Fam20C, is very low, if not completely absent [90]. It increases slowly, reaching the level observed in adults a few weeks after the normal time of delivery [25,91]. Interestingly, a top-down study of the salivary proteome performed in a group of children with autism spectrum disorders evidenced significantly lower phosphorylation levels of four salivary peptides, in comparison with age- and gender-matched healthy controls, providing a clue as to the molecular pathogenesis responsible for these disorders [92].

4. Sulfation

Sulfonation (or sulfation) is usually an irreversible PTM. It consists of the transfer of a sulfate group (-SO$_3^{-1}$) from the only known sulfate donor, i.e., 3′-phosphoadenosine-5′-phospho-sulfate (PAPS), to endogenous substances such as proteins, carbohydrates, and catecholamines as well as estrogenic steroids and xenobiotics [93]. Moreover, 3′-phosphoadenosine-5′-phosphosulfate synthase (PAPSS) is the enzyme responsible for biosynthesizing PAPS during two reactions: inorganic sulfate is first converted to adenosine-5-phosphosulfate (APS) by ATP sulfurylase (EC 2.7.7.4), and this intermediate molecule is then phosphorylated by the APS kinase (EC 2.7.1.25) to form PAPS [94]. Both APS and PAPS are activated sulphuryl donors that possess a phospho-sulfate anhydride bond [95]. In prokaryotes, fungi, and plants, the synthesis of PAPS is performed by two separate enzymes [94]. In the animal kingdom, however, the ATP sulfurylase and the APS kinase are encoded by the same gene and translated into a single polypeptide that forms the dual-functional enzyme PAPSS [95]. PAPSS1 and PAPSS2 are two characterized isoforms of this enzyme, according to their different localization. Moreover, a relation of various pathological conditions to deficiencies of PAPSS (both isoforms) has been demonstrated [96]. While PAPSS1 and PAPSS2 are responsible for the bioactivation of sulfate, sulfo-conjugation reactions are catalyzed by enzymes known as sulfotransferases [97]. Sulfotransferases are mainly divided into two groups, as they are either cytosolic or membrane-bound [98]. Cytosolic sulfotransferases constitute the superfamily of enzymes known as SULTs, which are involved in the sulfonation of xenobiotics and small endogenous compounds such as neurotransmitters and hormones [99]. The membrane-bound sulfotransferases are found in the Golgi apparatus and are responsible for the post-translational sulfation of endogenous macromolecules such as proteins, lipids, and glycosaminoglycans [100]. Currently, 12 SULT isoforms have been identified and detected in human tissues [101]. The availability of PAPS

in different tissues can highly affect the sulfation pathway by modifying the affinity of sulfotransferases [100–102].

In human saliva, we were able to determine that Hst1 is partly sulfated on the last four tyrosines (out of five) of its sequence [103], and until now, this was the only phospho-sulfo-peptide detected in human saliva. As previously reported, the phosphorylation of Hst1 is not a complete event, because in whole human saliva, it is possible to detect about 10% of the non-phosphorylated peptide. Further studies performed by high-resolution HPLC- MS apparatus suggested that: (a) as supposed in our work, the sulfation process is hierarchical, with Tyr_{27} being the first residue sulfated, followed by Tyr_{30}, $Tyr_{34,}$ and Tyr_{36}. Indeed, MS CID fragmentation data on the mono-sulfated (non-phosphorylated) Hst1 clearly indicated Tyr_{27} as the unique residue involved; (b) the sulfation process is strictly confined to the submandibular gland, because the sublingual gland does not express Hst1, and Hst1 secreted by the parotid gland is not sulfated; (c) the phosphorylation and the sulfation processes are independent, because, by the study of the neutral losses generated during the MS/MS CID-induced fragmentation, it is possible to discriminate non-phosphorylated poly-sulfated derivatives from the phosphorylated ones; (d) the percentages of the different polysulfated derivatives vary sensibly in the human submandibular saliva of humans; (e) the sulfation of Hst1 is not detectable in children until puberty; and (f) potential changes in physio-pathological conditions and during pharmacological treatments require extensive statistical analysis in a large population.

5. S-Modifications

The thiol side chain of cysteine residues present in the proteins is characterized by high redox sensitivity. This redox sensitivity derives from the pKa of the thiol group, which due to the environment of the protein, even at physiological pH, may be partly present in the thiolate form ($RS^−$) and more reactive than the thiol form (RSH). Indeed, $RS^−$ is prone to donating an electron pair to target reagents, generating the oxidized derivative sulfenic acid (RSOH), mainly present as sulfenate ($RSO^−$), by a reversible reaction. Then, sulfenate may undergo other reversible and irreversible oxidative modifications (Figure 5) [104]. For instance, it can rapidly react with another thiols (RSH) by generating the corresponding disulfide (RSSR), or with GSH or free Cys to form S-glutathionyl (RSSG) or S-cysteinyl derivatives. This is not the only formation pathway for disulfide derivatives, since the latter can also be generated in a process termed thiol-disulfide exchange, in which a thiol form reacts with a disulfide to generate a different thiol and disulfide.

In the presence of strong oxidants, sulfenate can undergo further oxidation and generate sulfinic (RSO_2H) and sulfonic acid (RSO_3H) derivatives sequentially. The latter two modifications are regarded as irreversible and are associated with oxidative damage. Sulfenate can also react with hydrogen sulfide ($HS^−$) to form a cysteinyl persulfide ($R–SS^−$), which can be oxidized to cysteinyl thiosulfate($R–SSO_3^−$). The thiol side chain of cysteine may also react with nitric oxide, resulting in the formation of nitrosothiol derivatives [105]. These modifications have been observed and characterized on specific salivary proteins as cystatins and S100A proteins. Cabras and colleagues evidenced that in human whole saliva, cystatin B is present mostly as S-modified derivatives on Cys_3, being the S-unmodified proteoform rarely detectable in the saliva of healthy adults. Generally, more than half of cystatin B (55%) is found to be S-glutathionylated, 30% is present in the dimeric form, and the remaining 15% is S-cysteinylated [106]. In an in-depth study, performed both in the acidic supernatant of whole saliva and in enriched fractions obtained by preparative RP-HPLC, the structures of all these S-modified proteoforms of cystatins were confirmed by an integrated top-down/bottom-up proteomic approach [57]. In the study, a carboxymethylated Cys_3 derivative of cystatin B was also detected and characterized. Carboxymethylation, like the formation of sulfinic and sulfonic acid derivatives, is a nonenzymatic, irreversible, and stable modification generated by the endogenously formed glyoxal with Cys sulfhydryl groups (Figure 5). This modification was novel not only for cystatin B but also for any other human salivary protein. It should be outlined that Cys residue present in the N-terminal

region at the C3 position of cystatin B is involved in the formation of these and other derivatives. The N-terminal region of cystatin B is crucial for the biological function of the protein, being involved in the binding of cysteine proteinases [107]. Indeed, it has been shown that Cys_3 is the most important residue for the interaction with papain and cathepsin H and is also a major contributor to cathepsin-L binding [108]. Differently from adults, the S-unmodified cystatin B represented the main proteoform in the saliva of preterm newborns. Interestingly, it was observed that the high relative amount the unmodified cystatin B observed in preterm newborns decreased as a function of the PCA, reaching values like those determined in at-term newborns, children, and adults at the normal term of delivery [60]. As reported in Section 2.1.6 when describing the proteolytic cleavages, in the same study, high relative amounts of fragments 1−53 and 54−98 of cystatin B were detected in very-preterm newborns. The N-terminal 1–53 fragment was detected both unmodified and modified at the level of Cys_3 residue, with the percentage of the different forms being like those determined for the intact protein [60].
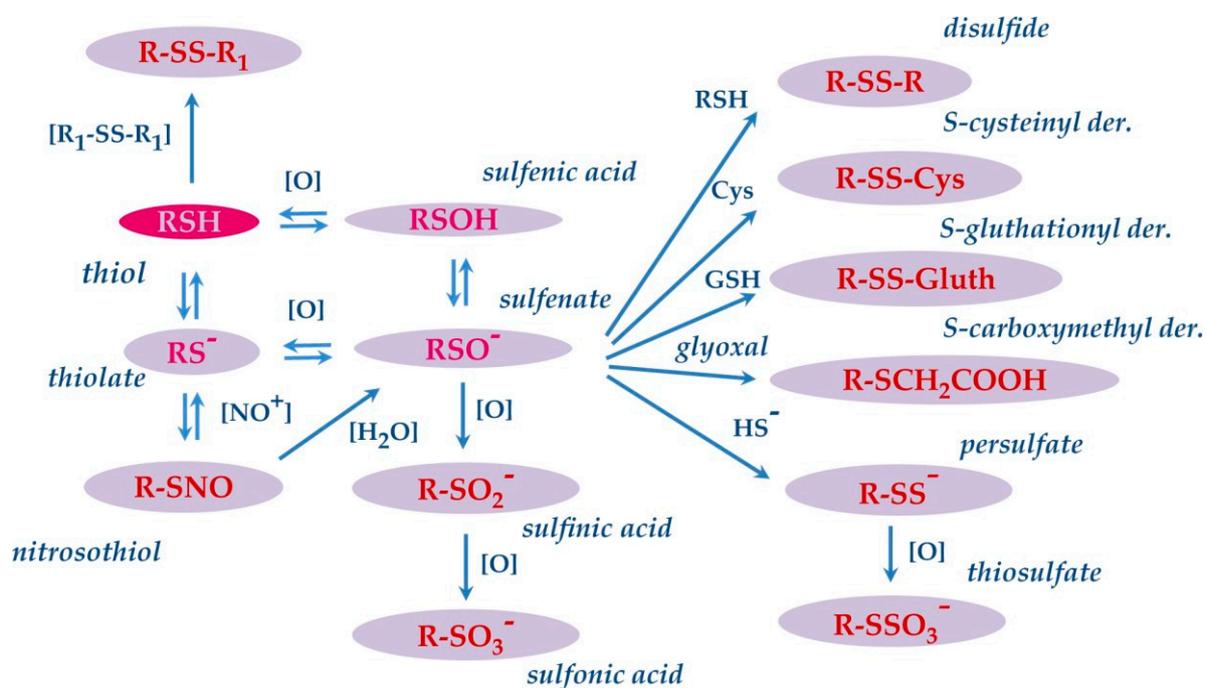
Figure 5. Map of reversible and irreversible oxidative modifications that can occur on the thiol group of the Cys residue.

S100 proteins constitute the largest family of calcium-binding proteins, with more than 3000 related entries in the NCBI Reference Sequences Data Bank. They are EF-hand calcium-binding proteins and bind calcium via helix–loop–helix motifs often present in multiple copies. The "S100" name originates from the solubility of the first identified S100 proteins in 100% ammonium sulfate solution. The members of the EF-hand superfamily can be divided according to their calcium affinity and their ability to change conformation following the binding of calcium. Intracellular functions of S100 proteins include: (i) regulation of protein phosphorylation by interaction with the substrates of the kinases, thus playing a role in signal transduction; and (ii) regulation of enzyme activity [109].

In the extracellular milieu, S100 proteins do not function as Ca^{2+} sensors as they are saturated by the mM range Ca^{2+} concentration; however, they are recognized as playing an important role in mediating inflammatory responses through the activation of several cell surface receptors, after their release from activated or necrotic cells [110]. S100A8, S100A9, and S100A12 are constitutively expressed in high amounts in neutrophils and are inducible in macrophages (cells that generate high amounts of ROS), while the expression

of the complex S100A8/S100A9 (calprotectin), S100A12, and S100A7 can be induced in keratinocytes, endothelial cells, and epithelial cells during inflammation [111].

Human S100A9 is encoded by a single copy gene with two isoforms—full-length and -truncated S100A9—and is translated from an alternate start site at codon 4 of the full-length form and lacks the single Cys_3 residue, making it less susceptible to oxidation. In the studies of Lim and colleagues [112,113] the many pro-inflammatory functions described for S100A8 and S100A9, as well as their anti-inflammatory roles in wound healing and protection against excessive oxidative tissue damage, are discussed, and an explanation that oxidative modifications may act as a regulatory switch for the disparate functional roles of S100A8 and S100A9 is suggested.

Concerning biofluids, in a series of studies carried out to investigate possible salivary biomarkers of pathologies both confined to the oral cavity and those systemic in nature, different oxidized derivatives of S100A8 and S100A9 proteins were revealed and characterized both in healthy and diseased subjects. The structure of these new proteoforms was established by high-resolution HPLC-ESI-MS/MS analysis of the mixture of peptides, obtained by trypsin digestion of salivary fractions enriched with S100A8 and S100A9 oxidized derivatives [68]. In this study, the long form of S100A9 was detected as being glutathionylated at Cys_3 in 15/32 subjects and cysteinylated only in 3/32, while S100A8 was sporadically detected as unmodified S100A8 (4/32), and as a sulfonic derivative at Cys_{42} (5/32).

Glutathionylated and cysteinylated S100A9 derivatives were also detected in the saliva of human preterm newborns [114]. Finally, the nitrosylated derivative of S100A8 was also observed in saliva from adult healthy subjects, but sporadically [73,74].

