Elsevier Editorial System(tm) for Journal of Proteomics
Manuscript Draft

Manuscript Number:

Title: The intriguing heterogeneity of human salivary proline-rich proteins

Article Type: SI:Protein Species

Section/Category: Review Article

Keywords: Acidic Proline-rich proteins;
Basic Proline-rich proteins;
N-/O-glycosylated proline-rich proteins;
Phosphorylation;
N-pyroglutaminylation

Corresponding Author: Dr. Irene Messana, PhD

Corresponding Author's Institution: Cagliari University

First Author: Barbara Manconi, PhD

Order of Authors: Barbara Manconi, PhD; Massimo Castagnola, PhD; Tiziana Cabras, PhD; Alessandra Olianas, PhD; Alberto Vitali, PhD; Claudia Desiderio, PhD; Maria T Sanna, PhD; Irene Messana, PhD

Abstract: The most heterogeneous family of human salivary proteins is represented by proline-rich proteins (PRPs) divided in acidic, basic, and basic glycosylated (aPRPs, bPRPs, gPRPs). They are encoded by six genes, clustered on chromosome 12p13.2: PRH1-2 encode aPRPs, PRB1-4 encode bPRPs and gPRPs. Each gene exists in different allelic forms: two for PRH2, three for PRH1, PRB2, and PRB4, four for PRB1, and PRB3. During granule maturation, PRP proproteins undergo proteolysis by the action of convertases and carboxypeptidases. Differently from bPRPs, proteolysis of aPRPs is not complete, and, besides fragments, entire protein species are also secreted. Maturation process generates ten aPRPs (PRP-1, PRP-2, PIF-s, Db-s, Pa, PRP-3, PRP-4, PIF-f, Db-f, P-C), and at least 18 bPRPs (II-2, P-E, IB-6, Ps-1, Ps-2, IB-1, P-J, IB-8a, P-F, P-H, P-D, II-1, protein glycosylated A, CD-IIg, and Gl1-4). In addition, single nucleotide and length polymorphisms, and differentially spliced transcripts originate several natural variants. Phosphorylation, N-pyroglutaminylation, dimerization, and N-/O-glycosylation also occur during maturation, enlarging the number of protein species, further increased by proteolytic events governed by carboxy- and endo-peptidases during and after secretion, and giving rise to a huge number of small peptides. The PRPs functional role is still poorly understood.

Suggested Reviewers: Francisco M Amado PhD
Prof, Dep of Chemistry, University of Aveiro
famado@ua.pt
He is an expert in the field of human saliva proteomics

To the Editors of the Special Issue on Protein Species - Journal of Proteomics

Drs. Jungblut, Schlüter and Thiede

Dear Editors,

I am submitting the review entitled: " The intriguing heterogeneity of human salivary proline-rich proteins" by B. Manconi, et al. for the special issue on Protein Species.

The review is focused on proline-rich proteins, a family of salivary proteins with high structural heterogeneity and inter-individual variability. They are products of six genes clustered on chromosome 12p13.2 and comprise a mixture of entire, truncated, phosphorylated, glycosylated and dimerized protein/peptide species, sharing large part of their sequences, and possibly involved in different biological activities.

In the hope that the paper may be considered for the publication, I send You my best regards. Yours sincerely.

Irene Messana

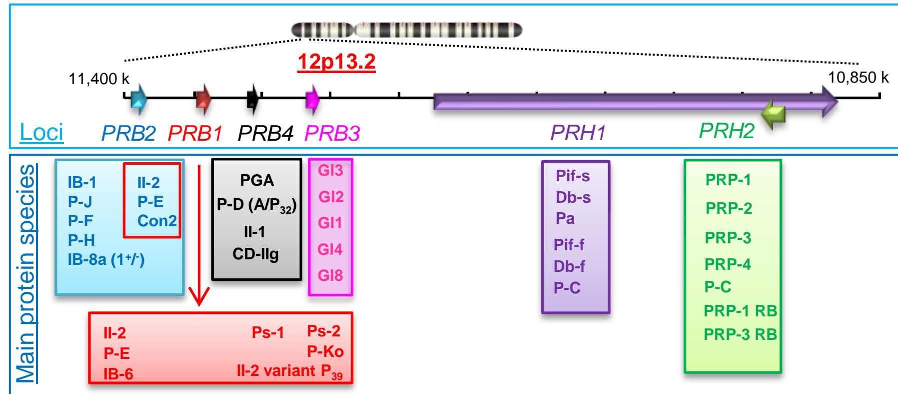
The intriguing heterogeneity of human salivary proline-rich proteins

Short title: Salivary proline-rich protein species

Significance

The high polymorphism of PRPs gives an important contribution to the high heterogeneity and inter-individual variability of the human salivary proteome. The products of six genes clustered on chromosome 12p13.2 comprise a mixture of entire, truncated, phosphorylated, glycosylated and dimerized protein/peptide species, sharing large part of their sequences, and possibly involved in different biological activities. Whatever the role of PRP species is, it should be crucial, given that PRPs are the most conserved oral salivary proteins among mammals.

Graphical Abstract



Highlights

- Proline-rich proteins are a large family of salivary proteins of secretory origin.
- Polymorphic stop codons and alternative splicing generate multiple protein species.
- Heterogeneity is increased by several pre- and post-secretory modifications.
- Sequences of human PRP species are largely shared.
- Top-down proteomics allowed unraveling this complex family.

*Manuscript
Click here to download Manuscript: Manconi et al. manuscript.docx Click here to view linked References

The intriguing heterogeneity of human salivary proline-rich proteins

Short title: Salivary proline-rich protein species

Barbara Manconi[a1], Massimo Castagnola[b,c1], Tiziana Cabras[a], Alessandra Olianas[a], Alberto Vitali[c], Claudia Desiderio[c], Maria Teresa Sanna[a], Irene Messana[a*]

[a]Department of Life and Environmental Sciences, Biomedical section, University of Cagliari, Cagliari, Italy
[b]Biochemistry and Clinical Biochemistry Institute, Medicine Faculty, Catholic University of Rome, Roma, Italy
[c]Institute of Chemistry of the Molecular Recognition, CNR, Roma, Italy

*Corresponding author at:

Department of Life and Environmental Sciences

University of Cagliari

Monserrato Campus

09042 Monserrato (CA) Italy

Tel: +390706754520

E-mail: imessana@unica.it

[1] Manconi, Barbara and Castagnola, Massimo contributed equally to this manuscript

Abstract

The most heterogeneous family of human salivary proteins is represented by proline-rich proteins (PRPs) divided in acidic, basic, and basic glycosylated (aPRPs, bPRPs, gPRPs). They are encoded by six genes, clustered on chromosome 12p13.2: *PRH1-2* encode aPRPs, *PRB1-4* encode bPRPs and gPRPs. Each gene exists in different allelic forms: two for *PRH2,* three for *PRH1*, *PRB2*, and *PRB4*, four for *PRB1*, and *PRB3*. During granule maturation, PRP proproteins undergo proteolysis by the action of convertases and carboxypeptidases. Differently from bPRPs, proteolysis of aPRPs is not complete, and, besides fragments, entire protein species are also secreted. Maturation process generates ten aPRPs (PRP-1, PRP-2, PIF-s, Db-s, Pa, PRP-3, PRP-4, PIF-f, Db-f, P-C), and at least 18 bPRPs (II-2, P-E, IB-6, Ps-1, Ps-2, IB-1, P-J, IB-8a, P-F, P-H, P-D, II-1, protein glycosylated A, CD-IIg, and Gl1-4). In addition, single nucleotide and length polymorphisms, and differentially spliced transcripts originate several natural variants. Phosphorylation, N-pyroglutaminylation, dimerization, and *N*-/*O*-glycosylation also occur during maturation, enlarging the number of protein species, further increased by proteolytic events governed by carboxy- and endo-peptidases during and after secretion, and giving rise to a huge number of small peptides. The PRPs functional role is still poorly understood.

Significance

The high polymorphism of PRPs gives an important contribution to the high heterogeneity and inter-individual variability of the human salivary proteome. The products of six genes clustered on chromosome 12p13.2 comprise a mixture of entire, truncated, phosphorylated, glycosylated and dimerized protein/peptide species, sharing large part of their sequences, and possibly involved in different biological activities. Whatever the role of PRP species is, it should be crucial, given that PRPs are the most conserved oral salivary proteins among mammals.

Keywords. Acidic Proline-rich proteins; Basic Proline-rich proteins; *N*-/*O*-glycosylated proline-rich proteins; Phosphorylation; N-pyroglutaminylation

1. Introduction

Proline is unique among the common 20 protein amino acids because the cyclization of the side-chain onto the nitrogen atom of the backbone limits its conformation (Φ angle about 65°) and restricts the arrangements of the preceding residue with a prominent preference for β-sheet conformation [1]. As a consequence, polyproline peptides are prone to adopt a structure called polyproline (PP) II helix, distinctive for a continuous hydrophobic strip around the surface of the helix. PP II helices are very common motifs in globular proteins, where they are generally solvent exposed [2]. The relative rigidity connected to the low conformational entropy upon binding affords weak and very fast on- and off-rates for binding with other widespread motifs, such as SH3 and WW domains [3]. Weak, but specific, interactions in intracellular signaling pathways are at the basis of the preponderance of proline-based recognition motifs in the eukaryotic proteome [4]. Indeed, proline-rich sequences are largely represented in proteins/peptides involved in processes requiring fast recruitment or interchange with cognate proteins, such as cytoskeletal rearrangements, signaling cascade or transcription initiation, and thus they are very common both in the animal and vegetable kingdoms [5]. In humans, proline-rich proteins (PRPs) or proline-rich domains are very common and participate to almost all biological processes, such as RNA splicing and processing [6], modulation of signaling pathways [7], misfolded protein binding [8], collagen fibril organization and chondrocyte development [9], to cite a few. Proteins involved in the formation of the epithelial cornified envelope (cornifins, involucrins and small PRPs) [10], and salivary PRPs are the largest and most heterogeneous families of human PRPs. Indeed, PRPs represent more than 20-30% (w/w) of total proteins in whole human saliva and more than 50-60% (w/w) of proteins of parotid saliva [11]. They are commonly divided in three classes: acidic PRPs (aPRPs), basic PRPs (bPRPs), and glycosylated (basic) PRPs (gPRPs) [12]. The six genes encoding salivary PRPs are localized in a cluster on chromosome 12p13.2 (Fig. 1), two (*PRH1* and *PRH2*) encoding for aPRPs and four (from *PRB1* to *PRB4*) for bPRPs. Proline is the predominant amino

acid in salivary PRP sequences (25-40% of all amino acids), but Gly and Gln are also highly represented, and globally these three amino acids account from 70 to 88% of all the residues [13]. Salivary PRPs are unique among the PRP families for the complete absence of hydroxyproline, hydroxylysine, and aromatic amino acids. As it will be better described in the following sections, the major aPRPs are 150 residue-long and the acidic portion is restricted to the first 30 residues for the presence of many Asp and Glu residues. The remaining part of the sequence shows high similarities with bPRPs and is highly repetitive, although aPRP repeats differ slightly from bPRP repeats. Due to these structural features, bPRPs and aPRPs elute as distinct chromatographic clusters in RP-HPLC separations (Fig. 2). While aPRPs are secreted by both parotid and submandibular/sublingual glands (in different percentages), bPRPs are secreted only by parotid glands. A further distinction between aPRPs and bPRPs is that while aPRPs can be found in saliva both as intact and truncated proteoforms, bPRPs encoded by *PRB1*, *PRB2* and *PRB4* genes are detectable in saliva only as fragments of the bigger proproteins. bPRPs are more polymorphic than aPRPs. Alleles of bPRPs exhibit tandem repeat length variation, and single nucleotide polymorphisms (SNP) in the coding region as well as polymorphic cleavage sites. Polymorphic stop codons and alternative splicing result in multiple transcript variants encoding distinct proteins [14-17]. In addition to proteolytic fragmentation, other PTMs, such as glycosylation, phosphorylation, and N-terminal pyroglutamic acid (pGlu) formation, occur before, during, and after secretion on these proteins. As a result, PRPs are characterized by a very high heterogeneity and inter-individual variability that have made their identification and characterization demanding. Top-down proteomics represents the solely approach able to investigate the naturally occurring protein species present in a biological sample [18], including the human salivary PRPs. In fact, in bottom-up proteomics, the presence of a protein in a complex sample is inferred by LC-MS/MS detection of proteotypic peptides generated by enzymatic digestions of the sample before the analysis. This approach reveals its limits when the goal is to evidence peptides or proteins carrying multiple PTMs in sites separated by the digestion. Moreover, it is inapplicable to investigate naturally occurring

proteolytic fragmentations. Conversely, the top-down approach, based on the analysis of the "intact" proteome, represents an exclusive strategy for the comprehensive observation of protein species, but suffers from several limits, the major linked to the detection of high molecular size and/or poorly soluble proteins.

In this review the advances in the characterization, mainly performed by top-down MS platforms, of human salivary PRPs and their most relevant PTMs, contributing to enlarge their numerous protein species, will be summarized and discussed.

2. Salivary acidic proline-rich proteins

Two loci, *PRH1* and *PRH2*, encode for aPRPs [19]. In western population *PRH2* locus is commonly biallelic and gives rise to PRP-1 and PRP-2, two protein species of 150 amino acid residues differing only at position 50 (Asn/Asp, respectively) (Fig. 3). Three different alleles of *PRH1* locus give rise to the parotid isoelectric-focusing variant slow (PIF-s), the parotid acidic protein (Pa), both 150 residues long, and the double band isoform slow (Db-s), 171 amino acid residues long. The names, deriving from electrophoretic and isoelectric focusing separation of human parotid salivary proteins, are confusing because all the different aPRPs are secreted by both parotid and submandibular/sublingual glands, with a relative contribution of about 80% and 20%, respectively [20]. The three protein species encoded by *PRH1* locus differ: a) for the residue at position 26 that is Leu for Db-s and Pa, and Ile for PIF-s; b) for the insertion in Db-s of a 21 amino acid residues repeat after position 81; c) for the residue at position 103 that is Cys in Pa and Arg in PIF-s. In the Db-s isoform the Arg is shifted to the 124 position. Pa is commonly present in human saliva as a Pa-dimer, generated by the formation of a disulfide bond between the Cys_{103} residues of the monomers (Fig. 3).

The possible combinations of the most common *PRH1* and *PRH2* alleles generate 18 different phenotypes in the western population. Concerning the *PRH1* locus, the *PIF* allele is more frequent

(about 67% in the adult population) than the *Pa* (18%) and the *Db* (15%) ones [19,21,22]. Indeed, PIF homozygosis is the most common phenotype detectable in adults (about 45%), followed by PIF-Pa heterozygosis (24%), PIF-Db heterozygosis (20%), Db-Pa heterozygosis (6%), and Pa (3%) and Db (2%) homozygosis. Concerning the *PRH-2* locus, the frequency of the *PRP-1* allele is about 71% and that of *PRP-2* allele is 29% [19] because PRP-1 homozygosis is the most common phenotype (50%) followed by PRP-1/PRP-2 heterozygosis (41%) and PRP-2 homozygosis (9%). Other uncommon allelic products of the *PRH1* locus (At, Au, Aw) were found in Asiatic populations, but were not structurally characterized [23,24]. Recently, during a screening of the acid-soluble fraction of about 200 human adult whole saliva samples, a new isoform called PRP-1 Roma-Boston $pSer_{22}$→Phe (PRP-1 RB) has been characterized by top-down high-resolution HPLC-ESI-MS/MS in three subjects [25] (Table 1, Fig. 3). This isoform has been detected at low frequency (0.7%) and might correspond to one of the two uncommon structurally not characterized isoforms detected by Hay et al. [19] during a survey of human salivary samples by anion-exchange HPLC.

Top-down HPLC low-resolution ESI–MS analysis is not able to distinguish PRP-1, with an avarage mass (Mav) of 15354 Da, PIF-s (Mav 15354 Da), and PRP-2 (Mav 15355 Da), which were therefore referred to as PRP-1 type proteins (Fig. 3) [26,27]. On the other hand, Db-s (Mav 17473 Da) and Pa 2-mer (Mav 30682 Da) can be easily discriminated.

Before granule storage, PRP-1, PRP-2, PIF-s and Db-s undergo a partial cleavage at Arg_{106} residue (Arg_{127} in Db-s) [20] by a microsomal proprotein convertase recognizing the RPPR↓ consensus sequence [28], originating a common 44-amino acid residues peptide (named P-C) and four truncated derivatives, PRP-3, PRP-4, PIF-f and Db-f (f stands for fast) [19,29]. The cleavage is only partial, leading to the presence in saliva of both full-length and truncated aPRPs [30]. Pa does not undergo this cleavage for the loss of the RPPR consensus sequence, due to the substitution Arg_{103}→Cys [31] (Fig. 3).

PRP-3 (Mav 11001 Da), PIF-f (Mav 11001 Da), and PRP-4 (Mav 11002 Da) cannot be distinguished by HPLC low-resolution ESI-MS analysis, and consequently, these proteins have been referred to as PRP-3 type (Fig. 3) [26].

It should be outlined that anionic PAGE of human parotid saliva succeeded in separating PRP-2 from PRP-1/PIF-s, and PRP-4 from PRP-3/PIF-f [32], and visualizing all the bands after staining and destaining of the gel in acetic acid as reported by Beeley et al.[33]. Top-down high-resolution HPLC-ESI-MS by Orbitrap demonstrated the capability of discriminating PRP-1/PIF-s from PRP-2 and PRP-3/PIF-f from PRP-4 [25] (Table 1).

In several studies focused on the proteomic characterization of whole human saliva by 2-DE and MALDI-TOF-MS analysis, among the high number of proteins identified, no aPRPs were revealed [34,35] with the exception of one paper that assigned different pink spots from whole, parotid and submandibular/sublingual saliva samples to aPRPs on the basis of the characteristic metachromatic reaction of salivary PRPs after Coomassie Blue R-250 staining [36].