S-Modifications of Salivary Proteins in Pathologies

A wider variety and abundance of S-modified proteoforms have been revealed in the saliva of subjects affected by various pathologies. For instance, a study performed on saliva collected from patients with schizophrenia and bipolar disorder, compared to healthy non-smokers and smokers control groups, revealed a more than 10-fold increase in salivary levels of S-cysteinylated and S-glutathionylated cystatin B, in addition to α-defensins 1–4, S100A12, and cystatin A, suggesting the dysregulation of the peripheral white-blood-cell immune pathway associated with the pathologies [115]. S-glutathionylated cystatin B was also found at high levels, predominantly in antibody deficiencies [116]. In WD patients, top-down proteomic analysis of saliva revealed significantly higher levels of S100A9 and S100A8 and some of its oxidized proteoforms, with respect to controls [68]. S100A8 oxidizing at Cys_{42} to sulfinic acid ($S100A8-SO_2H$) was only detected in patients, while the sulfonic derivative ($S100A8-SO_3H$) oxidizing at Trp_{54} was detected both in patients and healthy controls, even if with different levels. The latter proteoforms were subjected to further oxidation to give the so-called hyperoxidized S100A8, with the second oxidation located at either Met_1 or Met_{78} (position to be determined). S100A8 was also shown to undergo glutathionylation, nitrosylation, and the formation of a disulfide bridge with the Cys_3 of long S100A9 (S100A8/A9-SS dimer), but only in the patient group. In the same group, for the first time, the homodimer proteoform of long S100A9 was characterized. Overall, the salivary proteome of WD patients reflected the oxidative stress and inflammatory conditions characteristic of the pathology, highlighting differences that could be useful clues of disease exacerbation [68].

Oxidative cross-linking via disulfide bonds of S100A9 and S100A8 has also been observed by Hoskin et al. in saliva from healthy adults and in lavage fluid from the lungs of patients with respiratory diseases [117]. The authors showed that reactive halogen species promote the cross-linking of the non-covalent heterodimer of S100A8/S100A9 and hypothesized that the cross-linking detected in the saliva samples was most likely mediated by hypothiocyanous acid produced by lactoperoxidase. They also demonstrated that the formation of the disulfide cross-linked derivative enhanced susceptibility to proteolysis by neutrophil proteases. Furthermore, Gomes et al. observed that sulfinic and sulfonic

acid derivatives of monomeric S100A8, together with novel oxathiazolidine oxide/dioxide forms, were present in asthmatic sputum [118].

The salivary proteome of Alzheimer's disease (AD) patients also highlighted elevated levels of some S100A8 and S100A9 oxidized proteoforms with respect to age- and gender-matched healthy controls. They were the hyperoxidized proteoforms of S100A8, S100A8-SNO, and glutathionylated long S100A9 [77]. This finding was not surprising since oxidative modifications of proteins are common in neurological disorders, due to the strongly oxidizing characteristics of the extracellular milieu because of the generation of ROS and re-active nitric oxide species [119]. Higher levels of S100A8-SNO were also determined in MuSc patients with respect to healthy controls, while levels of mono- and di-oxidized cystatin SN, mono- and di-oxidized cystatin S1, and mono-oxidized cystatin SA were lower [73]. The reduced level in the patients of oxidized derivatives of S-type cystatins was an intriguing result. Indeed, high levels of oxidative stress markers and lower antioxidant status have been reported in the saliva and plasma of MuSc patients under corticosteroid therapy by Karlik and colleagues [120]. However, different results were obtained by Schipper and colleagues [121], which highlighted a positive effect of immunotherapy on oxidative stress, evidenced by a reduction of oxidative stress markers in MuSc patients. Manconi et al. speculated that reduced levels of cystatin oxidation in MuSc patients could be related to the treatment of 32/49 patients enrolled in the study. Indeed, S100A8-SNO was found at a significantly higher level only in the group of untreated patients with respect to controls [73].

Recently, analysis of the salivary proteome of patients affected by AD compared to healthy adult and elderly controls has highlighted significantly higher levels of S100A8-SNO and hyperoxidized S100A8, as well as glutathionylated S100A9 (long), in healthy adults with respect to the elderly, and in AD patients with respect to healthy elderly individuals. The same trend was observed for both glutathionylated and dimeric cystatin B, but not for the cysteinylated proteoform, which showed a significant difference only by comparing patients and elderly subjects [75].

6. Transglutamination

Several salivary proteins are involved in the formation of proteins layers, i.e., the so-called "oral pellicles" The general term for these layers is pellicles, but due to the different characteristics of the coated surfaces, the enamel pellicle and mucosal pellicle are their own entities. These protein films have a dual role because one, "the acquired enamel," is important for the integrity of the tooth, and the second, the "mucosal pellicle," is important for the protection of the oral mucosa. There is considerable information on the enamel pellicle: many proteins and their fragments are involved with a particular concern for PRPs, statherin, histatins, and the P-B peptide [122,123]. On the contrary, only limited data are available on the mucosal pellicle [124]. This can be attributed to the difficulty to develop a standardized preparation of this latter biological structure. However, it has a completely different ultrastructure as compared with the enamel pellicle. Since it is comprised of larger glycoproteins retaining water, it might be considered a hydrogel, and it appears to have a lower tenacity than the enamel pellicle. Maturation and turnover are influenced by the delivery of salivary proteins, the flow of saliva, and the underlying desquamating oral epithelium. Its probable functions include lubrication and moisture retention. Furthermore, interactions between mucosal pellicle proteins and bacterial surfaces are responsible for the specificity of the bacterial colonization during the earliest stage of plaque formation [125]. The in vivo pellicle is thought to be an insoluble network of proteins generated by the post-secretory processing of proteins, mainly due to cross-linking. Cross-links for the formation of oral mucosal pellicles were demonstrated firstly by Bradway and colleagues, which highlighted the existence of a network of proteins formed by components of saliva adsorbed onto buccal epithelial-cell surfaces that cover the oral mucosal surface [126]. Recently, a new proteomic protocol was optimized to investigate the proteins participating

in the composition of the oral mucosal pellicle, and among them, the proteins of the PLUNC family were identified [127].

The oral mucosal pellicle is a thin lubricating layer generated by the binding of saliva proteins on epithelial oral cells [128]. This protein molecular network interacts with the oral epithelial-cell plasma membrane and its associated cytoskeleton and contributes to mucosal epithelial flexibility and turnover. It was demonstrated that acidic-proline-rich proteins, statherin, the major histatins, and mucins are substrates of oral transglutaminase 2 (TG2) and participate in cross-linking reactions [126] as putative pellicle precursor proteins. Whatever the structure of these protein networks may be, oral transglutaminases (mainly type-2 transglutaminase) are the pivotal enzymes for pellicle formation. TG2 can cross-link acidic PRP-1 and statherin in vitro [122]. TG2 is the ubiquitous tissue enzyme also expressed in oral epithelial cells [129], which catalyzes different biological processes and generates a cross-link between two peptide chains, typically between the ε-amine of the lysine residue (acting as the lone-pair donor) and a glutamine residue (the lone-pair acceptor), and the reaction is accomplished by the loss of an ammonia molecule. TG2 is a Ca^{2+}-dependent enzyme released by the epithelial oral cells; it is negatively modulated by GTP [129] and affected by the reversible formation of an intramolecular disulfide bridge [130]. Recently, a study of our group [131] showed that also bPRPs and the P-C peptide are potential substrates of TG2. Nonetheless, they showed a very different reactivity for monodansyl-cadaverine (used as a lone pair donor). Mass-spectrometry analyses of the reaction products highlighted that P-C, P-H, and P-D (both Pro_{32} and Ala_{32} variants) peptides were active substrates of TG2; II-2 was less reactive, while P-F and P-J peptides showed negligible activity [131]. MS characterization suggested that the consensus sequence for the linking is connected more to the environment of glutamine residue than to the donor ammine. The pivotal residues characterized for P-H, II-2, and both variants of P-D evidenced . . .GNPQ. . . as the consensus sequence recognized by TG2 on bPRP peptides [131]. This consensus sequence is not present in P-F and P-J peptides and, probably for this reason, they were poor substrates for TG2.

P-C, P-H, and P-D peptides formed cyclo-derivatives after the TG2 reaction, and only specific glutamine residues were involved in the cycle formation and reacted with specific monodansyl-cadaverines [131]. The stereospecificity of TG2 was at first recognized on statherin, which under the action of TG2 forms a cyclic derivative involving almost only Gln_{37} (out of seven glutamine residues) and Lys_6 (as a unique lysine residue) in the sequence (Figure 4) [132]. The detection of a small amount of cyclo-statherin in adult-human whole saliva [132] and the high reactivity of secondary glutamine residues after the formation of cyclo-statherin and cyclo-P-C were suggestive of the in vivo formation of ring structures with a pivotal role in the architecture of the oral mucosal and enamel pellicles [133,134]. Whatever the molecular mechanism towards the structure of the oral pellicles, they have a relevant role in the protection of the mucosa due to the mechanical and thermal high stresses that the mouth undergoes during the human life.

7. Glycosylation

7.1. Non-Enzymatic Glycosylation (Glycation)

Non-enzymatic glycosylation or glycation is a ubiquitous process involving proteins, peptides, lipids, and nucleic acids. It consists of a nucleophilic addition of sugar (commonly glucose) to the free amino group of a biological molecule (in protein, typically the ε-amino group of lysine). The first studies regarding glycation processes were carried out by the food industry for the quality control of foodstuffs [135], and only forty years later, this process was observed in the human body for monitoring the trend of glycated hemoglobin in diabetic patients [136–138]. The glycation process starts with the formation of a reversible but unstable Shiff's base, followed by intermolecular rearrangement into stable Amadori products. When large amounts of Amadori products are formed, they rearrange to form a heterogeneous group of protein-bound moieties, termed advanced glycated end products (AGEs) [139]. Fluorescent spectroscopy is usually used to monitor the rate of formation of

AGEs, considering that most of the advanced products of glycation are fluorophores [140]. The rate of these reactions is quite slow, and only proteins with large amounts of lysine residues undergo glycation, with significant amounts of AGEs leading to reduced protein functionality and stability [141,142].

Proteins with a long metabolic turnover, as extracellular matrix proteins, are the most susceptible to non-enzymatic glycosylation, compared to most proteins that counteract the impact of glycation thanks to their high turnover rate and short half-life [143]. For instance, the fast hematic turnover of serum albumin offers the chance to monitor short-term diabetes [144]. It is known that proteins from saliva show a very quick turnover; nevertheless, few studies point out how the loss of functional activity of some salivary proteins is to be related to glycation processes. The antibacterial activity of lactoferrin or lysozyme and the antioxidant capacity of superoxide dismutase, all detectable in human saliva, decreased after the formation of AGEs [142,145,146].

AGEs accumulate in the human body by binding to AGE-specific receptors (RAGEs), which play an important role in the development of age-related organ hypofunction [139]. The presence of RAGEs was also evidenced in the oral cavity, particularly in the minor salivary glands [147]. Manig and colleagues, in 2019 [148], carried out the first study focused on monitoring the free salivary glycated compounds predominant during later reaction stages of AGEs production in saliva. Specifically, by LC-MS-MS analysis, the concentration of Nε-carboxymethyllysine (CML), pirralyne (Pyr), methylglyoxal-derived hydroimidazolone 1 (MG-H 1), and Nε-carboxyethyllysine (CEL) was evaluated [148]. The presence of AGEs in saliva increases the risk of developing periodontitis and pulpitis [147]. Furthermore, the expression of RAGEs appeared to be correlated with the histologic differentiation of oral squamous cell carcinomas [149]. Maciejczyk M. and colleagues [150] evidenced the role of saliva as a diagnostic biofluid alternative to blood for assessing the severity of protein glycation, correlated to the aging processes. Several proteomic studies, mainly performed by top-down mass-spectrometry approaches, have shown that saliva can be a mirror of the aging process. Indeed, age-related changes in salivary proteome profiles have been observed in individuals from 180 days to the elderly [75]. Among the salivary proteins and peptides affected by the aging process, S100A8 and S100A9 have been found to participate in the toll-like receptor-4- or RAGE-mediated multiple inflammatory pathways [75]. In fact, S100A8 and S100A9 represent the main protein content of the neutrophils and actively participate in modulating inflammation [151] by also acting as ROS/RNS scavengers. The maintenance of normal salivary function is important for a healthy oral environment. In this context, the study of glycated proteins and peptides in saliva could be, given the low-invasiveness and easy collection of this biofluid, a potential tool to study the aging process as well as disease progression or patient follow-up during treatments [152].

It is relevant to remark that improper sample manipulation can generate an artefactual detection of the non-enzymatic glycation of proteins. For instance, freeze-drying salivary solutions containing glucose and/or fructose can generate artefactual amounts of glycated proteins [22].

7.2. Enzymatic Glycosylation

Enzymatic glycosylation is one of the most-common and prominent PTMs, modulating the activity of cellular and extracellular proteins, implicated in complex and articulate physiological processes [153,154] such as cell adhesion, signaling, and cell-to-cell interaction, to cite a few, and many proteins detectable in human saliva are enzymatically glycosylated. The two principal types of protein glycosylation in humans are N- and O-glycosylation, while the C-mannosylation of tryptophan occurs rarely, with its function still not well understood [155].

N-glycosylation occurs as a co-translational modification on the Asn residue at the Asn-X-Ser/Thr sequon, even though some recent glycoproteomic studies evidenced that several N-glycosylation sites do not adhere to this canonical consensus sequence [156]. The

glycan structure is not under direct genetic control; rather, it results from the cooperative and competitive interactions between several glycosyltransferases. These enzymes transfer a monosaccharide unit from an activated sugar donor (frequently a nucleotide sugar) to an acceptor (such as an amino acid, a lipid, or another sugar). Consequently, glycosylation must be considered a stochastic rather than a deterministic process. The biological role of glycans is frequently mediated by sugar-binding molecules (lectins), which, upon recognition of specific carbohydrate structures, trigger a broad range of cellular effects, including proliferation, apoptosis, and cell migration [157]. Galectins (beta-galactoside binding proteins) and siglecs (sialic acid-binding immunoglobulin-type lectins) are among these sugar-binding molecules. The activities of specific galectins and siglecs can be inhibited or promoted by non-coding RNAs (ncRNAs), such as those implicated in tumor invasion and cell proliferation [158], as well as by many glycosyltransferases [158]. Indeed, ncRNAs form an extremely complex non-deterministic network of gene expression regulation, and they have a key role in modulating the expression of several glycosyltransferase genes, especially promoting those implicated in the progression and invasion of various kind of tumors [158]. The relevant role of several classes of ncRNAs is nowadays recognized in regulating both physiological and pathological mechanisms. Moreover, the ncRNA aberrant expression appears to be associated with different human disorders, such as reproductive disorders and cancer [159]. The term ncRNAs comprises a wide and heterogeneous group of RNA molecules differing in length, biological function, and cellular localization. ncRNAs are classified, based on their size, into long ncRNAs (lncRNA, >200 nucleotides) and short ncRNAs (<200 nucleotides), which include, among others, microRNAs (miR-NAs). Circulating miRNA are largely studied as the next-generation class of diagnostic and prognostic biomarkers. lncRNAs are tissue-specific and can directly interact with other macromolecules, such as DNA, RNA, and protein, and moreover, they can regulate gene expression both at the transcriptional and the post-transcriptional level.

In addition, glycosylation has been demonstrated, on the other hand, to modulate ncRNA expression in some instances. Finally, short ncRNAs have recently been shown to undergo canonical N-glycosylation [160]. These glyco-RNAs bear terminal sialic acid and fucose residues and can interact with sugar-binding molecules, such as siglecs.

Thus, the ncRNA network and glycosylation can be considered as two stochastic mechanisms affecting the role of proteins: the first and the second acting before and after the protein axis biosynthesis, respectively. The N-linked glycans present a common pentasaccharide core to which different types of monosaccharide residues can be linked in different positions, generating diverse classes of branched glycans [161].

O-glycosylation occurs on either Ser or Thr residue exposed on the protein surface. The O-linked glycans can be very simple and are constituted by one or a few monosaccharide residues, such as N-acetyl-galactosamine, mannose, glucose, and fucose. The glycosylation of salivary proteins was studied several years ago using a bottom-up pipeline, identifying about 1500 glycosites [162].

Due to the stochastic mechanism of the synthesis described above, glycoproteins are characterized by high microheterogeneity, because each glycoprotein may exist as a complex family of glycoforms sharing an amino-acid sequence and differing for various features: (a) occupancy of the multiple glycosites; (b) type of glycan (only N-linked, only O-linked, N- and O-linked on the same protein); and (c) composition and stereoisomery of the glycans that can change on the same glycosite.

We were able to characterize using an integrated top-down/bottom-up strategy the structure of six glycoforms of IB-8a (Con 1+), so called because they interact with concanavalin A [163]. MS analyses on the intact glycoproteins before and after N-deglycosylation with the PNGase F and MS/MS sequencing of peptides and glycopeptides from tryptic digests allowed for the structural characterization of the glycan moieties and the polypeptide backbone, as well as the establishment of the glycosylation site at the asparagine residue at the 98th position. Five of the glycoforms carry a biantennary N-linked glycan fucosylated in the innermost N-acetylglucosamine of the core and showed from zero to four

additional fucoses in the antennal region. The sixth glycoform carries a monoantennary monofucosylated oligosaccharide. The glycoform cluster was detected in 28 out of the 71 adult saliva specimens analyzed [163].

The level of fucosylation showed interindividual variability, with major relative abundance for the tri-fucosylated glycoform. Non-glycosylated IB-8a (Con 1+) and the variant IB-8a (Con 1−), lacking the glycosylation site for the substitution, have been also detected in human saliva.

The surprising detection of the gPRP protein 3M (encoded by the PRB3 human gene) in the saliva of preterm newborns at an age when the other bPRPs are absent in human saliva [20,25,26] allowed us to purify the 3M protein and to establish the structures of the intact proteoforms before and after N-deglycosylation with peptide-N-glycosidase F and the MS/MS sequencing of peptides obtained after endoproteinase GluC digestion [164]. The heterogeneous mixture of the proteoforms derives from the combination of 8 different neutral and sialylated glycans O-linked to Thr_{50} and 33 different glycans N-linked to various Asn residues at positions 66, 87, 108, 129, 150, 171, 192, and 213 of the 3M protein. In a study carried out to evidence potential salivary diagnostic markers of MuSc [72], we were able to evidence that prolactin-inducible protein exists in two glycosylated proteoforms: one with an N-acetyl-hexosamine residue and the other with a further deoxyhexose moiety. Although the site of modification was not established, these two glycosylated proteoforms were significantly increased in the patients with MuSc with respect to healthy controls. Indeed, altered glycoprotein expression and glycosylation aberrations, consisting of increased glycan size, extra branching of glycan chains, over-sialylation, or over-fucosylation, were associated with many diseases, especially cancer and neurological disorders [165,166]. A recent magistral review [158] evidences the complexity of the relationships between enzymatic glycosylation and other factors and their molecular alterations as potential diagnostic markers in many diseases. Therefore, glycoproteomic studies are of great interest for the development of disease biomarkers. Nonetheless, they are a complex and challenging field of research, requiring the evaluation of the glycoprotein profiling of a biological sample in physiological conditions and the characterization of the qualitative and quantitative changes related to a specific pathological condition [167,168].

8. Citrullination

Citrullination is a PTM consisting of the conversion of peptidylarginine to peptidyl-citrulline in a calcium-dependent reaction catalyzed by peptidylarginine deiminases (PADs), a family of five isoenzymes (PAD 1–4 and 6), with tissue-specific expression [169,170]. The reaction of the PAD enzyme with the arginine residue of a peptide/polypeptide chain forms an adduct, with the release of ammonia and the subsequent formation of a ketone functional group following the cleavage of the adduct by a water molecule. The result of the deimination reaction is the loss of a positive charge, i.e., the imino moiety, resulting in a generation of a neutral amino acid, citrulline, lacking the strong basic character proper of arginine. It produces the delta mass increase of the protein/peptide molecular mass of +0.9840276 and +0.98476 Da (monoisotopic and average, respectively). Interestingly, the rate of peptidylarginine citrullination by PADs enzymes depends on the amino-acid position along the protein/peptide sequence, with about 80–90% of citrullinated arginines positioned after aspartic acid residues. Arginines close to glutamic acid or to the N-terminal sequence trait are also largely citrullinated [169–171].

The citrullination affects the physico-chemical properties of a protein because, at a neutral pH, the positive charge of arginine is lost by the modification and causes changes in the overall charge and charge distribution of the protein, altering the isoelectric point, ionic bonds, protein structure, activity, and protein–protein interactions [169–171]. The increase in hydrophobicity produced by citrullination allows to easily distinguish the unmodified from the modified form of a small protein or a peptide by reverse-phase liquid chromatographic separations. In addition, because citrulline is not included in the list of natural amino acids incorporated into proteins, it was hypothesized to possibly induce

an immune response, and it was suggested to deeply investigate its role in autoimmune inflammatory diseases such as rheumatoid arthritis (RA) [172]. In fact, citrullination was studied in relation to various physio-pathological processes in addition to RA, such as apoptosis, multiple sclerosis, Alzheimer's disease, psoriasis, systemic lupus erythematosus, periodontitis, COVID-19, cancer, and thromboembolism [169–171]. Citrullination has been the topic of a very recent review that outlines the diagnostic as well as therapeutic approaches based on this PTM and the future perspectives of "citrullinome" characterization and disclosure [172]. Studies and clinical trials are nowadays based on the detection of anticitrullinated protein antibodies as a high-specificity tool for RA diagnosis, before clinical evidence [173,174].

Several are the proteins that are substrates of PAD enzymes and that can exhibit citrullination PTM in physiological, as well as in inflammatory, pathological states, and they include filaggrin, keratin, fibronectin, actin, tubulin, vimentin, glial fibrillary acidic protein, and histones [173]. PAD 2 is the isoenzyme acting in salivary and secretory glands, as well as in other cell types and tissues, targeting myelin basic protein, C-X-C motif chemokine 10 and 11, vimentin, actin, glial fibrillary acidic protein, S100-A3, and histones H3 and H4 [173]. The importance of the detection of anti-citrullinated protein antibodies (ACPAs) in serum for the diagnosis of RA [173,175] has increased the interest of their application to saliva analysis for the characterization of citrullinated proteins.

Yasuda et al. identified citrullinated cytokeratin 13 in saliva samples of RA patients and healthy subjects by two-dimensional electrophoresis, silver staining, and immunoprecipitation Western blotting [175]. In the same paper, the origin of the pre- or post-secretion of citrullinated cytokeratin 13 in saliva is also discussed based on calcium concentration data inside salivary gland cells and secreted saliva, since PAD enzyme catalysis is calcium-dependent. According to the Genotype-Tissue Expression database, PAD 1–4 and six calcium-binding enzymes are reported to be expressed in human minor salivary glands [176]. Citrullination could take place either in saliva, i.e., post-secretion because the biofluid contains calcium levels proper to PAD enzyme activity, or intracellularly, dependent on the calcium concentration increasing. Indeed, intracellular citrullination of histones and other proteins was recognized [177]. Citrullination of salivary proteins could therefore occur either intracellularly, followed by the secretion of the modified protein, or extracellularly in the biofluid, after secretion.

A recent paper compared the levels of citrullinated proteins levels in the saliva of RA patients with respect to healthy controls, also studying their correlation with periodontal status and a temporomandibular joint-associated disorder [178]. Confirming previous data [175], no differences in citrullinated protein levels were found between RA patients and healthy controls, apart from the differences in periodontal status observed between groups. However, in RA, the salivary anti-cyclic citrullinated peptide levels were in correlation with periodontal disease staging, considering that the bacteria involved in its pathogenesis produce non-physiological citrullination by a family of PAD enzymes very similar to human PADs [179].

One study investigated the influence of inflammation from periodontal disease using RA-model mice analysis in serum and saliva samples of both anti-citrullinated proteins (using antibody analysis) and citrullinated proteins (using gel electrophoresis) [180]. Saliva and serum showed similar results, evidencing protein citrullination around the 55 kDa molecular-weight electrophoretic band. Periodontitis exacerbated RA symptoms, indicating a relationship between the *Porphyromonas gingivalis* periodontitis bacterial infection and RA.

Patients affected by Sjögren's syndrome have been found to be positive for ACPAs and showed increased levels of the PAD 2 enzyme, suggesting a possible role of citrullination in the disease that could be responsive to autoantigens triggering [180]. However, in the same study, it was outlined that only approximately 7.5% of Sjögren's syndrome patients are positive for ACPAs. Despite the poor knowledge of the protein pattern that led to citrullination in saliva, the indications of the high potential of this biofluid are clear in terms of the development of new high-specificity and -sensitivity diagnostic tools of low

invasiveness targeting selected citrullinated proteins in different pathologies. However, deeper research still needs to be conducted to decipher the saliva citrullinome associated with physiological as well as pathological states and to clearly understand where the modification occurs, i.e., inside salivary gland cells, extracellularly, or both. To the best of our knowledge, combined top-down/bottom-up proteomic platforms have never been applied to the study of the saliva citrullinome. Their application coupled with high-resolution mass-spectrometry detection could finally characterize the proteins subjected to citrullination in saliva in a wide molecular range and abundance, and, especially, could localize the modification inside the sequence and better define its molecular features.

9. N-Terminal Modifications

The N-terminal of a polypeptide chain is a structural site that must often be modified and protected during ribosome synthesis. The positive charge of the terminal amino group can be an obstacle to the proper structural folding of the protein, which starts in the nascent protein [181]. The most common modification in humans is the excision of the initiatory methionine (iMet), which is commonly carried out by enzymes called methionine aminopeptidases (MAPs). In any case, the polypeptide chain, with or without iMet, can be submitted to N-terminal acetylation induced by enzymes called N-terminal acetyltransferases (NATs). To date, 12 NATs have been identified, possessing different compositions, substrate specificities, and modes of regulation. In humans, NATs are present as a group of six enzymes (called from NAT-A to NAT-F), with slightly different specificities [182].