In adult human saliva both full-length and truncated aPRPs present a pGlu at the N-terminus and are mainly diphosphorylated on Ser_8 and Ser_{22} (Fig. 3), by the action of Fam20C, a physiological casein kinase that phosphorylates multiple secreted proteins within a SXE/pS consensus sequence [37,38]. It should be outlined that Fam20C phosphorylates Ser_{22} by recognizing the minor consensus motif (S(X)3/4(E/D/pS)3). Even before the identification of this kinase, in a human submandibular cell line transfected with a vector encoding PRP it was shown that phosphorylation of Ser_8 or Ser_{22} in aPRPs was not hierarchical, since it did not depend on the other site being phosphorylated [39]. Later on, it was shown, directly on human saliva, that monophosphorylated aPRP derivatives were represented by a mixture of phosphorylated aPRPs either on Ser_8 or on Ser_{22} [21]. Low percentages of triphosphorylated (on Ser_8, Ser_{22} and Ser_{17}) and monophosphorylated (either on Ser_8 or Ser_{22}) PRP-1 type, or Db-s have been detected, while non-phosphorylated PRP-1 type resulted undetectable (Fig. 3) [21]. Phosphorylation on Ser_{17} is probably hierarchical, because it requires the previous phosphorylation of Ser_{22} to generate the secondary consensus sequence recognized by

Fam20C. This hypothesis is in agreement with the lack of detection of diphosphorylated protein species of PRP-1 Roma-Boston (or PRP-3 RB) variant, which shows a Phe residue in substitution to the phoshorylated Ser_{22}. PRP-1 and PRP-3 RB proteins are phosphorylated only on Ser_8 (Table 1).

In adult saliva, diphosphorylated aPRPs truncated at C-terminal Arg ($desArg_{106}$) have been always detected and corresponded to about 10% of the parent aPRP, while P-C $desGln_{44}$ (truncated at the C terminus) was the 2% and P-C $desPro_{43}Gln_{44}$ the 1.5% [21].

In addition, in saliva of healthy subjects, P-C peptide undergoes further proteolysis, generating different fragments [40], mainly originating from cleavages at the XPQ↓G site (with X being preferably K, but also S or R), and probably related to a glutamine endoproteinase of microbial origin active in the oral cavity [41]. P-C peptide fragmentation has been found increased in type 1 diabetic patients, indicating an higher activity of this proteolytic enzyme in the oral cavity of these patients [42].

Studies devoted to the characterization of the age-related aPRP expression evidenced a ten times lower concentration of aPRPs in newborns saliva with respect to adults, showing an increasing trend after birth [26]. Additionally, the protein species encoded by *PRH2* have been observed to appear early during foetal life with respect to the protein species encoded by *PRH1*, expressed later, after the normal term of delivery [26]. The increase of aPRPs with age has been confirmed by both Manconi et al. [43], studying the salivary proteome of children aged between 0 and 48 months, and Cabras et al. [44], showing that aPRP concentration reached a minimum around 6–9 years of age, probably in concomitance with replacement of deciduous dentition.

By measuring the percentages of nonphosphorylated and monophosphorylated aPRPs in saliva samples from subjects of a wide age range, starting from preterm newborns up to 57-years-old adults, their exponential decrease with the increase of age was revealed [27]. The study additionally highlighted that the Fam20C kinase involved in aPRP phosphorylation is not fully active even at birth, reaching an activity comparable to that of adults at around 2 years of age.

aPRPs show a high affinity for hydroxyapatite [45] and are involved in the formation of acquired enamel pellicle [46]. aPRP binding to hydroxyapatite involves the acidic N-terminal domain and exposes the proline-rich C-terminal domain to oral bacteria binding. aPRPs play also an important role in the protection of tooth enamel by inhibiting calcium phosphate precipitation, and thus promoting calcium homeostasis [47].

3. Salivary basic proline-rich proteins

bPRPs, constituting more than 30% of the proteins secreted by the human parotid glands [11, 20], are the most composite family of salivary proteins [48-54], without considering the several glycoforms of some bPRPs, and the numerous small peptides generated by post-secretory fragmentation [20,40,55,56].

The cluster of genes encoding for bPRPs includes *PRB1-PRB4*, each one existing in several allelic forms: at least four alleles (S, small; M, medium; L, large; and VL, very large) are present in the western population at *PRB1* and *PRB3* loci, and three (S, M, L) at *PRB2* and *PRB4* loci (Fig. 1) [16, 57-59]. These alleles exhibit tandem repeat length variations and SNPs in the coding region, polymorphic cleavage sites and polymorphic stop codons. Moreover, alternative splicing generates multiple transcript variants encoding distinct protein species, and some alleles are still pending for their characterization [14-17].

Genetic variability, PTMs implicated in the pre-secretory maturation processes and further transformations occurring in the oral environment give a contribution to the heterogeneity of bPRPs. The proteolytic cleavage is the main post-translational event occurring on bPRPs; indeed, the pre-proproteins encoded by each allele, after peptide-signal removal, are completely cleaved by proprotein convertases before secretion, except for the protein encoded by the *PRB3* locus that originates gPRPs [20,27,60]. A presecretory cleavage of bPRP proproteins occurs at the consensus sequence KSRSXR↓, where X may be P, S, or A [59,61], recognized by proprotein convertases.

This presecretory event generates the following main bPRP species: (a) II-2 peptide from PRB1-(S, M, L) proproteins, P-E peptide, and IB-6 protein from PRB1-S, Ps-1 protein from PRB1-M, Ps-2 protein from PRB1-L (Table 2) [27,58,59,62]; (b) IB-1, P-J, P-H, P-F peptides, and IB-8a protein from PRB2-L proprotein [27,58, 59,62] among these, P-F, P-J, IB-8a are further processed during the granule maturation with removal of the C-terminal arginine residue [20,44,63,64] by the action of carboxypeptidase Z [65] (Table 3); (c) P-D peptide from PRB4-(S, M, L) proproteins, together with glycosylated proteins [27,58,59,62] (Table 4).

During granule maturation, II-2 and IB-1 undergo both cyclization of the N-terminal Gln residue (N-pyroglutaminylation), and phosphorylation of $Ser_{8,}$ generating the mature forms [49,64] (Table 2-3). N-pyroglutaminylation catalyzed by glutaminyl cyclase [66] protects peptides and proteins from the proteolytic action of amino peptidases [67]. It has also been demonstrated that the pGlu moiety is necessary for both activity and receptor binding ability of several neuropeptides and hormones [68]. The phosphorylation is almost a complete event, since only low percentages (less than 1%) of the nonphosphorylated forms were detected in parotid granules, parotid and whole saliva, and probably occurs after the cleavage of the proprotein [20,44,64]. Phosphorylation of II-2 and IB-1 occurs by the Fam20C kinase, [38], acting also on aPRPs and many other salivary peptides and proteins, such as histatin 1, statherin and cystatin S [61].

The complexity of the bPRP species is enhanced by the existence of several variants generated by nucleotide polymorphisms or alternative splicing affecting bPRP alleles. Indeed, the bPRP with a Mav of 10433.5 Da detected in whole saliva and in parotid secretory granules [20, 64], and named P-Ko by Halgand et al. [49], is encoded by *cP4*, a differentially spliced transcript of *PRB1L* allele [14] (Table 2). cP4 proprotein lacks the sequence 106-299 of PRB1-L (P04280), and its cleavage generates II-2 peptide and P-Ko protein. Another differentially spliced transcript of *PRB1L* is *cP5* [14] encoding a proprotein lacking the sequence 106-319. The cP5 proprotein should generate II-2 peptide and a peptide with an average mass of 8391.2 Da, never detected in saliva up to now. A variant of II-2 peptide, lacking the proline residue at position 39 was also recently found [49]. The

substitution Ser_{100}→Pro due to a SNP determines the existence of two species of the IB-8a protein [58]. IB-8a $Con1^+$ carries a Ser at position 100, and its glycosylated forms on Asn_{98} will be discussed in the next section. IB-8a with Pro_{100} is named $Con1^-$ and it is not glycosylated on Asn_{98}. Both protein species were detected in parotid granules and whole saliva by top-down proteomics [20,48,63]. Halgand et al. found in parotid saliva a protein species with the molecular mass of the mature form of IB-8a $Con1^-$ (theoretical Mav 11896.2 Da). The sequence established on the basis of MS/MS experiments corresponded to that of IB-8a protein lacking the two C-terminal serine residues, carrying the substitution Gln_{115}→His, one hexose bound to Ser_{120}, and a methylation on the C-terminal Arg_{121} [49] (Table 3). The UniprotKB database reports another potential natural variant of IB-8a originated from the substitution Gln_{59}→Arg. This protein species has never been observed and its existence has been postulated on the basis of a DNA sequence present in the dbSNP database. Two protein species of P-D peptide, differing for the single substitution Ala_{32}→Pro, may derive from all the three alleles on *PRB4* locus and are detectable in parotid and whole saliva [48,58] (Table 4). Other three potential protein species originate from *PRB4L* allele if the SNPs reported on UniprotKB are considered. As for the *PRB2L* allele, the sequences of these natural variants are deduced from DNA sequences present on dbSNP.