Another possibility, less common, of N-terminal modification occurs when the terminal residue is either a glutamic acid or a glutamine, which can generate a cyclic amide, called pyroglutamic N-terminus, with the loss of a molecule of water (Glu) or ammonia (Gln), respectively. This PTM can proceed spontaneously, at reasonable rates, or by the action of enzymes called glutaminyl cyclases, in a reaction commonly much faster with N-terminal glutamines than with glutamate residues [183].

Our top-down mass-spectrometry pipelines allowed us to verify that these three PTMs are largely detectable in human saliva, while C-terminal modifications are rarely detectable, probably because at the end of their synthesis the proteins have assumed a folding close to the functional conformation. A few minor examples of C-terminal modifications of human salivary proteins will be reported at the end of this section.

9.1. Excision of Initiatory Methionine and Nt-Acetylation

It is relevant to outline that when the methionine residue is at the first position of a leader peptide, in turn responsible for the docking of the protein to the ribosome at the endoplasmic reticulum, acetylation is not carried out. This is observed for cystatins A, B, and S-type (S, SA, and SN), and all the PRPs, suggesting the absence of NATs in the cisternae of the cytoplasmic reticulum. In all other proteins, the action of MAPs is not mandatory and seems to depend on the type of the residue present in the second position in the sequence. Therefore, Nt-acetylation may occur either on the iMet or on the first residue after iMet excision by MAPs. The N-termini of proteins with small amino-acid residues (Ser, Ala, Thr, Val, Gly, and Cys) in the second position are mostly processed by MAPs, and the newly generated N-termini may be acetylated by NAT-A, as observed for the serine residue of thymosins β4 and β10; of S100A7; and of the two proteoforms of small proline-rich protein 3 [87,184–186]. The N-termini of proteins with larger amino-acid residues in the second position are not cleaved by MAPs but potentially acetylated directly on the iMet by a variety of NATs, depending on the N-terminal sequence. NAT-B potentially acetylates Met-N-termini when the second residue is Asp, Glu, or Asn. NAT-C potentially acetylates Met-N-termini when the second residue is a hydrophobic amino acid (mainly Leu, Ile, and Phe). Hydrophobic terminals are also recognized by NAT-E and NAT-F, suggesting that redundancy in activity exists between NATs [187]. Moreover, NAT-F showed a broader specificity than NAT-C in acetylating Met-N-terminals when the second residue was Met,

Lys, or Gln, and the potential to acetylate the same N-terminals was recognized by NAT-A in proteins in which iMet has not been cleaved [188]. Examples of human salivary proteins are represented by cystatin A, which has an isoleucine residue in position 2 and is probably acetylated on Met1 by NAT-C (or NAT-F), and cystatin B, which has a second methionine residue in position 2 and is acetylated on Met_1 probably by NAT-F [57,87].

A peculiar example concerns the acetylation of S100A9, which among the five N-terminal residues has two Met residues—one at position 1 and the other at position 5 (MTCKM. . .). Because of this N-terminal sequence, MAP catalyzes the cleavage not only at the level of the first Met residue, generating the long proteoform (1–113 residues), but also at the level of the fifth Met, giving rise to the short proteoform (1–109 residues), with the percentages of the two proteins being similar in human saliva [25,87]. The Cys_3 residue, present only in the long form, is responsible for the formation of the S-cysteinyl and S-glutathionyl derivatives of S100A9, described in Section 5. Since these four proteoforms of S100A9 can also be phosphorylated in the penultimate threonine residue of the sequence, the number of possible proteoforms of S100A9 detectable in human saliva is eight [188]. As far as we are aware, no information exists about the specific function of these different proteoforms of S100A9.

9.2. Pyroglutamic Acid Modification

As above reported, several human salivary proteins undergo the N-terminal cyclization of glutamine and glutamic acid as a protection for the N-terminal amino group. This PTM has been characterized for all the proteoforms of aPRPs (PRP-1 and PRP-3 types) [90,189], bPRPs (II-2, IB-1), all the gPRPs [23], and the P-B peptide [41]. It is relevant to remark that although this PTM allows for the integrity of the protein structure by protecting it from the action of aminopeptidases [190], on the other hand, it significantly alters its hydrophobicity and solubility [183], facilitating protein aggregation. For this reason, many authors hypothesized that it plays a relevant role in several amyloidopathies [191]. Indeed, experiments carried out in a mouse model using glutaminyl cyclase inhibitors demonstrated a reduced level of pGlu-modified Aβ peptides that appeared to attenuate Alzheimer's disease [192]. Nonetheless, the enzyme does not seem essential for the cyclization, which according to Bersin and colleagues [193], is spontaneous and is also observed in the absence of the enzyme.

10. C-Terminal Modifications

Except for the removal of C-terminal residues by carboxy-peptidases, described in Section 2.1.2, only three C-terminal modifications were detected during our top-down experiments. One is the formation of a cyclic anhydride bond for the loss of one water molecule from two carboxylic groups in the presence of a C-terminal Glu residue of the fragment 1–26 of β-globin; the second is the formation of one cyclic ester for the loss of a water molecule from the C-terminal serine of the fragment 1–42 of P-C peptide; and the third is the formation of a cyclic N-substituted amide for the loss of one water molecule from the C-terminal glutamine of the fragment 1–41 of the P-C peptide.

11. Conclusions

This review, which provides many answers regarding the complexity of the PTMs that human salivary proteins undergo, nevertheless prompts even more questions regarding the roles of many of them.

The first question concerns the possible functional role played by the products generated by the multiple proteolytic cleavages to which the components of almost all protein families are subjected.

It is relevant to underline that they could be evidenced only by a top-down strategy. Since evolution is addressed towards the selection of molecular events relevant to the functions of cells, organs, and tissues, particularly puzzling is the functional meaning of the cleavages that undergo acidic and basic PRPs, as well as the challenging role of

the P-B peptide. Particularly challenging is defining the role of the myriad recurrent fragments generated from bigger salivary proteins by exogenous proteinases. Therefore, the development of new strategies to study the biological role of the fragments is demanding.

The role of almost all the other PTMs described is also obscure. Proteomic searches, devoted to the discovery of new disease biomarkers, are observational studies, supported by powerful statistical tools; they are able to suggest potential biomarkers but not able to suggest their specific physio-pathological roles. Therefore, new experimental designs are needed to help researchers connect the information obtained on structural modifications and the role of these changes in health and disease. This knowledge is fundamental for the biotechnological utilization of biologically active salivary proteins [194].

The recent pandemic evidenced the utility of saliva as a human biofluid for painless, non-invasive, free, at-home diagnostic purposes [1,195,196]. Indeed, saliva is currently also used for the fast check of illicit drug abuse on the road [196]. We hope that this review could be a stimulus for further investigations and clinical applications.

Author Contributions: Conceptualization, M.C., I.M., G.F., T.C., A.F. and B.M.; data curation, all authors; writing—original draft preparation, all the authors; writing—review and editing, all the authors. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

References

1. Boroumand, M.; Olianas, A.; Cabras, T.; Manconi, B.; Fanni, D.; Faa, G.; Desiderio, C.; Messana, I.; Castagnola, M. Saliva, a bodily fluid with recognized and potential diagnostic applications. *J. Sep. Sci.* **2021**, *44*, 3677–3690. [CrossRef]
2. Helmerhorst, E.J.; Oppenheim, F.G. Saliva: A dynamic proteome. *J. Dent. Res.* **2007**, *86*, 680–693. [CrossRef]
3. Lau, W.W.; Hardt, M.; Zhang, Y.H.; Freire, M.; Ruhl, S. The Human Salivary Proteome Wiki: A Community-Driven Research Platform. *J. Dent. Res.* **2021**, *100*, 1510–1519. [CrossRef] [PubMed]
4. Iavarone, F.; Desiderio, C.; Vitali, A.; Messana, I.; Martelli, C.; Castagnola, M.; Cabras, T. Cryptides: Latent peptides everywhere. *Crit. Rev. Biochem. Mol. Biol.* **2018**, *53*, 246–263. [CrossRef] [PubMed]
5. Castle, D.; Castle, A. Intracellular transport and secretion of salivary proteins. *Crit. Rev. Oral Biol. Med.* **1998**, *9*, 4–22. [CrossRef] [PubMed]
6. Messana, I.; Cabras, T.; Pisano, E.; Sanna, M.T.; Olianas, A.; Manconi, B.; Pellegrini, M.; Paludetti, G.; Scarano, E.; Fiorita, A.; et al. Trafficking and postsecretory events responsible for the formation of secreted human salivary peptides: A proteomics approach. *Mol. Cell. Proteom.* **2008**, *7*, 911–926. [CrossRef]
7. Chan, M.; Bennick, A. Proteolytic processing of a human salivary proline-rich protein precursor by proprotein convertases. *Eur. J. Biochem.* **2001**, *268*, 3423–3431. [CrossRef]
8. Cai, K.; Bennick, A. Processing of acidic proline-rich proprotein by human salivary gland convertase. *Arch. Oral Biol.* **2004**, *49*, 871–879. [CrossRef]
9. Inzitari, R.; Cabras, T.; Onnis, G.; Olmi, C.; Mastinu, A.; Sanna, M.T.; Pellegrini, M.G.; Castagnola, M.; Messana, I. Different isoforms and post-translational modifications of human salivary acidic proline-rich proteins. *Proteomics* **2005**, *5*, 805–815. [CrossRef]
10. Kauffman, D.L.; Keller, P.J. The basic proline-rich proteins in human parotid saliva from a single subject. *Arch. Oral Biol.* **1979**, *24*, 249–256. [CrossRef]
11. Kauffman, D.; Wong, R.; Bennick, A.; Keller, P. Basic proline-rich proteins from human parotid saliva: Complete covalent structure of protein IB-9 and partial structure of protein IB-6, members of a polymorphic pair. *Biochemistry* **1982**, *21*, 6558–6562. [CrossRef] [PubMed]
12. Kauffman, D.; Hofmann, T.; Bennick, A.; Keller, P. Basic proline-rich proteins from human parotid saliva: Complete covalent structures of proteins IB-1 and IB-6. *Biochemistry* **1986**, *25*, 2387–2392. [CrossRef]
13. Kauffman, D.L.; Bennick, A.; Blum, M.; Keller, P.J. Basic proline-rich proteins from human parotid saliva: Relationships of the covalent structures of ten proteins from a single individual. *Biochemistry* **1991**, *30*, 3351–3356. [CrossRef]
14. Saitoh, E.; Isemura, S.; Sanada, K. Complete amino acid sequence of a basic proline-rich peptide, P-D, from human parotid saliva. *J. Biochem.* **1983**, *93*, 495–502. [CrossRef]

15. Saitoh, E.; Isemura, S.; Sanada, K. Complete amino acid sequence of a basic proline-rich peptide, P-F, from human parotid saliva. *J. Biochem.* **1983**, *93*, 883–888. [CrossRef]

16. Saitoh, E.; Isemura, S.; Sanada, K. Further fractionation of basic proline-rich peptides from human parotid saliva and complete amino acid sequence of basic proline-rich peptide P-H. *J. Biochem.* **1983**, *94*, 1991–1999. [CrossRef]

17. Isemura, S.; Saitoh, E.; Sanada, K. Isolation and amino acid sequences of proline-rich peptides of human whole saliva. *J. Biochem.* **1979**, *86*, 79–86.