Protein and DNA databases are often a puzzling source of information for the scientists engaged in the identification of bPRP species. Indeed, several other potential bPRP species should be detected in human saliva given the number of protein sequences obtained from cDNA or genomic DNA large-scale studies. After a search at UniprotKB site on “human PRB1” at least 6 results are obtained, two concerning *PRB1L* (P04280), and *PRB1M* (Q86YA1) alleles, and the others deriving from nucleotide polymorphisms of *PRB1* gene and corresponding to A5D903 [69], G3V1R1, G59X6 and G3V1M9 codes [70]. A similar situation occurs for the *PRB4* gene, with new species that should originate from the nucleotide polymorphisms A2VCL9 [69], E9PAL0 and E7EXA8 [70]. The proproteins deriving from these polymorphisms have one or two convertase consensus

sequences, and during the maturation process should generate smaller peptides. However, the expected molecular masses have never been observed and could be objects of future investigations.

During secretion and after their release in the oral cavity, bPRPs undergo further proteolytic events (Table 2-4). For instance, P-E desArg$_{61}$, II-2 desArg$_{75}$, and IB-1 desArg$_{96}$ peptides, generated by the removal of the C-terminal arginine residue, were detected in parotid saliva collected from the duct, largely in whole saliva, but never in parotid granule content [20]. Instead, II-2 des$_{72-75}$ and IB-1 des$_{93-96}$ protein species, deriving from removal of the four C-terminal residues, have been detected only in whole saliva. Sporadically the P-J des$_{60-61}$, named IB-7 peptide and deriving from removal of the three C-terminal residues in P-J peptide [44], was also observed in parotid and in whole saliva [20,71]. These data could suggest a release of carboxypeptidase from ductal cells or a transfer of carboxypeptidase activity from secretory granules to oral cavity, where the peptides II-2 and IB-1 seem to be, among bPRPs, the most susceptible substrates.

The action of endo-proteases, also of microbial origin, in the oral environment gives rise to further fragmentation of the main bPRP species generating peptides 7-20 residue-long. The presence of multiple repeats in the bPRP sequence makes difficult to assign these fragments to univocal bPRP parent. However, II-2, P-E, IB-6, Ps-1 and Ps-2, deriving from *PRB1*, and IB-1, P-J, P-F, IB-8a and P-H, deriving from *PRB2L,* appeared to be the most susceptible to the proteolysis [20,40]. A great number of bPRP fragments were detected in whole saliva by different top-down proteomics approaches [55]. In a recent paper Vitorino et al. characterized 25 bPRP fragments carrying N-terminal pGlu and 8 carrying phosphorylation [72]. Cleavages occurred at the sites Q/G, P/Q, P/P, Q/K, Q/P, S/G, K/P, G/N, G/D, and R/P, the Q/G resulting predominant [20,40,55,56,71,73], being the XPQ consensus sequence (where X is prominently K), the most frequently detected in fragments [40]. It should be outlined that the deep knowledge on the multiple bPRP species detectable in saliva, including their natural variants, has been possible thanks to the application of top-down proteomics and peptidomics platforms, for their ability to investigate complex protein mixtures in their naturally occurring forms [20,44,48,49,55,63,64,72].

Some of bPRP fragments have been detected as recurrent non-covalent bound components of the acquired enamel surface, suggesting a functional significance for this proteolytic event. Vitorino et al. characterized six fragments from IB-8a ($Con1^{+}/Con1^{-}$), one from P-E, one from IB-1, and one from each PRB1-S, M, L proteins [74]. Different enamel bounding peptides were identified by Siqueira et al., nine deriving from PRB2-L-related protein species and three from PRB1-S, M, L-related protein species [75].

The functional roles of bPRPs is far to be completely understood. However, the information obtained up to now evidenced that, in spite of structural similarities, the protein species generated from the same proprotein may show distinct biological functions. A good example is represented by II-2 peptide and Ps-1 protein, encoded by *PRB1M* allele, which have been shown to be involved in the PROP bitter taste responsiveness [76,77]. II-2 peptide is also encoded by *PRB1S*, together with P-E peptide and IB-6 protein, and by *PRB1L* together with Ps-2 protein. However, the functional role of P-E, IB-6, and Ps-2 is completely different. Indeed, P-E, and IB-6 together with IB-1, P-H [78,79] and P-D peptides [80,81] are involved in precipitation of the harmful tannins. In addition, it has been demonstrated that IB-6 protein acts as receptor for *Candida albicans* adhesion on hydroxyapatite surface [82].

Interestingly, the salivary level of the major bPRP species is strictly connected to age. Indeed, they are almost undetectable in saliva before 3 years of age [27] and their amount increases with age reaching the adult level during adolescence [44]. Processes of growth and hormonal maturation may influence salivary gland function, and changes in the diet occurring with age might partly influence bPRP salivary levels. In this respect a higher number of bPRP small fragments deriving from Q/G cleavage were detected in 6-month-old infants receiving solid food compared to infants on a milk-based diet [73]. These data suggested that diet variation, causing a major load of microbial flora, leads to an increased activity of the endoglutamine protease responsible for fragment generation.

4. Salivary glycosylated proline-rich proteins

A contribution to bPRP heterogeneity derives from *N*- and *O*-glycosylations that give rise to the family of gPRPs. *N*-glycan moieties are attached to the amide group of asparagine residues within the sequon NXS/T, where X may be any amino acid excluding proline [83,84], and share a common pentasaccharide core Man3GlcNAc2, which, after processing in the Golgi network, originates the three main classes of *N*-linked glycan classes: High-mannose, Hybrid and Complex [85]. Unlike *N*-linked glycosylation, no consensus sequence has been found for *O*-linked glycosylation [86], termed 'mucin-type *O*-glycosylation' being mucins the major example of *O*-glycoproteins. *O*-acetylgalactosamine (*O*-GalNAc) and *O*-acetylglucosamine (*O*-GlcNAc) are the most common *O*-sugars of the core; *O*-GalNAc is attached to the hydroxyl group of serine or threonine residues through an α-linkage, while *O*-GlcNAc is attached through a β-linkage [86]. Thus, characterization of glycoprotein species is a difficult task, due to their high heterogeneity deriving from the combination of multiple glycosylation sites and different oligosaccharide structures. Identification of the glycosylation sites and of the glycan structures can be accomplished after releasing of the glycans from the protein, followed by mapping of the glycosylation sites by protease digestion. Another strategy is based on intact glycopeptide MS analysis, and this approach allows simultaneous identification of both the peptide sequence and the site-specific glycan structures [87]. However, by this approach it is not possible to define how multiple glycosylation sites and different oligosaccharides are combined, and thus the glycoprotein species present in the sample remain uncharacterized. To date the complete characterization of the different glycosylated protein species present in saliva has been achieved only for IB-8a Con1$^+$. Six glycosylated protein species of IB-8a, *N*-glycosylated at Asn_{98} have been characterized together with the nonglycosylated protein, in adult human saliva by HPLC-ESI-MS [63]. Five of the glycosylated species carry a biantennary *N*-linked glycan fucosylated in the innermost N-acetylglucosammine of the core and show from zero to four additional fucoses in the antennal region. The sixth glycoform carries a monoantennary

monofucosylated oligosaccharide (Fig. 4). The frequency of the different protein species of IB-8a was determined in 71 subjects [63] (Table 5). The *PRB2*-like Con2$^+$ glycoprotein has a single glycosylation site and it is encoded by a hybrid *PRB1M CON2*$^+$ allele originated by a conversion between *PRB1* and *PRB2* genes (Table 3). Amino acid sequencing of Con2$^+$ evidenced that the portion of the protein surrounding the glycosylation site is the same of Con1$^+$, but no information about the linked glycan moieties exists [58].

Unlike IB-8a protein species the other gPRPs possess a variable number of glycosites, and some of them are both *N*- and *O*-glycosylated making the protein species characterization a difficult task. The PRB3 gene encodes highly glycosylated proteins known as parotid salivary glycoproteins Gl (Table 6). At least nine size variants of Gl proteins can be found in different populations [88-90]. In black and white populations the four allelic size variants *S, M, L* and *VL* encode for the corresponding Gl protein size variants Gl4/*PRB3VL* > Gl1/*PRB3L* > Gl2/*PRB3M* > Gl3/*PRB3S* [62]. The Gl8 glycoprotein derives from a single nucleotide insertion in the *PRB3S*^{Cys} allele, which converts Arg_{15} to Cys. Gl8 protein is electrophoretically distinct from the other Gl protein variants because they form a disulfide-bonded heterodimer with salivary peroxidase [89]. Each Gl protein carries a different number of putative glycosylation sites depending of the length of the polypeptide backbone (Table 6) [59]. Gl proteins possess both *O*- and *N*-glycosylation sites [91,92]. The most abundant *N*-linked oligosaccharide was the highly fucosylated species represented by a biantennary asialosaccharide containing two fucose moieties on one antenna and an unsubstituted terminal lactosamine moiety on the other [92].

Obviously, the approach utilized did not allow defining the number of protein species deriving from the combination of the different glycans at each glycosyte.

To date the possible glycoprotein species associated to II-1, from *PRB4M* [58,59,62], as well as to protein glycosylated A (PGA), and to CD-IIg, potentially assigned to *PRB4S*, and *L,* respectively, [14,62,93,94] (Table 4) have not been characterized.

LC-MALDI-TOF/TOF analysis of the low molecular weight fraction of whole saliva evidenced *O*-glycosylation sites at Ser_{87} and Ser_{330} of *PRB1*-related protein species [72]. In particular, the *O*-glycosylation site at Ser_{330} corresponds to Ser_{178} of Ps-1 and Ser_{239} of Ps-2 from PRB1-M and PRB1-L proproteins, respectively. The same results were obtained by Carpenter et al. that evidenced *O*-linked sialylated glycoforms of both Ps-1 and Ps-2 proteins [91].

gPRPs are recognized as the “first line of oral defense” against the detrimental effects of polyphenols in the diet [95] and pathogen infections. Indeed, gPRPs may bind different oral microorganisms, and the glycan structure plays a crucial role in discriminating the different microorganism. For instance, Gillece-Castro et al. [92] demonstrated that the major glycan structure linked to Gl proteins, containing an unsubstituted lactosamine unit (Gal $\beta1\rightarrow$ 4GlcNAc), is a receptor for *Fusobacterium nucleatum,* a periodontal pathogen, as confirmed by abolished interaction *in vitro* between *F. nucleatum* and saliva in *PRB3* null mutations (resulting in absence of Gl2 and Gl4 proteins) [96]. Conversely the presence of sialic acid linked to Gal, or Fuc linked to Gal or GlcNAc, has been demonstrated to prevent adhesion of *F. nucleatum*. *N*-glycans linked to Gl proteins bind to a variety of oral microorganisms other than *F. nucleatum,* such as *Streptococcus mitis* and *S. sanguis* [97,98], both involved in dental plaque formation. Galactose-mediated adhesion of *S. mitis* to gPRP species has been suggested by different authors [99], while binding of *S. sanguis* could be inhibited by removal of terminal sialic acid and by complex oligosaccharides moieties [97,100]. These findings suggest that the high heterogeneity of gPRP species is functional to the protection against different pathogen infections, and that individual variation in the ability of gPRP to support microorganism adherence is most likely due to combination of glycan structures unique for a particular glycoprotein species [100].

5. Concluding remarks

The complexity of PRP species originates at different levels along the pathway from genes to proteins via transcripts, and the presence of multiple repeats in the sequences, together with the extensive fragmentation occurring before and after secretion, makes the structural characterization difficult unless sophisticated and high-resolution proteomic approaches are applied. In addition, a point that definitely requires attention and contributes to increase the complexity of PRP world, concerns the too complex and confusing nomenclature of PRPs, which should be rationalized. For instance, bPRPs names were assigned in different studies following diverse criteria. Ten bPRP peptides characterized by Kaufmann et al. [50] were named by the authors according to the name of the chromatographic fractions. A different nomenclature was proposed by Isemura and coll. [51,54] that assigned to each salivary peptide the name P-X, where X corresponds to a letter from A to I. In the literature, bPRPs are named according to both nomenclatures, and the correspondences are shown in Tables 2-4: P-D peptide corresponds to IB-5, P-E to IB-9, P-F to IB-8c, P-H to IB-4 and P-I to IB-6. A new basic peptide recently characterized by our group was called P-J [44,64]. Other components of the PRP family reported in Tables 1-4 were named with different acronyms without any connection with the previous researches increasing the difficulties to understand the nomenclature for the non-experts in this topic.

The complex mosaic of human salivary PRP species proteomic studies strongly pushes toward a deep understanding of the specific function associated to single mosaic pieces and to the comprehension of the interplay between the various components in preserving the integrity of the oral cavity [61,101]. Even though it is known that PRP species are involved, in coordination with other salivary proteins, in the modulation of oral and gastro-intestinal microbiota, further studies are needed to define how the various phenotypes fulfill microbiota selection.

Still unclear is the biological significance of aPRP partial cleavage, as well as the role of the P-C peptide. The involvement of aPRPs in the formation of acquired enamel and oral epithelial protein pellicle [46] has been well established, but no information about specific functions of entire and truncated aPRPs is available. Furthermore, with the exception of the suggested correlation between

Db allele and caries experience [102], there is no information about the physiological (or pathological) significance of the different aPRPs phenotypes.

Whatever the role of both aPRPs and bPRPs is, it must be crucial, given that PRP are the most conserved oral salivary proteins among mammals [103].

In conclusion, proteomic studies strongly contributed to the elucidation of the intriguing heterogeneity of salivary PRPs, but further investigations are needed to clarify the connection between the several protein species and different functions in the oral cavity.

Acknowledgment

The authors acknowledge the financial support of Cagliari University, Catholic University of Rome, Italian National Research Council (CNR), and Nando Peretti Foundation.

Abbreviations: acidic proline-rich proteins (aPRPs); average mass (Mav); basic proline-rich proteins (bPRPs); glycosylated proline-rich proteins (gPRPs); *O*-acetylgalactosamine (*O*-GalNAc); and *O*-acetylglucosamine (*O*-GlcNAc); proline-rich proteins (PRPs); protein glycosylated A (PGA); PRP-1 (PRP-3) Roma-Boston Ser_{22}(Phos)→ Phe variant (PRP-1/PRP-3 RB); pyroglutamic acid (pGlu); single nucleotide polymorphisms (SNPs).

References

[1] MacArthur MW, Thornton JM. Influence of proline residues on protein conformation. J Mol Biol 1991;218:397-412.

[2] Zafra-Ruano A, Luque I. Interfacial water molecules in SH3 interactions: Getting the full picture on polyproline recognition by protein-protein interaction domains. FEBS Lett 2012;586:2619-30.

[3] Kay BK, Williamson MP, Sudol M. The importance of being proline: the interactions of proline-rich motifs in signaling proteins with their cognate domains FASEB J 2000;14:231-241.

[4] Neduva V, Russell RB. Proline-rich regions in transcriptional complexes: heading in many directions. Sci STKE 2007;2007:pe1.

[5] Mandal A, Mandal S, Park MH. Genome-Wide Analyses and Functional Classification of Proline Repeat-Rich Proteins: Potential Role of eIF5A in Eukaryotic Evolution. PLoS One 2014;9:e111800.

[6] Komuro A, Saeki M, Kato S. Association of two nuclear proteins, Npw38 and NpwBP, via the interaction between the WW domain and a novel proline-rich motif containing glycine and arginine. J Biol Chem 1999;274:36513-19.

[7] Huang B, Porter G. Expression of proline-rich Akt-substrate PRAS40 in cell survival pathway and carcinogenesis. Acta Pharmacol Sin 2005;26:1253-8.

[8] Banerji J, Sands J, Strominger JL, Spies T. A gene pair from the human major histocompatibility complex encodes large proline-rich proteins with multiple repeated motifs and a single ubiquitin-like domain. Proc Natl Acad Sci USA 1990;87:2374-8.

[9] Arndt S, Bosserhoff AK. TANGO is a tumor suppressor of malignant melanoma. Int J Cancer 2006;119:2812-20.

[10] Henry J, Toulza E, Hsu CY, Pellerin L, Balica S, Mazereeuw-Hautier J, Paul C, Serre G, Jonca N, Simon M. Update on the epidermal differentiation complex. Front Biosci 2012;17:1517-32.

[11] Bennick A. Salivary proline-rich proteins. Mol Cell Biochem 1982;45:83-99.

[12] Oppenheim FG, Salih E, Siqueira WL, Zhang W, Helmerhorst EJ. Salivary proteome and its genetic polymorphisms. Ann NY Acad Sci 2007;1098:22-50.

[13] Bennick A. Structural and genetic aspects of proline-rich proteins. J Dent Res 1987;66:457-61.

[14] Maeda N, Kim HS, Azen EA, Smithies OJ. Differential RNA splicing and post-translational cleavages in the human salivary proline-rich protein gene system. J Biol Chem 1985;260:11123-30.

[15] Kim HS, Lyons KM, Saitoh E, Azen EA, Smithies O, Maeda N. The structure and evolution of the human salivary proline-rich protein gene family. Mamm Genome 1993;4:3-14.

[16] Azen EA. Genetics of salivary protein polymorphisms. Crit. Rev Oral Biol Med 1993;4:479-85.

[17] Azen EA, Latreille P, Niece RL. PRBI gene variants coding for length and null polymorphisms among human salivary Ps, PmF, PmS, and Pe proline-rich proteins (PRPs). Am J Hum Genet 1993;53:264-78.

[18] Cabras T, Iavarone F, Manconi B, Olianas A, Sanna MT, Castagnola M, Messana I. Top-down analytical platforms for the characterization of the human salivary proteome. Bioanalysis 2014; 6:563-81.

[19] Hay DI, Ahern JM, Schluckebier SK, Schlesinger DH. Human salivary acidic proline-rich protein polymorphisms and biosynthesis studied by high-performance liquid chromatography. J Dent Res 1994;73:1717-26.

[20] Messana I, Cabras T, Pisano E, Sanna MT, Olianas A, Manconi B, Pellegrini M, Paludetti G, Scarano E, Fiorita A, Agostino S, Contucci AM, Calò L, Picciotti PM, Manni A, Bennick A, Vitali A, Fanali C, Inzitari R, Castagnola M. Trafficking and postsecretory events responsible for the formation of secreted human salivary peptides: a proteomics approach. Mol Cell Proteomics 2008;7:911-26.

[21] Inzitari R, Cabras T, Onnis G, Olmi C, Mastinu A, Sanna MT, Pellegrini MG, Castagnola M, Messana I. Different isoforms and post-translational modifications of human salivary acidic proline-rich proteins. Proteomics 2005;5:805-15.