18. Isemura, S.; Saitoh, E.; Sanada, K. Fractionation and characterization of basic proline-rich peptides of human parotid saliva and the amino acid sequence of proline-rich peptide P-E. *J. Biochem.* **1982**, *91*, 2067–2075. [CrossRef] [PubMed]

19. Messana, I.; Cabras, T.; Inzitari, R.; Lupi, A.; Zuppi, C.; Olmi, C.; Fadda, M.B.; Cordaro, M.; Giardina, B.; Castagnola, M. Characterization of the human salivary basic proline-rich protein complex by a proteomic approach. *J. Proteome Res.* **2004**, *3*, 792–800. [CrossRef]

20. Cabras, T.; Pisano, E.; Boi, R.; Olianas, A.; Manconi, B.; Inzitari, R.; Fanali, C.; Giardina, B.; Castagnola, M.; Messana, I. Age-dependent modifications of the human salivary secretory protein complex. *J. Proteome Res.* **2009**, *8*, 4126–4134. [CrossRef] [PubMed]

21. Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O. Differential RNA splicing and post-translational cleavages in the human salivary proline-rich protein gene system. *J. Biol. Chem.* **1985**, *260*, 11123–11130. [CrossRef] [PubMed]

22. Messana, I.; Inzitari, R.; Fanali, C.; Cabras, T.; Castagnola, M. Facts and artifacts in proteomics of body fluids. What proteomics of saliva is telling us? *J. Sep. Sci.* **2008**, *31*, 1948–1963. [CrossRef] [PubMed]

23. Padiglia, A.; Orrù, R.; Boroumand, M.; Olianas, A.; Manconi, B.; Sanna, M.T.; Desiderio, C.; Iavarone, F.; Liori, B.; Messana, I.; et al. Extensive Characterization of the Human Salivary Basic Proline-Rich Protein Family by Top-Down Mass Spectrometry. *J. Proteome Res.* **2018**, *17*, 3292–3307. [CrossRef] [PubMed]

24. Manconi, B.; Castagnola, M.; Cabras, T.; Olianas, A.; Vitali, A.; Desiderio, C.; Sanna, M.T.; Messana, I. The intriguing heterogeneity of human salivary proline-rich proteins: Short title: Salivary proline-rich protein species. *J. Proteom.* **2016**, *134*, 47–56. [CrossRef]

25. Messana, I.; Cabras, T.; Iavarone, F.; Manconi, B.; Huang, L.; Martelli, C.; Olianas, A.; Sanna, M.T.; Pisano, E.; Sanna, M.; et al. Chrono-proteomics of human saliva: Variations of the salivary proteome during human development. *J. Proteome Res.* **2015**, *14*, 1666–1677. [CrossRef]

26. Manconi, B.; Cabras, T.; Pisano, E.; Sanna, M.T.; Olianas, A.; Fanos, V.; Faa, G.; Nemolato, S.; Iavarone, F.; Castagnola, M.; et al. Modifications of the acidic soluble salivary proteome in human children from birth to the age of 48 months investigated by a top-down HPLC-ESI-MS platform. *J. Proteom.* **2013**, *91*, 536–543. [CrossRef]

27. Lu, Y.; Bennick, A. Interaction of tannin with human salivary proline-rich proteins. *Arch. Oral Biol.* **1998**, *43*, 717–728. [CrossRef]

28. Cannon, R.D.; Chaffin, W.L. Oral colonization by Candida albicans. *Crit. Rev. Oral Biol. Med.* **1999**, *10*, 359–383. [CrossRef] [PubMed]

29. Ruhl, S.; Sandberg, A.L.; Cisar, J.O. Salivary receptors for the proline-rich protein-binding and lectin-like adhesins of oral actinomyces and streptococci. *J. Dent. Res.* **2004**, *83*, 505–510. [CrossRef] [PubMed]

30. Cabras, T.; Melis, M.; Castagnola, M.; Padiglia, A.; Tepper, B.J.; Messana, I.; Tomassini Barbarossa, I. Responsiveness to 6-n-propylthiouracil (PROP) is associated with salivary levels of two specific basic proline-rich proteins in humans. *PLoS ONE* **2012**, *7*, e30962. [CrossRef]

31. Vitorino, R.; Calheiros-Lobo, M.J.; Williams, J.; Ferrer-Correia, A.J.; Tomer, K.B.; Duarte, J.A.; Domingues, P.M.; Amado, F.M. Peptidomic analysis of human acquired enamel pellicle. *Biomed. Chromatogr.* **2007**, *21*, 1107–1117. [CrossRef] [PubMed]

32. Palmerini, C.A.; Mazzoni, M.; Radicioni, G.; Marzano, V.; Granieri, L.; Iavarone, F.; Longhi, R.; Messana, I.; Cabras, T.; Sanna, M.T.; et al. Antagonistic Effect of a Salivary Proline-Rich Peptide on the Cytosolic Ca^{2+} Mobilization Induced by Progesterone in Oral Squamous Cancer Cells. *PLoS ONE* **2016**, *11*, e0147925. [CrossRef] [PubMed]

33. Macias, M.J.; Wiesner, S.; Sudol, M. WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett.* **2002**, *513*, 30–37. [CrossRef] [PubMed]

34. Righino, B.; Pirolli, D.; Radicioni, G.; Marzano, V.; Longhi, R.; Arcovito, A.; Sanna, M.T.; De Rosa, M.C.; Paoluzi, S.; Cesareni, G.; et al. Structural studies and SH3 domain binding properties of a human antiviral salivary proline-rich peptide. *Biopolymers* **2016**, *106*, 714–725. [CrossRef]

35. Azen, E.A. Genetics of salivary protein polymorphisms. *Crit. Rev. Oral Biol. Med.* **1993**, *4*, 479–485. [CrossRef]

36. Sabatini, L.M.; Carlock, L.R.; Johnson, G.W.; Azen, E.A. cDNA cloning and chromosomal localization (4q11-13) of a gene for statherin, a regulator of calcium in saliva. *Am. J. Hum. Genet.* **1987**, *41*, 1048–1060.

37. 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.* **1992**, *267*, 5968–5976. [CrossRef]

38. Niemi, L.D.; Johansson, I. Salivary statherin peptide-binding epitopes of commensal and potentially infectious Actinomyces spp. delineated by a hybrid peptide construct. *Infect. Immun.* **2004**, *72*, 782–787. [CrossRef]

39. Gibbons, R.J.; Hay, D.I. Human salivary acidic proline-rich proteins and statherin promote the attachment of Actinomyces viscosus LY7 to apatitic surfaces. *Infect. Immun.* **1988**, *56*, 439–445. [CrossRef]

40. Isemura, S. Nucleotide sequence of gene PBII encoding salivary proline-rich protein P-B. *J. Biochem.* **2000**, *127*, 393–398. [CrossRef]

41. Inzitari, R.; Cabras, T.; Rossetti, D.V.; Fanali, C.; Vitali, A.; Pellegrini, M.; Paludetti, G.; Manni, A.; Giardina, B.; Messana, I.; et al. Detection in human saliva of different statherin and P-B fragments and derivatives. *Proteomics* **2006**, *6*, 6370–6379. [CrossRef]

42. Strawich, E.; Glimcher, M.J. Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albumin. *Eur. J. Biochem.* **1990**, *191*, 47–56. [CrossRef]

43. Ayad, M.; Van Wuyckhuyse, B.C.; Minaguchi, K.; Raubertas, R.F.; Bedi, G.S.; Billings, R.J.; Bowen, W.H.; Tabak, L.A. The association of basic proline-rich peptides from human parotid gland secretions with caries experience. *J. Dent. Res.* **2000**, *79*, 976–982. [CrossRef] [PubMed]

44. Jensen, J.L.; Lamkin, M.S.; Troxler, R.F.; Oppenheim, F.G. Multiple forms of statherin in human salivary secretions. *Arch. Oral Biol.* **1991**, *36*, 529–534. [CrossRef] [PubMed]

45. Sabatini, L.M.; He, Y.Z.; Azen, E.A. Structure and sequence determination of the gene encoding human salivary statherin. *Gene* **1990**, *89*, 245–251. [CrossRef]

46. Vitorino, R.; Lobo, M.J.; Duarte, J.R.; Ferrer-Correia, A.J.; Domingues, P.M.; Amado, F.M. The role of salivary peptides in dental caries. *Biomed. Chromatogr.* **2005**, *19*, 214–222. [CrossRef] [PubMed]

47. Sabatini, L.M.; Azen, E.A. Histatins, a family of salivary histidine-rich proteins, are encoded by at least two loci (HIS1 and HIS2). *Biochem. Biophys. Res. Commun.* **1989**, *160*, 495–502. [CrossRef] [PubMed]

48. Oppenheim, F.G.; Xu, T.; McMillian, F.M.; Levitz, S.M.; Diamond, R.D.; Offner, G.D.; Troxler, R.F. 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.* **1988**, *263*, 7472–7477. [CrossRef]

49. Castagnola, M.; Inzitari, R.; Rossetti, D.V.; Olmi, C.; Cabras, T.; Piras, V.; Nicolussi, P.; Sanna, M.T.; Pellegrini, M.; Giardina, B.; et al. A cascade of 24 histatins (histatin 3 fragments) in human saliva. Suggestions for a pre-secretory sequential cleavage pathway. *J. Biol. Chem.* **2004**, *279*, 41436–41443. [CrossRef]

50. Torres, P.; Díaz, J.; Arce, M.; Silva, P.; Mendoza, P.; Lois, P.; Molina-Berríos, A.; Owen, G.I.; Palma, V.; Torres, V.A. The salivary peptide histatin-1 promotes endothelial cell adhesion, migration, and angiogenesis. *FASEB J.* **2017**, *31*, 4946–4958. [CrossRef]

51. Oudhoff, M.J.; Bolscher, J.G.; Nazmi, K.; Kalay, H.; van't Hof, W.; Amerongen, A.V.; Veerman, E.C. Histatins are the major wound-closure stimulating factors in human saliva as identified in a cell culture assay. *FASEB J.* **2008**, *22*, 3805–3812. [CrossRef] [PubMed]

52. Amado, F.; Lobo, M.J.; Domingues, P.; Duarte, J.A.; Vitorino, R. Salivary peptidomics. *Expert Rev. Proteom.* **2010**, *7*, 709–721. [CrossRef]

53. Castagnola, M.; Scarano, E.; Passali, G.C.; Messana, I.; Cabras, T.; Iavarone, F.; Di Cintio, G.; Fiorita, A.; De Corso, E.; Paludetti, G. Salivary biomarkers and proteomics: Future diagnostic and clinical utilities. *Acta Otorhinolaryngol. Ital.* **2017**, *37*, 94–101. [CrossRef] [PubMed]

54. Kos, J.; Lah, T.T. Cysteine proteases and their endogenous inhibitors: Target proteins for prognosis, diagnosis and therapy in cancer (review). *Oncol. Rep.* **1998**, *5*, 1349–13461. [CrossRef] [PubMed]

55. Henskens, Y.M.; Veerman, E.C.; Nieuw Amerongen, A.V. Cystatins in health and disease. *Biol. Chem. Hoppe Seyler* **1996**, *377*, 71–86. [CrossRef] [PubMed]

56. Calkins, C.C.; Sloane, B.F. Mammalian cysteine protease inhibitors: Biochemical properties and possible roles in tumor progression. *Biol. Chem. Hoppe Seyler* **1995**, *376*, 71–80. [PubMed]

57. Manconi, B.; Liori, B.; Cabras, T.; Vincenzoni, F.; Iavarone, F.; Castagnola, M.; Messana, I.; Olianas, A. Salivary cystatins: Exploring new post-translational modifications and polymorphisms by top-down mass spectrometry. *J. Proteome Res.* **2017**, *16*, 4196–4207. [CrossRef]

58. Lupi, A.; Messana, I.; Denotti, G.; Schininà, M.E.; Gambarini, G.; Fadda, M.B.; Vitali, A.; Cabras, T.; Piras, V.; Patamia, M.; et al. Identification of the human salivary cystatin complex by the coupling of high-performance liquid chromatography and ion-trap mass spectrometry. *Proteomics* **2003**, *3*, 461–467. [CrossRef]

59. Ryan, C.M.; Souda, P.; Halgand, F.; Wong, D.T.; Loo, J.A.; Faull, K.F.; Whitelegge, J.P. Confident assignment of intact mass tags to human salivary cystatins using top-down Fourier-transform ion cyclotron resonance mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 908–917. [CrossRef]

60. Iavarone, F.; Cabras, T.; Pisano, E.; Sanna, M.T.; Nemolato, S.; Vento, G.; Tirone, C.; Romagnoli, C.; Cordaro, M.; Fanos, V.; et al. Top-down HPLC-ESI-MS detection of S-glutathionylated and S-cysteinylated derivatives of cystatin B and its 1-53 and 54-98 fragments in whole saliva of human preterm newborns. *J. Proteome Res.* **2013**, *12*, 917–926. [CrossRef]

61. Requena, T.; Velasco, M. The human microbiome in sickness and in health. *Rev. Clin. Esp.* **2021**, *221*, 233–240. [CrossRef] [PubMed]

62. Chimenos-Küstner, E.; Giovannoni, M.L.; Schemel-Suárez, M. Dysbiosis as a determinant factor of systemic and oral pathology: Importance of microbiome. *Med. Clin.* **2017**, *149*, 305–309. [CrossRef]

63. Helmerhorst, E.J.; Sun, X.; Salih, E.; Oppenheim, F.G. Identification of Lys-Pro-Gln as a novel cleavage site specificity of saliva-associated proteases. *J. Biol. Chem.* **2008**, *283*, 19957–19966. [CrossRef] [PubMed]

64. Zamakhchari, M.; Wei, G.; Dewhirst, F.; Lee, J.; Schuppan, D.; Oppenheim, F.G.; Helmerhorst, E.J. Identification of Rothia bacteria as gluten-degrading natural colonizers of the upper gastro-intestinal tract. *PLoS ONE* **2011**, *6*, e24455. [CrossRef]

65. Asano, M.; Komiyama, K. Polymeric immunoglobulin receptor. *J. Oral Sci.* **2011**, *53*, 147–156. [CrossRef]

66. Kaetzel, C.S. The polymeric immunoglobulin receptor: Bridging innate and adaptive immune responses at mucosal surfaces. *Immunol. Rev.* **2005**, *206*, 83–99. [CrossRef] [PubMed]

67. Ramachandran, P.; Boontheung, P.; Xie, Y.; Sondej, M.; Wong, D.T.; Loo, J.A. Identification of N-linked glycoproteins in human saliva by glycoprotein capture and mass spectrometry. *J. Proteome Res.* **2006**, *5*, 1493–1503. [CrossRef]