[22] Azen EA, Maeda N. Molecular genetics of human salivary proteins and their polymorphisms. Adv Hum Genet 1988;17:141-99.

[23] Shintani M, Minaguchi K, Suzuki K, Lim KA. Allelic variants of acidic proline-rich proteins observed in Japanese, Chinese, and Malays. Biochem Genet 1990;28:173-84.

[24] Minaguchi K, Shintani M, Suzuki K. New allelic product of the PRH1 locus coding for salivary acidic proline-rich proteins. Hum Hered 1990;40:221-30.

[25] Iavarone F, D'Alessandro A, Tian N, Cabras T, Messana I, Helmerhorst EJ, Oppenheim FG, 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 Ser_{22} (Phos) → Phe variant. J Sep Sci 2014;37:1896-902.

[26] 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 preterm and at-term newborns. J Proteome Res 2007;6:1371-7.

[27] Messana I, Cabras T, Iavarone F, Manconi B, Huang L, Martelli C, Olianas A, Sanna MT, Pisano E, Sanna M, Arba M, D’Alessandro A, Desiderio C, Vitali A, Pirolli D, Tirone C, Lio A, Vento G, Romagnoli C, Cordaro M, Manni A, Gallenzi P, Fiorita A, Scarano E, Calò L, Passali GC, Picciotti PM, Paludetti G, Fanos V, Faa G, Castagnola M. Chrono-proteomics of human saliva: variations of the salivary proteome during human development. J Proteome Res 2015;14:1666-77.

[28] Cai K, Bennick A. Processing of acidic proline-rich proprotein by human salivary gland convertase. Arch Oral Biol 2004;49:871-9.

[29] Bennick A, McLaughlin AC, Grey AA, Madapallimattam G. The location and nature of calcium-binding sites in salivary acidic proline-rich phosphoproteins. J Biol Chem 1981;256:4741-6.

[30] Hay DI, Bennick A, Schlesinger DH, Minaguchi K, Madapallimattam G, Schluckebier SK. The primary structures of six human salivary acidic proline-rich proteins (PRP-1, PRP-2, PRP-3, PRP-4, PIF-s and PIF-f). Biochem J 1988;255:15-21.

[31] Azen EA, Kim HS, Goodman P, Flynn S, Maeda N. Alleles at the PRH1 locus coding for the human salivary-acidic proline-rich proteins Pa, Db, and PIF. Am J Hum Genet. 1987;41:1035-47.

[32] Oppenheim FG, Hay DI, Franzblau C. Proline-rich proteins from human parotid saliva. I. Isolation and partial characterization. Biochemistry 1971;10:4233-8.

[33] Beeley JA, Sweeney D, Lindsay JC, Buchanan ML, Sarna L, Khoo KS. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis of human parotid salivary proteins. Electrophoresis 1991;12:1032-41.

[34] Ghafouri B, Tagesson C, Lindahl M. Mapping of proteins in human saliva using two-dimensional gel electrophoresis and peptide mass fingerprinting. Proteomics 2003;3:1003-15.

[35] Yao Y, Berg EA, Costello CE, Troxler RF, Oppenheim FG. Identification of protein components in human acquired enamel pellicle and whole saliva using novel proteomics approaches. J Biol Chem 2003;278:5300-8.

[36] Walz A, Stühler K, Wattenberg A, Hawranke E, Meyer HE, Schmalz G, Blüggel M, Ruhl S. Proteome analysis of glandular parotid and submandibular-sublingual saliva in comparison to whole human saliva by two-dimensional gel electrophoresis. Proteomics 2006;6:1631-9.

[37] Tagliabracci VS, Engel JL, Wen J, Wiley SE, Worby CA, Kinch LN, Xiao J, Grishin NV, Dixon JE. Secreted kinase phosphorylates extracellular proteins that regulate biomineralization. Science 2012;336:1150-3.

[38] Tagliabracci VS, Pinna LA, Dixon JE. Secreted protein kinases. Trends Biochem Sci 2013;38:121-30.

[39] Drzymala L, Castle A, Cheung JC, Bennick A. Cellular phosphorylation of an acidic proline-rich protein, PRP1, a secreted salivary phosphoprotein. Biochemistry 2000;39:2023-31.

[40] Helmerhorst EJ, Sun X, Salih E, Oppenheim FG. Identification of Lys-Pro-Gln as a novel cleavage site specificity of saliva-associated proteases. J Biol Chem 2008;283:19957-66.

[41] Helmerhorst EJ, Zamakhchari M, Schuppan D, Oppenheim FG. Discovery of a novel and rich source of gluten-degrading microbial enzymes in the oral cavity. PLoS One 2010;5:e13264.

[42] Cabras T, Pisano E, Mastinu A, Denotti G, Pusceddu PP, 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 Proteomics 2010;9:2099-108.

[43] Manconi B, Cabras T, Pisano E, Sanna MT, Olianas A, Fanos V, Faa G, Nemolato S, Iavarone F, Castagnola M, Messana I. 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 Proteomics 2013;91:536-43.

[44] 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;8:4126-34.

[45] Moreno EC, Kresak M, Hay DI. Adsorption thermodynamics of acidic proline-rich human salivary proteins onto calcium apatites. J Biol Chem 1982;257:2981-9.

[46] Bennick A, Chau G, Goodlin R, Abrams S, Tustian D, Madapallimattam G. The role of human salivary acidic proline-rich proteins in the formation of acquired dental pellicle in vivo and their fate after adsorption to the human enamel surface. Arch Oral Biol 1983;28:19-27.

[47] Hay DI, Carlson ER, Schluckebier SK, Moreno EC, Schlesinger DH. Inhibition of calcium phosphate precipitation by human salivary acidic proline-rich proteins: structure-activity relationships. Calcif Tissue Int 1987;40:126-32.

[48] Castagnola M, Cabras T, Iavarone F, Vincenzoni F, Vitali A, Pisano E, Nemolato S, Scarano E, Fiorita A, Vento G, Tirone C, Romagnoli C, Cordaro M, Paludetti G, Faa G, Messana I. Top-

down platform for deciphering the human salivary proteome. J Matern Fetal Neonatal Med 2012; 25:27-43.

[49] Halgand F, Zabrouskov V, Bassilian S, Souda P, Loo JA, Faull KF, Wong DT, Whitelegge JP. Defining intact protein primary structures from saliva: a step toward the human proteome project. Anal Chem 2012;84:4383-95.

[50] Kauffman DL, Bennick A, Blum M, Keller PJ. Basic proline-rich proteins from human parotid saliva: relationships of the covalent structures of ten proteins from a single individual. Biochemistry 1991;30:3351-56.

[51] 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;91:2067-75.

[52] 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.

[53] 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;93:883-8.

[54] 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;94:1991-9.

[55] Vitorino R, Barros A, Caseiro A, Domingues P, Duarte J, Amado F. Towards defining the whole salivary peptidome. Proteomics Clin Appl 2009;3:528-40.

[56] Huq NL, Cross KJ, Ung M, Myroforidis H, Veith PD, Chen D, Stanton D, He H, Ward BR, Reynolds EC. A review of the salivary proteome and peptidome and saliva-derived peptide therapeutics. Int. J Pept Res Ther 2007;13:547-64.

[57] Lyons KM, Stein JH, Smithies O. Length polymorphisms in human proline-rich protein genes generated by intragenic unequal crossing over. Genetics 1988;120:267-78.

[58] Azen EA, Amberger E, Fisher S, Prakobphol A, Niece RL. PRB1, PRB2, and PRB4 coded polymorphisms among human salivary concanavalin-A binding, II-1, and Po proline-rich proteins. Am J Hum Genet 1996;58:143-53.

[59] Stubbs M, Chan J, Kwan A, So J, Barchynsky U, Rassouli-Rahsti M, Robinson R, Bennick A. Encoding of human basic and glycosylated proline-rich proteins by the PRB gene complex and proteolytic processing of their precursor proteins. Arch Oral Biol 1998;43:753-70.

[60] Chan M, Bennick A. Proteolytic processing of a human salivary proline-rich protein precursor by proprotein convertases. Eur J Biochem 2001;268:3423-31.

[61] 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;31:1948-63.

[62] Lyons KM, Azen EA, Goodman PA, Smithies O. Many protein products from a few Loci: Assignment of human salivary Proline-Rich Proteins to specific loci. Genetics 1988;120:255-265.

[63] 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;35:1079-86.

[64] Messana I, Cabras T, Inzitari R, Lupi A, Zuppi C, Olmi C, Fadda MB, 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.

[65] Fischer WH, Spiess J. Identification of a mammalian glutaminyl cyclase converting glutaminyl into pyroglutamyl peptides. Proc Natl Acad Sci USA 1987;84:3628-32.

[66] Song L, Fricker LD. Cloning and expression of human carboxypeptidase Z, a novel metallocarboxypeptidase. J Biol Chem 1997;272:10543-50.

[67] Van Coillie E, Proost P, Van Aelst I, Struyf S, Polfliet M, De Meester I, Harvey DJ, Van Damme J, Opdenakker G. Functional comparison of two human monocyte chemotactic protein-2 isoforms, role of the amino-terminal pyroglutamic acid and processing by CD26/dipeptidyl peptidase IV. Biochemistry 1998;37:12672-80.

[68] Kumar A, Bachhawat AK. Pyroglutamic acid: throwing light on a lightly studied metabolite. Current Science 2012; 102: 288-97.

[69] Gerhard DS, Wagner L, Feingold EA, Shenmen CM, Grouse LH, Schuler G, Klein SL, Old S, Rasooly R, Good P, Guyer M, Peck AM, Derge JG, Lipman D, Collins FS, Jang W, Sherry S, Feolo M, Malek J. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). Genome Res 2004;14:2121-7.

[70] Scherer SE, Muzny DM, Buhay CJ, Chen R, Cree A, Ding Y, Dugan-Rocha S, Gill R, Gunaratne P, Harris RA, Hawes AC, Hernandez J, Hodgson AV, Hume J, Jackson A, Khan ZM, Kovar-Smith C, Lewis LR. The finished DNA sequence of human chromosome 12. Nature 2006;440:346-51.

[71] Ayad M, Van Wuyckhuyse BC, Minaguchi K, Raubertas RF, Bedi GS, Billings RJ, Bowen WH, Tabak LA. The association of basic proline-rich peptides from human parotid gland secretions with caries experience. J Dent Res 2000;79:976-82.

[72] Vitorino R, Alves R, Barros A, Caseiro A, Ferreira R, Lobo MC, Bastos A, Duarte J, Carvalho D, Santos LL, Amado FL. Finding new posttranslational modifications in salivary proline-rich proteins. Proteomics 2010;10:3732-42.

[73] Morzel M, Jeannin A, Lucchi G, Truntzer C, Pecqueur D, Nicklaus S, Chambon C, Ducoroy P. Human infant saliva peptidome is modified with age and diet transition. J Proteomics 2012;75:3665-73.

[74] Vitorino R, Calheiros-Lobo MJ, Williams J, Ferrer-Correia AJ, Tomer KB, Duarte JA, Domingues PM, Amado FM. Peptidomic analysis of human acquired enamel pellicle. Biomed Chromatogr 2007;21:1107-17.

[75] Siqueira WL, Oppenheim FG. Small molecular weight proteins/peptides present in the in vivo formed human acquired enamel pellicle. Arch Oral Biol 2009;54:437-44.

[76] Cabras T, Melis M, Castagnola M, Padiglia A, Tepper BJ, 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.

[77] Melis M, Aragoni MC, Arca M, Cabras T, Caltagirone C, Castagnola M, Crnjar R, Messana I, Tepper BJ, Tomassini Barbarossa I. Marked increase in PROP taste responsiveness following oral supplementation with selected salivary proteins or their related free amino acids. PLoS One 2013;8:e59810.

[78] Lu Y, Bennick A. Interaction of tannin with human salivary proline-rich proteins. Arch Oral Biol 1998;43:717-28.

[79] Cai K, Hagerman AE, Minto RE, Bennick A. Decreased polyphenol transport across cultured intestinal cells by a salivary proline-rich protein. Biochem Pharmacol 2006;71:1570-80.

[80] Charlton AJ, Baxter NJ, Lilley TH, Haslam E, McDonald CJ, Williamson MP. Tannin interactions with a full-length human salivary proline-rich protein display a stronger affinity than with single proline-rich repeats. FEBS Lett 1996;382:289-92.

[81] Canon F, Paté F, Cheynier V, Sarni-Manchado P, Giuliani A, Pérez J, Durand D, Li J, Cabane B. Aggregation of the salivary proline-rich protein IB5 in the presence of the tannin EgCG. Langmuir 2013;29:1926-37.

[82] O'Sullivan JM, Cannon RD, Sullivan PA, Jenkinson HF. Identification of salivary basic proline-rich proteins as receptors for Candida albicans adhesion. Microbiology 1997;143:341-8.

[83] Marshall RD. The nature and metabolism of the carbohydrate peptide linkages of glycoproteins. Biochem Soc Symp 1974;40:17-26.

[84] Gavel Y, von Heijne G. Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. Protein Eng 1990;3:433-42.

[85] Petrescu AJ, Butters TD, Reinkensmeier G, Petrescu S, Platt FM, Dwek RA, Wormald MR. The solution NMR structure of glycosylated N-glycans involved in the early stages of glycoprotein biosynthesis and folding. EMBO J 1997;16:4302-10.

[86] Varki A. Essentials of Glycobiology. New York, Cold Spring Harbor Laboratory Press; 2nd edition. 2009.

[87] Mayampurath AM, Wu Y, Segu ZM, Mechref Y, Tang H. Improving confidence in detection and characterization of protein N-glycosylation sites and microheterogeneity. Rapid Commun Mass Spectrom 2011; 25:2007-19.

[88] Azen EA, Hurley CK, Denniston C. Genetic polymorphism of the major parotid salivary glycoprotein (Gl) with linkage to the genes for Pr, Db, and Pa. Biochem Genet 1979;17:257-79.

[89] Azen EA, Minaguchi K, Latreille P, Kim HS. Alleles at the PRB3 locus coding for a disulfide-bonded human salivary proline-rich glycoprotein (Gl 8) and a null in an Ashkenazi Jew. Am J Hum Genet 1990;47:686-97.

[90] Minaguchi K, Takaesu Y, Tsutsumi T, Suzuki K. Studies of genetic markers in human saliva. (VII). Frequencies of the major parotid salivary glycoprotein (GI) system in a Japanese population. Bull Tokyo Dent Coll 1981;22:1-6.

[91] Carpenter GH, Proctor GB. O-linked glycosylation occurs on basic parotid salivary proline-rich proteins. Oral Microbiol Immunol 1999;14:309-15.

[92] Gillece-Castro BL, Prakobphol A, Burlingame AL, Leffler H, Fisher SJ. Structure and bacterial receptor activity of a human salivary proline-rich glycoprotein. J Biol Chem 1991;266:17358-68.

[93] Levine MJ, Weill JC, Ellison SA. The isolation and analysis of a glycoprotein from parotid saliva. Biochim Biophys Acta 1969;188:165-7.

[94] Shimomura H, Kanai Y, Sanada K. Amino acid sequences of glycopeptides obtained from basic proline-rich glycoprotein of human parotid saliva. J Biochem 1983;93:857-63.

[95] Mehansho H, Carlson DM. Induction of protein and glycoprotein synthesis in rat submandibular glands by isoproterenol. J Biol Chem 1983;258:6616-20.

[96] Azen EA, Prakobphol A, Fisher SJ. PRB3 null mutations result in absence of the proline-rich glycoprotein Gl and abolish *Fusobacterium nucleatum* interactions with saliva in vitro. Infect Immun 1993;6:4434-9.

[97] Bergey EJ, Levine MJ, Reddy MS, Bradway SD, Al-Hashimi I. Use of the photoaffinity cross-linking agent N-hydroxysuccinimidyl-4-azidosalicylic acid to characterize salivary-glycoprotein-bacterial interactions. Biochem J 1986;234:43-8.

[98] Ruhl S, Sandberg AL, Cisar JO. Salivary receptors for the proline-rich protein-binding and lectin-like adhesins of oral actinomyces and streptococci. J Dent Res 2004;83:505-10.

[99] Nagata K, Nakao M, Shibata S, Shizukuishi S, Nakamura R, Tsunemitsu A. Purification and characterization of galactosephilic component present on the cell surfaces of *Streptococcus sanguis* ATCC 10557. J Periodontol 1983;54:163-72.

[100] Murray PA, Prakobphol A, Lee T, Hoover CI, Fisher SJ. Adherence of oral streptococci to salivary glycoproteins. Infect Immun 1992;60:31-8.

[101] Perez-Gregorio MR, Mateus N, de Freitas V. Rapid screening and identification of new soluble tannin-salivary protein aggregates in saliva by mass spectrometry (MALDI-TOF-TOF and FIA-ESI-MS). Langmuir 2014;30:8528-37.

[102] Zakhary GM, Clark RM, Bidichandani SI, Owen WL, Slayton RL, Levine M. Acidic proline-rich protein Db and caries in young children. J Dent Res 2007;86:1176-80.

[103] Castagnola M, Cabras T, Vitali A, Sanna MT, Messana I. Biotechnological implications of the salivary proteome. Trends Biotechnol 2011;29:409-18.

CAPTIONS TO FIGURES

Figure 1. Schematic representation of the human PRPs gene cluster. The six genes of PRPs (*PRB2*, ID: 653247; *PRB1*, ID: 5542; *PRB4*, ID: 5545; *PRB3*, ID: 5544; *PRH1*, ID: 5554; *PRH2*, ID: 5555) are contained within a ~ 0.5 Mb segment of the chromosome 12p13.2. The red box reports the main alleles found in Caucasian population.

Figure 2. Typical TIC HPLC-ESI-MS profile of the acidic soluble fraction of adult human whole saliva. The main families of secretory salivary peptides and proteins, i.e. aPRPs, bPRPs and gPRPs, histatins, statherin and PB peptides, S-type cystatins, due to the structural similarity of their members, elute in well-defined chromatographic clusters (only histatins are partly overlapped to PRPs).