68. Cabras, T.; Sanna, M.; Manconi, B.; Fanni, D.; Demelia, L.; Sorbello, O.; Iavarone, F.; Castagnola, M.; Faa, G.; Messana, I. Proteomic investigation of whole saliva in Wilson's disease. *J. Proteom.* **2015**, *128*, 154–163. [CrossRef] [PubMed]

69. Boroumand, M.; Iavarone, F.; Manconi, B.; Pieroni, L.; Greco, V.; Vento, G.; Tirone, C.; Desiderio, C.; Fiorita, A.; Faa, G.; et al. HPLC-ESI-MS top-down analysis of salivary peptides of preterm newborns evidenced high activity of some exopeptidases and convertases during late fetal development. *Talanta* **2021**, *222*, 121429. [CrossRef]

70. Cabras, T.; Pisano, E.; Mastinu, A.; Denotti, G.; Pusceddu, P.P.; Inzitari, R.; Fanali, C.; Nemolato, S.; Castagnola, M.; Messana, I. Alterations of the salivary secretory peptidome profile in children affected by type 1 diabetes. *Mol. Cell. Proteom.* **2010**, *9*, 2099–2108. [CrossRef]

71. Hardt, M.; Thomas, L.R.; Dixon, S.E.; Newport, G.; Agabian, N.; Prakobphol, A.; Hall, S.C.; Witkowska, H.E.; Fisher, S.J. Toward defining the human parotid gland salivary proteome and peptidome: Identification and characterization using 2D SDS-PAGE, ultrafiltration, HPLC, and mass spectrometry. *Biochemistry* **2005**, *44*, 2885–2899. [CrossRef] [PubMed]

72. Huq, N.L.; Cross, K.J.; Ung, M.; Myroforidis, H.; Veith, P.D.; Chen, D.; Stanton, D.; Huiling, H.; Ward, B.R.; Reynolds, E.C. A Review of the Salivary Proteome and Peptidome and Saliva-derived Peptide Therapeutics. *Int. J. Peptide Res. Ther.* **2007**, *13*, 547–564. [CrossRef]

73. Manconi, B.; Liori, B.; Cabras, T.; Vincenzoni, F.; Iavarone, F.; Lorefice, L.; Cocco, E.; Castagnola, M.; Messana, I.; Olianas, A. Top-down proteomic profiling of human saliva in multiple sclerosis patients. *J. Proteom.* **2018**, *187*, 212–222. [CrossRef] [PubMed]

74. Serrao, S.; Firinu, D.; Olianas, A.; Deidda, M.; Contini, C.; Iavarone, F.; Sanna, M.T.; Boroumand, M.; Amado, F.; Castagnola, M. Top-Down Proteomics of Human Saliva Discloses Significant Variations of the Protein Profile in Patients with Mastocytosis. *J. Proteome Res.* **2020**, *19*, 3238–3253. [CrossRef] [PubMed]

75. Contini, C.; Serrao, S.; Manconi, B.; Olianas, A.; Iavarone, F.; Bizzarro, A.; Masullo, C.; Castagnola, M.; Messana, I.; Diaz, G.; et al. Salivary Proteomics Reveals Significant Changes in Relation to Alzheimer's Disease and Aging. *J. Alzheimers Dis.* **2022**, *89*, 605–622. [CrossRef]

76. Olianas, A.; Guadalupi, G.; Cabras, T.; Contini, C.; Serrao, S.; Iavarone, F.; Castagnola, M.; Messana, I.; Onali, S.; Chessa, L.; et al. Top-Down Proteomics Detection of Potential Salivary Biomarkers for Autoimmune Liver Diseases Classification. *Int. J. Mol. Sci.* **2023**, *24*, 959. [CrossRef]

77. Contini, C.; Olianas, A.; Serrao, S.; Deriu, C.; Iavarone, F.; Boroumand, M.; Bizzarro, A.; Lauria, A.; Faa, G.; Castagnola, M.; et al. Top-Down Proteomics of Human Saliva Highlights Anti-inflammatory, Antioxidant, and Antimicrobial Defense Responses in Alzheimer Disease. *Front. Neurosci.* **2021**, *15*, 668852. [CrossRef]

78. Wilson, L.J.; Linley, A.; Hammond, D.E.; Hood, F.E.; Coulson, J.M.; MacEwan, D.J.; Ross, S.J.; Slupsky, J.R.; Smith, P.D.; Eyers, P.A.; et al. New Perspectives, Opportunities, and Challenges in Exploring the Human Protein Kinome. *Cancer Res.* **2018**, *78*, 15–29. [CrossRef]

79. Salih, E.; Siqueira, W.L.; Helmerhorst, E.J.; Oppenheim, F.G. Large-scale phosphoproteome of human whole saliva using disulfidethiol interchange covalent chromatography and mass spectrometry. *Anal. Biochem.* **2010**, *407*, 19–33. [CrossRef]

80. Stone, M.D.; Chen, X.; McGowan, T.; Bandhakavi, S.; Cheng, B.; Rhodus, N.L.; Griffin, T.J. Large-scale phosphoproteomics analysis of whole saliva reveals a distinct phosphorylation pattern. *J. Proteome Res.* **2011**, *10*, 1728–1736. [CrossRef]

81. Tagliabracci, V.S.; Pinna, L.A.; Dixon, J.E. Secreted protein kinases. *Trends Biochem. Sci.* **2013**, *38*, 121–130. [CrossRef]

82. Tagliabracci, V.S.; Wiley, S.E.; Guo, X.; Kinch, L.N.; Durrant, E.; Wen, J.; Xiao, J.; Cui, J.; Nguyen, K.B.; Engel, J.L.; et al. A Single Kinase Generates the Majority of the Secreted Phosphoproteome. *Cell* **2015**, *161*, 1619–1632. [CrossRef]

83. Isemura, S.; Saitoh, E.; Sanada, K.; Minakata, K. Identification of full-sized forms of salivary (S-type) cystatins (cystatin SN, cystatin SA, cystatin S, and two phosphorylated forms of cystatin S) in human whole saliva and determination of phosphorylation sites of cystatin S. *J. Biochem.* **1991**, *110*, 648–654. [CrossRef] [PubMed]

84. Madapallimattam, G.; Bennick, A. Phosphorylation of salivary proteins by salivary gland protein kinase. *J. Dent. Res.* **1986**, *65*, 405–411. [CrossRef] [PubMed]

85. Iavarone, F.; D'Alessandro, A.; Tian, N.; Cabras, T.; Messana, I.; Helmerhorst, E.J.; Oppenheim, F.G.; Castagnola, M. High-resolution high-performance liquid chromatography with electrospray ionization mass spectrometry and tandem mass spectrometry characterization of a new isoform of human salivary acidic proline-rich proteins named Roma-Boston Ser22 (Phos) Phe variant. *J. Sep. Sci.* **2014**, *37*, 1896–1902. [CrossRef]

86. Halgand, F.; Zabrouskov, V.; Bassilian, S.; Souda, P.; Loo, J.A.; Faull, K.F.; Wong, D.T.; Whitelegge, J.P. Defining intact protein primary structures from saliva: A step toward the human proteome project. *Anal. Chem.* **2012**, *84*, 4383–4395. [CrossRef]

87. Castagnola, M.; Inzitari, R.; Fanali, C.; Iavarone, F.; Vitali, A.; Desiderio, C.; Vento, G.; Tirone, C.; Romagnoli, C.; Cabras, T.; et al. The surprising composition of the salivary proteome of preterm human newborn. *Mol. Cell Proteom.* **2011**, *10*, M110.003467. [CrossRef] [PubMed]

88. Wahid, A.; Sohail, A.; Wang, H.; Guo, M.; Zhang, L.; Ji, Y.; Wang, P.; Xiao, H. Titanium(IV) immobilized affinity chromatography facilitated phosphoproteomics analysis of salivary extracellular vesicles for lung cancer. *Anal. Bioanal. Chem.* **2022**, *414*, 3697–3708. [CrossRef]

89. Sun, J.; Wang, X.; Ding, Y.; Xiao, B.; Wang, X.; Ali, M.M.; Ma, L.; Xie, Z.; Gu, Z.; Chen, G.; et al. Proteomic and phosphoproteomic landscape of salivary extracellular vesicles to assess OSCC therapeutical outcomes. *Proteomics* **2023**, *23*, e2200319. [CrossRef]

90. Inzitari, R.; Vento, G.; Capoluongo, E.; Boccacci, S.; Fanali, C.; Cabras, T.; Romagnoli, C.; Giardina, B.; Messana, I.; Castagnola, M. Proteomic analysis of salivary acidic proline-rich proteins in human pre-term and at-term new-borns. *J. Proteome Res.* **2007**, *6*, 1371–1377. [CrossRef]

91. Scarano, E.; Fiorita, A.; Picciotti, P.M.; Passali, G.C.; Calò, L.; Cabras, T.; Inzitari, R.; Fanali, C.; Messana, I.; Castagnola, M.; et al. Proteomics of saliva: Personal experience. *Acta Otorhinolaryngol. Ital.* **2010**, *30*, 125–130.

92. Castagnola, M.; Messana, I.; Inzitari, R.; Fanali, C.; Cabras, T.; Morelli, A.; Pecoraro, A.M.; Neri, G.; Torrioli, M.G.; Gurrieri, F. Hypo-phosphorylation of salivary peptidome as a clue to the molecular pathogenesis of autism spectrum disorders. *J. Proteome Res.* **2008**, *7*, 5327–5332. [CrossRef]

93. Nimmagadda, D.; Cherala, G.; Ghatta, S. Cytosolic sulfotransferases. *Indian J. Exp. Biol.* **2006**, *44*, 171–182.

94. Leyh, T.S. The physical biochemistry and molecular genetics of sulfate activation. *Crit. Rev. Biochem. Mol. Biol.* **1993**, *28*, 515–542. [CrossRef] [PubMed]

95. Venkatachalam, K.V. Human 3'-phosphoadenosine 5'-phosphosulfate (PAPS) synthase: Biochemistry, molecular biology and genetic deficiency. *IUBMB Life* **2003**, *55*, 1–11. [CrossRef] [PubMed]

96. Schröder, E.; Gebel, L.; Eremeev, A.A.; Morgner, J.; Grum, D.; Knauer, S.K.; Bayer, P.; Mueller, J.W. Human PAPS synthase isoforms are dynamically regulated enzymes with access to nucleus and cytoplasm. *PLoS ONE* **2012**, *7*, e29559. [CrossRef]

97. Gamage, N.; Barnett, A.; Hempel, N.; Duggleby, R.G.; Windmill, K.F.; Martin, J.L.; McManus, M.E. Human sulfotransferases and their role in chemical metabolism. *Toxicol. Sci.* **2006**, *90*, 5–22. [CrossRef]

98. Negishi, M.; Pedersen, L.G.; Petrotchenko, E.; Shevtsov, S.; Gorokhov, A.; Kakuta, Y.; Pedersen, L.C. Structure and function of sulfotransferases. *Arch. Biochem. Biophys.* **2001**, *390*, 149–157. [CrossRef]

99. Falany, C.N. Enzymology of human cytosolic sulfotransferases. *FASEB J.* **1997**, *11*, 206–216. [CrossRef] [PubMed]

100. Pedersen, L.C.; Yi, M.; Pedersen, L.G.; Kaminski, A.M. From steroid and drug metabolism to glycobiology, using sulfotransferase structures to understand and tailor function. *Drug Metab. Dispos.* **2022**, *50*, 1027–1041. [CrossRef] [PubMed]

101. Kurogi, K.; Rasool, M.I.; Alherz, F.A.; El Daibani, A.A.; Bairam, A.F.; Abunnaja, M.S.; Yasuda, S.; Wilson, L.J.; Hui, Y.; Liu, M.C. SULT genetic polymorphisms: Physiological, pharmacological and clinical implications. *Expert Opin. Drug Metab. Toxicol.* **2021**, *17*, 767–784. [CrossRef] [PubMed]

102. Riches, Z.; Stanley, E.L.; Bloomer, J.C.; Coughtrie, M.W. Quantitative evaluation of the expression and activity of five major sulfotransferases (SULTs) in human tissues: The SULT "pie". *Drug Metab. Dispos.* **2009**, *37*, 2255–2261. [CrossRef]

103. Cabras, T.; Fanali, C.; Monteiro, J.A.; Amado, F.; Inzitari, R.; Desiderio, C.; Giardina, B.; Castagnola, M.; Messana, I. Tyrosine polysulfation of human salivary histatin 1. A post-translational modification specific of the submandibular gland. *J. Proteome Res.* **2007**, *6*, 2472–2480. [CrossRef] [PubMed]

104. Poole, L.B. The basics of thiols and cysteines in redox biology and chemistry. *Free Radic. Biol. Med.* **2015**, *80*, 148–157. [CrossRef]

105. Winterbourn, C.C.; Hampton, M.B. Thiol chemistry and specificity in redox signaling. *Free Radic. Biol. Med.* **2008**, *45*, 549–561. [CrossRef] [PubMed]

106. Cabras, T.; Manconi, B.; Iavarone, F.; Fanali, C.; Nemolato, S.; Fiorita, A.; Scarano, E.; Passali, G.C.; Manni, A.; Cordaro, M.; et al. RP-HPLC-ESI-MS evidenced that salivary cystatin B is detectable in adult human saliva mostly as S-modified derivatives: S-glutathionyl, S-cysteinyl and S-S 2-mer. *J. Proteom.* **2012**, *75*, 908–913. [CrossRef]

107. Pol, E.; Björk, I. Contributions of individual residues in the N-terminal region of cystatin B (stefin B) to inhibition of cysteine proteinases. *Biochim. Biophys. Acta* **2003**, *1645*, 105–112. [CrossRef]

108. Grek, C.L.; Zhang, J.; Manevich, Y.; Townsend, D.M.; Tew, K.D. Causes and consequences of cysteine S-glutathionylation. *J. Biol. Chem.* **2013**, *288*, 26497–26504. [CrossRef]

109. Donato, R.; Cannon, B.R.; Sorci, G.; Riuzzi, F.; Hsu, K.; Weber, D.J.; Geczy, C.L. Functions of S100 proteins. *Curr. Mol. Med.* **2013**, *13*, 24–57. [CrossRef]

110. Leclerc, E.; Fritz, G.; Vetter, S.W.; Heizmann, C.W. Binding of S100 proteins to RAGE: An update. *Biochim. Biophys. Acta* **2009**, *1793*, 993–1007. [CrossRef]

111. Zackular, J.P.; Chazin, W.J.; Skaar, E.P. Nutritional Immunity: S100 Proteins at the Host-Pathogen Interface. *J. Biol. Chem.* **2015**, *290*, 18991–18998. [CrossRef] [PubMed]

112. 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.* **2009**, *86*, 577–587. [CrossRef] [PubMed]

113. Lim, S.Y.; Raftery, M.J.; Goyette, J.; Geczy, C.L. S-glutathionylation regulates inflammatory activities of S100A9. *J. Biol. Chem.* **2010**, *285*, 14377–14388. [CrossRef]

114. Boroumand, M.; Manconi, B.; Serrao, S.; Iavarone, F.; Olianas, A.; Cabras, T.; Contini, C.; Pieroni, L.; Sanna, M.T.; Vento, G.; et al. Investigation by top-down high-performance liquid chromatography–mass spectrometry of glutathionylation and cysteinylation of salivary S100A9 and cystatin B in preterm newborns. *Sep. Sci. Plus* **2022**, *5*, 17–27. [CrossRef]

115. Iavarone, F.; Melis, M.; Platania, G.; Cabras, T.; Manconi, B.; Petruzzelli, R.; Cordaro, M.; Siracusano, A.; Faa, G.; Messana, I.; et al. Characterization of salivary proteins of schizophrenic and bipolar disorder patients by top-down proteomics. *J. Proteom.* **2014**, *103*, 15–22. [CrossRef] [PubMed]

116. Contini, C.; Firinu, D.; Serrao, S.; Manconi, B.; Olianas, A.; Cinetto, F.; Cossu, F.; Castagnola, M.; Messana, I.; Del Giacco, S.; et al. RP-HPLC-ESI-IT Mass Spectrometry Reveals Significant Variations of the Human Salivary Protein Profile Associated with Predominantly Antibody Deficiencies. *J. Clin. Immunol.* **2020**, *40*, 329–339. [CrossRef]

117. Hoskin, T.S.; Crowther, J.M.; Cheung, J.; Epton, M.J.; Sly, P.D.; Elder, P.A.; Dobson, R.C.J.; Kettle, A.J.; Dickerhof, N. Oxidative cross-linking of calprotectin occurs in vivo, altering its structure and susceptibility to proteolysis. *Redox Biol.* **2019**, *24*, 101202. [CrossRef]

118. Gomes, L.H.; Raftery, M.J.; Yan, W.X.; Goyette, J.D.; Thomas, P.S.; Geczy, C.L. S100A8 and S100A9-oxidant scavengers in inflammation. *Free Radic. Biol. Med.* **2013**, *58*, 170–186. [CrossRef]

119. Smith, T.S.; Bennett, J.P., Jr. Mitochondrial toxins in models of neurodegenerative diseases. I: In vivo brain hydroxyl radical production during systemic MPTP treatment or following microdialysis infusion of methylpyridinium or azide ions. *Brain Res.* **1997**, *765*, 183–188. [CrossRef]

120. Karlík, M.; Valkovič, P.; Hančinová, V.; Krížová, L.; Tóthová, L.; Celec, P. Markers of oxidative stress in plasma and saliva in patients with multiple sclerosis. *Clin. Biochem.* **2015**, *48*, 24–28. [CrossRef]

121. Schipper, H.M.; Arnold, D.; Grand'Maison, F.; Melmed, C.; Moore, F.; Levental, M.; Su, H.; Constantin, M.; Stril, J.L.; Godin, J. Tolerability and safety of combined glatiramer acetate and N-acetylcysteine in relapsing-remitting multiple sclerosis. *Clin. Neuropharmacol.* **2015**, *38*, 127–131. [CrossRef] [PubMed]

122. Yao, Y.; Berg, E.A.; Costello, C.E.; Troxler, R.F.; Oppenheim, F.G. Identification of protein components in human acquired enamel pellicle and whole saliva using novel proteomics approaches. *J. Biol. Chem.* **2003**, *278*, 5300–5308. [CrossRef] [PubMed]

123. Schüpbach, P.; Oppenheim, F.G.; Lendenmann, U.; Lamkin, M.S.; Yao, Y.; Guggenheim, B. Electron-microscopic demonstration of proline-rich proteins, statherin, and histatins in acquired enamel pellicles in vitro. *Eur. J. Oral Sci.* **2001**, *109*, 60–68. [CrossRef] [PubMed]

124. Hannig, C.; Hannig, M.; Kensche, A.; Carpenter, G. The mucosal pellicle—An underestimated factor in oral physiology. *Arch. Oral Biol.* **2017**, *80*, 144–152. [CrossRef] [PubMed]

125. Presland, R.B.; Dale, B.A. Epithelial structural proteins of the skin and oral cavity: Function in health and disease. *Crit. Rev. Oral Biol. Med.* **2000**, *11*, 383–408. [CrossRef]

126. 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.* **1989**, *261*, 887–896. [CrossRef]

127. Cabiddu, G.; Maes, P.; Hyvrier, F.; Olianas, A.; Manconi, B.; Brignot, H.; Canon, F.; Cabras, T.; Morzel, M. Proteomic characterization of the mucosal pellicle formed in vitro on a cellular model of oral epithelium. *J. Proteom.* **2020**, *222*, 103797. [CrossRef]

128. Neyraud, E.; Morzel, M. Biological films adhering to the oral soft tissues: Structure, composition, and potential impact on taste perception. *J. Texture Stud.* **2019**, *50*, 19–26. [CrossRef]

129. Perez Alea, M.; Thomas, V.; Martin, G.; El Alaoui, S. Identification of human salivary transglutaminases. *Amino Acids* **2013**, *44*, 245–250. [CrossRef]

130. Wang, Z.; Griffin, M. TG2, a novel extracellular protein with multiple functions. *Amino Acids* **2012**, *42*, 939–949. [CrossRef]

131. Boroumand, M.; Olianas, A.; Manconi, B.; Serrao, S.; Iavarone, F.; Desiderio, C.; Pieroni, L.; Faa, G.; Messana, I.; Castagnola, M.; et al. Mapping of transglutaminase 2 sites of human salivary small basic proline-rich proteins by HPLC high resolution ESI-MS/MS. *J. Proteome Res.* **2020**, *19*, 300–313. [CrossRef] [PubMed]

132. Cabras, T.; Inzitari, R.; Fanali, C.; Scarano, E.; Patamia, M.; Sanna, M.T.; Pisano, E.; Giardina, B.; Castagnola, M.; Messana, I. HPLC-MS characterization of cyclo-statherin Q-37, a specific cyclization product of human salivary statherin generated by transglutaminase 2. *J. Sep. Sci.* **2006**, *29*, 2600–2608. [CrossRef] [PubMed]

133. Hannig, C.; Hannig, M.; Attin, T. Enzymes in the acquired enamel pellicle. *Eur. J. Oral Sci.* **2005**, *113*, 2–13. [CrossRef] [PubMed]

134. Siqueira, W.L.; Custodio, W.; McDonald, E.E. New insights into the composition and functions of the acquired enamel pellicle. *J. Dent. Res.* **2012**, *91*, 1110–1118. [CrossRef] [PubMed]

135. Zhang, Q.; Ames, J.M.; Smith, R.D.; Baynes, J.W.; Metz, T.O. A Perspective on the Maillard Reaction and the Analysis of protein Glycation by Mass Spectrometry: Probing the Pathogenesis of Chronic Disease. *J. Proteome Res.* **2009**, *8*, 754–769. [CrossRef]

136. Rahbar, S.; Blumenfe, O.; Ranney, H.M. Studies of an unusual hemoglobin in patients with diabetes mellitus. *Biochem. Biophys. Res. Commun.* **1969**, *36*, 838–843. [CrossRef]

137. Nakamoto, I.; Morimoto, K.; Takeshita, T.; Toda, M. Correlation between Saliva Glycated and Blood Glycated Proteins. *Environ. Health Prev. Med.* **2003**, *8*, 95–99. [CrossRef]

138. Goldstein, D.E.; Little, R.R.; Lorenz, R.A.; Malone, J.I.; Nathan, D.; Peterson, C.M. Tests of glycemia in diabetes. *Diabetes Care* **1995**, *18*, 896–909. [CrossRef]

139. Ramasamy, R.; Vannucci, S.J.; Yan, S.S.; Herold, K.; Yan, S.F.; Schmidt, A.M. Advanced glycation end products and RAGE: A common thread in aging, diabetes, neurodegeneration, and inflammation. *Glycobiology* **2005**, *15*, 16R–28R. [CrossRef]

140. Corica, D.; Pepe, G.; Currò, M.; Aversa, T.; Tropeano, A.; Ientile, R.; Wasniewska, M. Methods to investigate advanced glycation end-product and their application in clinical practice. *Methods* **2022**, *203*, 90–102. [CrossRef]

141. McCarthy, A.D.; Cortizo, A.M.; Giménez Segura, G.; Bruzzone, L.; Etcheverry, S.B. Non-enzymatic glycosylation of alkaline phosphatase alters its biological properties. *Mol. Cell. Biochem.* **1998**, *181*, 63–69. [CrossRef] [PubMed]

142. Muraoka, M.Y.; Justino, A.B.; Caixeta, D.C.; Queiroz, J.S.; Sabino-Silva, R.; Espindola, F.S. Fructose and methylglyoxal-induced glycation alters structural and functional properties of salivary proteins, albumin and lysozyme. *PLoS ONE* **2022**, *17*, e0262369. [CrossRef] [PubMed]

143. Tessier, F.J. The Maillard reaction in the human body. The main discoveries and factors that affect glycation. *Pathol. Biol.* **2010**, *58*, 214–219. [CrossRef] [PubMed]

144. Shin, A.; Connolly, S.; Kabytaev, K. Protein glycation in diabetes mellitus. *Adv. Clin. Chem.* **2023**, *113*, 101–156. [CrossRef]

145. Li, Y.M. Glycation ligand binding motif in lactoferrin. Implication in diabetic infection. *Adv. Exp. Med. Biol.* **1998**, *443*, 57–63. [CrossRef]

146. Belce, A.; Uslu, E.; Kucur, M.; Umut, M.; Ipbuker, A.; Seymen, H.O. Evaluation of salivary sialic acid level and Cu-Zn superoxide dismutase activity in type 1 diabetes mellitus. *Tohoku J. Exp. Med.* **2000**, *192*, 219–225. [CrossRef]

147. Ilea, A.; Băbţan, A.M.; Boşca, B.A.; Crişan, M.; Petrescu, N.B.; Collino, M.; Sainz, R.M.; Gerlach, J.Q.; Câmpian, R.S. Advanced Glycation end products (AGEs) in oral pathology. *Arch. Oral Biol.* **2018**, *93*, 22–30. [CrossRef]

148. Manig, F.; Hellwig, M.; Pietz, F.; Henle, T. Quantitation of free glycation compounds in saliva. *PLoS ONE* **2019**, *14*, e0220208. [CrossRef]

149. Landesberg, R.; Woo, V.; Huang, L.; Cozin, M.; Lu, Y.; Bailey, C.; Qu, W.; Pulse, C.; Schmidt, A.M. The expression of the receptor for glycation endproducts (RAGE) in oral squamous cell carcinomas. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* **2008**, *105*, 617–624. [CrossRef]

150. Maciejczyk, M.; Nesterowicz, M.; Szulimowska, J.; Zalewska, A. Oxidation, Glycation, and Carbamylation of Salivary Biomolecules in Healthy Children, Adults, and Elderly: Can Saliva Be Used in the assessment of Aging? *J. Inflamm. Res.* **2022**, *15*, 2051–2073. [CrossRef]

151. Goyette, J.; Geczy, C.L. Inflammation-associated S100 proteins: New mechanisms that regulate function. *Amino Acids* **2011**, *41*, 821–842. [CrossRef] [PubMed]

152. Agho, E.T.; Owotade, F.J.; Kolawole, B.A.; Oyetola, E.O.; Adedeji, T.A. Salivary inflammatory biomarkers and glycated haemoglobin among patients with type 2 diabetic mellitus. *BMC Oral Health* **2021**, *21*, 101. [CrossRef] [PubMed]

153. Helenius, A.; Aebi, M. Intracellular functions of N-linked glycans. *Science* **2001**, *291*, 2364–2369. [CrossRef] [PubMed]

154. Xu, C.; Ng, D.T.W. Glycosylation-directed quality control of protein folding. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 742–752. [CrossRef] [PubMed]

155. Minakata, S.; Manabe, S.; Inai, Y.; Ikezaki, M.; Nishitsuji, K.; Ito, Y.; Ihara, Y. Protein *C*-Mannosylation and *C*-Mannosyl Tryptophan in Chemical Biology and Medicine. *Molecules* **2021**, *26*, 5258. [CrossRef]

156. Franc, V.; Yang, Y.; Heck, A.J. Proteoform profile mapping of the human serum complement component C9 revealing unexpected new features of N-, O-, and C-glycosylation. *Anal. Chem.* **2017**, *89*, 3483–3491. [CrossRef]

157. Siew, J.J.; Chern, Y. Microglial Lectins in Health and Neurological Diseases. *Front. Mol. Neurosci.* **2018**, *11*, 158. [CrossRef]

158. Duca, M.; Malagolini, N.; Dall'Olio, F. The Mutual Relationship between Glycosylation and Non-Coding RNAs in Cancer and Other Physio-Pathological Conditions. *Int. J. Mol. Sci.* **2022**, *23*, 15804. [CrossRef]

159. Borga, C.; Meeran, S.M.; Fassan, M. Non-coding RNAs, a real Next-Gen Class of Biomarkers? *Noncoding RNA Res.* **2019**, *4*, 80–81. [CrossRef]

160. Flynn, R.A.; Pedram, K.; Malaker, S.A.; Batista, P.J.; Smith, B.A.H.; Johnson, A.G.; George, B.M.; Majzoub, K.; Villalta, P.W.; Carette, J.E.; et al. Small RNAs are modified with N-glycans and displayed on the surface of living cells. *Cell* **2021**, *184*, 3109–3124. [CrossRef]

161. Morelle, W.; Canis, K.; Chirat, F.; Faid, V.; Michalski, J.C. The use of mass spectrometry for the proteomic analysis of glycosylation. *Proteomics* **2006**, *6*, 3993–4015. [CrossRef] [PubMed]

162. Sondej, M.; Denny, P.A.; Xie, Y.; Ramachandran, P.; Si, Y.; Takashima, J.; Shi, W.; Wong, D.T.; Loo, J.A.; Denny, P.C. Glycoprofiling of the Human Salivary Proteome. *Clin. Proteom.* **2009**, *5*, 52–68. [CrossRef] [PubMed]

163. Cabras, T.; Boi, R.; Pisano, E.; Iavarone, F.; Fanali, C.; Nemolato, S.; Faa, G.; Castagnola, M.; Messana, I. 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.* **2012**, *35*, 1079–1086. [CrossRef]

164. Manconi, B.; Cabras, T.; Sanna, M.; Piras, V.; Liori, B.; Pisano, E.; Iavarone, F.; Vincenzoni, F.; Cordaro, M.; Faa, G.; et al. N- and O-linked glycosylation site profiling of the human basic salivary proline-rich protein 3M. *J. Sep. Sci.* **2016**, *39*, 1987–1997. [CrossRef] [PubMed]

165. Liu, S.; Wang, H.; Jiang, X.; Ji, Y.; Wang, Z.; Zhang, Y.; Wang, P.; Xiao, H. Integrated *N*-glycoproteomics Analysis of Human Saliva for Lung Cancer. *J. Proteome Res.* **2022**, *21*, 1589–1602. [CrossRef]

166. Peng, W.; Kobeissy, F.; Mondello, S.; Barsa, C.; Mechref, Y. MS-based glycomics: An analytical tool to assess nervous system diseases. *Front. Neurosci.* **2022**, *16*, 1000179. [CrossRef]

167. Robertson, L.A.; Moya, K.L.; Breen, K.C. The potential role of tau protein O-glycosylation in Alzheimer's disease. *J. Alzheimers Dis.* **2004**, *6*, 489–495. [CrossRef]

168. Pinho, S.S.; Reis, C.A. Glycosylation in cancer: Mechanisms and clinical implications. *Nat. Rev. Cancer* **2015**, *15*, 540–555. [CrossRef]

169. György, B.; Tóth, E.; Tarcsa, E.; Falus, A.; Buzás, E.I. Citrullination: A posttranslational modification in health and disease. *Int. J. Biochem. Cell Biol.* **2006**, *38*, 1662–1677. [CrossRef]

170. Zhai, Q.; Wang, L.; Zhao, P.; Li, T. Role of citrullination modification catalyzed by peptidylarginine deiminase 4 in gene transcriptional regulation. *Acta Biochim. Biophys. Sin.* **2017**, *49*, 567–572. [CrossRef]

171. Alghamdi, M.; Alasmari, D.; Assiri, A.; Mattar, E.; Aljaddawi, A.A.; Alattas, S.G.; Redwan, E.M. An Overview of the Intrinsic Role of Citrullination in Autoimmune Disorders. *J. Immunol. Res.* **2019**, *2019*, 7592851. [CrossRef]

172. van Venrooij, W.J.; Pruijn, G.J. Citrullination: A small change for a protein with great consequences for rheumatoid arthritis. *Arthritis Res.* **2000**, *2*, 249–251. [CrossRef]

173. Ciesielski, O.; Biesiekierska, M.; Panthu, B.; Soszyński, M.; Pirola, L.; Balcerczyk, A. Citrullination in the pathology of inflammatory and autoimmune disorders: Recent advances and future perspectives. *Cell. Mol. Life Sci.* **2022**, *79*, 94. [CrossRef]

174. Bay-Jensen, A.C.; Sand, J.M.B.; Genovese, F.; Siebuhr, A.S.; Nielsen, M.J.; Leeming, D.J.; Manon-Jensen, T.; Karsdal, M.A. Chapter 31—Structural Biomarkers. In *Biochemistry of Collagens, Laminins and Elastin*, 1st ed.; Karsdal, M., Ed.; Elsevier: Amsterdam, The Netherlands, 2016; pp. 203–233.

175. Yasuda, T.; Tahara, K.; Sawada, T. Detection of salivary citrullinated cytokeratin 13 in healthy individuals and patients with rheumatoid arthritis by proteomics analysis. *PLoS ONE* **2022**, *17*, e0265687. [CrossRef] [PubMed]

176. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **2020**, *369*, 1318–1330. [CrossRef] [PubMed]

177. Anzilotti, C.; Pratesi, F.; Tommasi, C.; Migliorini, P. Peptidylarginine deiminase 4 and citrullination in health and disease. *Autoimmun. Rev.* **2010**, *9*, 158–160. [CrossRef]

178. Tar, I.; Csősz, É.; Végh, E.; Lundberg, K.; Kharlamova, N.; Soós, B.; Szekanecz, Z.; Márton, I. Salivary citrullinated proteins in rheumatoid arthritis and associated periodontal disease. *Sci. Rep.* **2021**, *11*, 13525. [CrossRef] [PubMed]

179. Sakaguchi, W.; To, M.; Yamamoto, Y.; Inaba, K.; Yakeishi, M.; Saruta, J.; Fuchida, S.; Hamada, N.; Tsukinoki, K. Detection of anti-citrullinated protein antibody (ACPA) in saliva for rheumatoid arthritis using DBA mice infected with Porphyromonas gingivalis. *Arch. Oral. Biol.* **2019**, *108*, 104510. [CrossRef]

180. Herrera-Esparza, R.; Rodríguez-Rodríguez, M.; Pérez-Pérez, M.E.; Badillo-Soto, M.A.; Torres-Del-Muro, F.; Bollain-Y-Goytia, J.J.; Pacheco-Tovar, D.; Avalos-Díaz, E. Posttranslational Protein Modification in the Salivary Glands of Sjögren's Syndrome Patients. *Autoimmune Dis.* **2013**, *2013*, 548064. [CrossRef]

181. Giglione, C.; Fieulaine, S.; Meinnel, T. N-terminal protein modifications: Bringing back into play the ribosome. *Biochimie* **2015**, *114*, 134–146. [CrossRef]

182. Deng, S.; Marmorstein, R. Protein N-Terminal Acetylation: Structural Basis, Mechanism, Versatility, and Regulation. *Trends Biochem. Sci.* **2021**, *46*, 15–27. [CrossRef] [PubMed]

183. Vijayan, D.K.; Zhang, K.Y.J. Human glutaminyl cyclase: Structure, function, inhibitors and involvement in Alzheimer's disease. *Pharmacol. Res.* **2019**, *147*, 104342. [CrossRef] [PubMed]

184. Castagnola, M.; Cabras, T.; Iavarone, F.; Vincenzoni, F.; Vitali, A.; Pisano, E.; Nemolato, S.; Scarano, E.; Fiorita, A.; Vento, G.; et al. Top-down platform for deciphering the human salivary proteome. *J. Matern. Fetal Neonatal Med.* **2012**, *25*, 27–43. [CrossRef] [PubMed]

185. Manconi, B.; Cabras, T.; Pisano, E.; Nemolato, S.; Inzitari, R.; Iavarone, F.; Fanali, C.; Sanna, M.T.; Tirone, C.; Vento, G.; et al. Characterization of two isoforms of SPRR3 from preterm human newborn saliva, autoptic fetal oral mucosa, parotid and submandibular gland samples. *Biochem. Biophys. Res. Commun.* **2010**, *398*, 477–481. [CrossRef]

186. Cabras, T.; Iavarone, F.; Manconi, B.; Olianas, A.; Sanna, M.T.; Castagnola, M.; Messana, I. Top-down analytical platforms for the characterization of the human salivary proteome. *Bioanalysis* **2014**, *6*, 563–581. [CrossRef]

187. Van Damme, P.; Hole, K.; Pimenta-Marques, A.; Helsens, K.; Vandekerckhove, J.; Martinho, R.G.; Gevaert, K.; Arnesen, T. NatF Contributes to an Evolutionary Shift in Protein N-Terminal acetylation and is important for normal chromosome segregation. *PLoS Genet.* **2011**, *7*, e1002169. [CrossRef]

188. Varland, S.; Osberg, C.; Arnesen, T. N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificity and biological effects. *Proteomics* **2015**, *15*, 2385–2401. [CrossRef]

189. Vitorino, R.; Alves, R.; Barro, A.; Caseiro, A.; Ferreira, R.; Lobo, M.C.; Bastos, A.; Duarte, J.; Carvalho, D.; Santos, L.L.; et al. Finding new translational modifications in salivary proline-rich proteins. *Proteomics* **2010**, *10*, 3732–3742. [CrossRef]

190. Schilling, S.; Wasternack, C.; Demuth, H.U. Glutaminyl cyclase from animals and plants: A case of functionally convergent protein evolution. *Biol. Chem.* **2008**, *389*, 983–991. [CrossRef]

191. Schlenzig, D.; Manhart, S.; Cinar, Y.; Kleinschmidt, M.; Hause, G.; Willbold, D.; Funke, S.A.; Schilling, S.; Demuth, H.U. Pyroglutamate formation influences solubility and amyloidogenicity of amyloid peptides. *Biochemistry* **2009**, *48*, 7072–7078. [CrossRef]

192. Wirths, O.; Breyhan, H.; Cynis, H.; Schilling, S.; Demuth, H.U.; Bayer, T.A. Intraneuronal pyroglutamate-Abeta 3-42 triggers neurodegeneration and lethal neurological deficits in a transgenic mouse model. *Acta Neuropathol.* **2009**, *118*, 487–496. [CrossRef] [PubMed]

193. Bersin, L.M.; Patel, S.M.; Topp, E.M. Effect of 'pH' on the rate of pyroglutamate formation in solution and lyophilized solids. *Mol. Pharm.* **2021**, *18*, 3116–3124. [CrossRef] [PubMed]

194. Castagnola, M.; Cabras, T.; Vitali, A.; Sanna, M.T.; Messana, I. Biotechnological implication of the salivary proteome. *Trends Biotechnol.* **2011**, *29*, 409–418. [CrossRef] [PubMed]

195. Schulte, S.J.; Huang, J.; Pierce, N.A. Hybridization Chain Reaction Lateral Flow Assays for Amplified Instrument-Free At-Home SARS-CoV-2 Testing. *ACS Infect. Dis.* **2023**, *9*, 450–458. [CrossRef]

196. Yu, H.; Lee, H.; Cheong, J.; Woo, S.W.; Oh, J.; Oh, H.K.; Lee, J.H.; Zheng, H.; Castro, C.M.; Yoo, Y.E.; et al. A rapid assay provides on-site quantification of tetrahydrocannabinol in oral fluid. *Sci. Transl. Med.* **2021**, *13*, eabe2352. [CrossRef] [PubMed]