Figure 3. Schematic representation of the most common human salivary aPRP protein species detectable in adult saliva of western population (Modified from ref. [26]). PRP-1, PRP-2, PIF-s, and Db-s are partially cleaved (bold arrows) at Arg_{106} (Arg_{127} in Db isoform) generating the four truncated protein species reported on the bottom left of the figure and the P-C peptide. The Pa isoform, carrying the substitution Arg_{103}→Cys is not cleaved, and it is usually present in human saliva as a 2-mer. Some entire or truncated protein species can partially undergo carboxypeptidase removal of C-terminal residues. <Q: N-terminal pyroglutamic acid; <u>S</u>: pSer (Ser_8 and Ser_{22}); S*: minor site of phosphorylation (Ser_{17}); S22**: $pSer_{22}$→Phe variation in PRP-1 (and PRP-3) Roma-Boston variant. For other information, see Section 2.

Figure 4. HPLC-ESI-MS analysis of IB-8a glycoprotein species. Amino acid sequence and deconvoluted mass spectrum showing the average mass values of the glycoprotein species carrying (left to right) the monofucosylated monoantennary form, and the monofucosylated, bifucosylated,

trifucosylated, tetrafucosylated and pentafucosylated biantennary glycans linked to Asn_{98} (bold amino acid in the sequence).

Table 1. List of the theoretical and experimental monoisotopic $[M+H]^{1+}$ values of the protein species originated from *PRH1* and *PRH2* genes, and detected in the acid soluble fraction of whole saliva of the three subjects carrying the PRP-1 (PRP-3) RB variant. Modified from ref. [25].

			Subject Boston I	Subject Boston II	Subject Roma
Genes	Protein species	$[M+H]^{1+}$ Theor.	$[M+H]^{1+}$ Exp.	$[M+H]^{1+}$ Exp.	$[M+H]^{1+}$ Exp.
PRH2	PRP-1 mono P	*15426.27*	-	15426.47	15426.22
	PRP-1 di P	*15506.24*	-	15506.24	15506.28
	PRP-1 tri P	*15586.21*	-	15586.23	15586.32
	PRP-2 mono P	*15427.26*	15427.27	-	-
	PRP-2 di P	*15507.22*	15507.16	-	-
	PRP-2 tri P	*15587.19*	15587.20	-	-
	PRP-3 mono P	*11076.11*	-	11075.98	11076.09
	PRP-3 di P	*11156.08*	-	11156.12	11156.10
	PRP-3 tri P	*11236.04*	-	11236.06	11236.07
	PRP-4 mono P	*11077.09*	11077.09	-	-
	PRP-4 di P	*11157.06*	11157.09	-	-
	PRP-4 tri P	*11237.03*	11237.12	-	-
	PRP-1 RB non P	*15406.34*	15406.24	15406.39	15406.39
	PRP-1 RB mono P (Ser_8)	*15486.31*	15486.27	15486.31	15486.33
	PRP-1 RB di P	*15566.27*	-	-	-
	PRP-3 RB di P	*11216.12*	-	-	-
PRH1	Pif-s mono P	*15426.27*	15426.33	15426.47	15426.22
	Pif-s di P	*15506.24*	15506.31	15506.24	15506.28
	Pif-s tri P	*15586.21*	15586.23	15586.23	15586.32
	Db-s mono P	*17543.31*	-	-	-
	Db-s di P	*17623.28*	-	17623.30	-
	Db-s tri P	*17703.25*		17703.32	
	Pa 2-mer tri P	*30824.32*	30825.39	-	-
	Pa 2-mer tetra P	*30904.28*	30904.36		
	Pif-f mono P	*11076.11*	11076.14	11075.98	11076.09
	Pif-f di P	*11156.08*	11156.07	11156.12	11156.10
	Pif-f tri P	*11236.04*	11236.00	11236.06	11236.07
	Db-f mono P	*13193.17*	-	13193.17	-
	Db-f di P	*13273.12*	-	13273.16	-
	Db-f tri P	*13353.08*	-	-	-
PRH2 and *PRH1*	P-C peptide	*4369.19*	4369.19	4369.19	4369.19

Table 2. bPRPs species from *PRB1* alleles

Gene	Proprotein		Protein species			
PRB1 Alleles	# Convertase cleavages	UniprotKB codes	Protein Name	# Residues	Presecretory PTMs	Postsecretory PTMs*
PRB1S	2	*P04280*	II-2 (Pe/DEAEII-2)[a] P-E (IB-9)[a] IB-6 (P-I)[a]	75 61 118	N-term. pGlu, $pSer_8$ P-E $desArg_{62}$	II-2 $desArg_{75;}$ II-2 $des_{72\text{-}75}$ P-E $desArg_{61}$
PRB1M	1	*Q86YA1*	II-2 (Pe/DEAEII-2)[a] Ps-1	75 240	N-term. pGlu, $pSer_8$ *O*-glycosylation	
PRB1L	1	*P04280*	II-2 (Pe/DEAEII-2)[a] Ps-2	75 300	N-term. pGlu, $pSer_8$ *O*-glycosylation	
PRB1L/ cP4 (alternative transcript)	1	*P04280*	II-2 (Pe/DEAEII-2)[a] P-Ko	75 107	N-term. pGlu, $pSer_8$	
PRB1S, M, or L	2 or 1		II-2 variant lacking Pro_{39}	74		

[a] Alternative names; *small fragments are generated by endoproteinase cleavages of the different bPRPs; pSer: phosphorylated Ser; pGlu: pyroglutamic acid

Table 3. bPRPs species from *PRB2* alleles

Gene	Proprotein		Protein species			
PRB2 Alleles	# Convertase cleavages	UniprotKB codes	Protein Name	# Residues	Presecretory PTMs	Postsecretory PTMs*
PRB2L *CON1*$^{-}$	4	*P02812*	IB-1 P-J P-F (IB-8c)[a] P-H (IB-4)[a] IB-8a (P_{100}) (Con1^{-})[a]	96 61 61 56 123	N-term. pGlu, $pSer_8$ C-term. removal of Arg_{62} C-term. removal of Arg_{62} C-term. removal of Arg_{124}	IB-1 $desArg_{96}$, IB-1 des_{93-96} P-J des_{60-61} (IB-7)
PRB2L *CON1*$^{+}$	4	*P02812*	IB-1 P-J P-F (IB-8c)[a] P-H (IB-4)[a] IB-8a (S_{100}) (Con1^{+})[a]	96 61 61 56 123	N-term. pGlu, $pSer_8$ C-term. removal of Arg_{62} C-term. removal of Arg_{62} C-term. removal of Arg_{124}, 1 *N*-glycosite	IB-1 $desArg_{96}$, IB-1 des_{93-96} P-J des_{60-61} (IB-7)
PRB1M *CON2*	4		II-2 (Pe/DEAEII-2)[a] P-E (IB-9)[a] P-F (IB-8c)[a] P-H (IB-4)[a] Con2	75 61 61 56 60	N-term. pGlu, $pSer_8$ P-E $desArg_{62}$ C-term. removal of Arg_{62} 1 *N*-glycosite	II-2 $desArg_{75}$; II-2 des_{72-75} P-E $desArg_{61}$

[a] Alternative names; *small fragments are generated by endoproteinase cleavages of the different bPRPs;
pSer: phosphorylated Ser; pGlu: pyroglutamic acid

Table 4. bPRPs species from *PRB4* alleles.

Gene	Proprotein		Protein species			
PRB4 Alleles	# Convertase cleavage	UniprotKB codes	Protein Name	# Residues	Presecretory PTMs	Postsecretory PTMs
PRB4S	1	*P10163*	PGA P-D (Ala_{32}) (IB-5)[a] *or* P-D (Pro_{32})	161 70 70	5 potential *N*-glycosytes	small fragments generated by endoproteinase cleavages
PRB4M	1	*P10163*	II-1 P-D (Ala_{32}) (IB-5)[a] *or* P-D (Pro_{32})	182 70 70	7 potential *N*-glycosytes	
PRB4L	1	*P10163*	CD-IIg P-D (Ala_{32}) (IB-5)[a] *or* P-D (Pro_{32})	182 70 70	8 potential *N*-glycosytes	

[a] Alternative names

Table 5. Phenotypic frequencies (percentages) of the glycosylated and nonglycosylated IB-8a protein species determined in 71 subjects.

Con 1$^+$ glycosylated	Con 1$^+$ glycosylated and Con 1$^+$ nonglycosylated	Con 1$^+$ nonglycosylated	Con 1$^-$	Con 1$^+$ glycosylated and Con 1$^-$
3 (4%)	19 (27%)	3 (4%)	26 (37 %)	6 (8%)

14 subjects (20%) did not show IB-8a protein species.

Table 6. bPRPs species from *PRB3* alleles

Gene	Protein species				
PRB3 Alleles	UniprotKB codes	Protein Name	# Residues	Presecretory PTMs	Postsecretory PTMs
PRB3S	*Q04118*	Gl3 (PRG)[a]	230	5 potential *N*-glycosytes	
PRB3M	*Q04118*	Gl2 (PRG)[a]	293	8 potential *N*-glycosytes	
PRB3L		Gl1 (PRG)[a]	313	9 potential *N*-glycosytes	small fragments generated by endoproteinase cleavages
PRB3VL		Gl4 (PRG)[a]	----		
PRB3S^{Cys}		Gl8 (Cys_{15})	230	5 potential *N*-glycosytes	

[a] Alternative names

Figure 1

Click here to download Figure: Fig 1.pptx

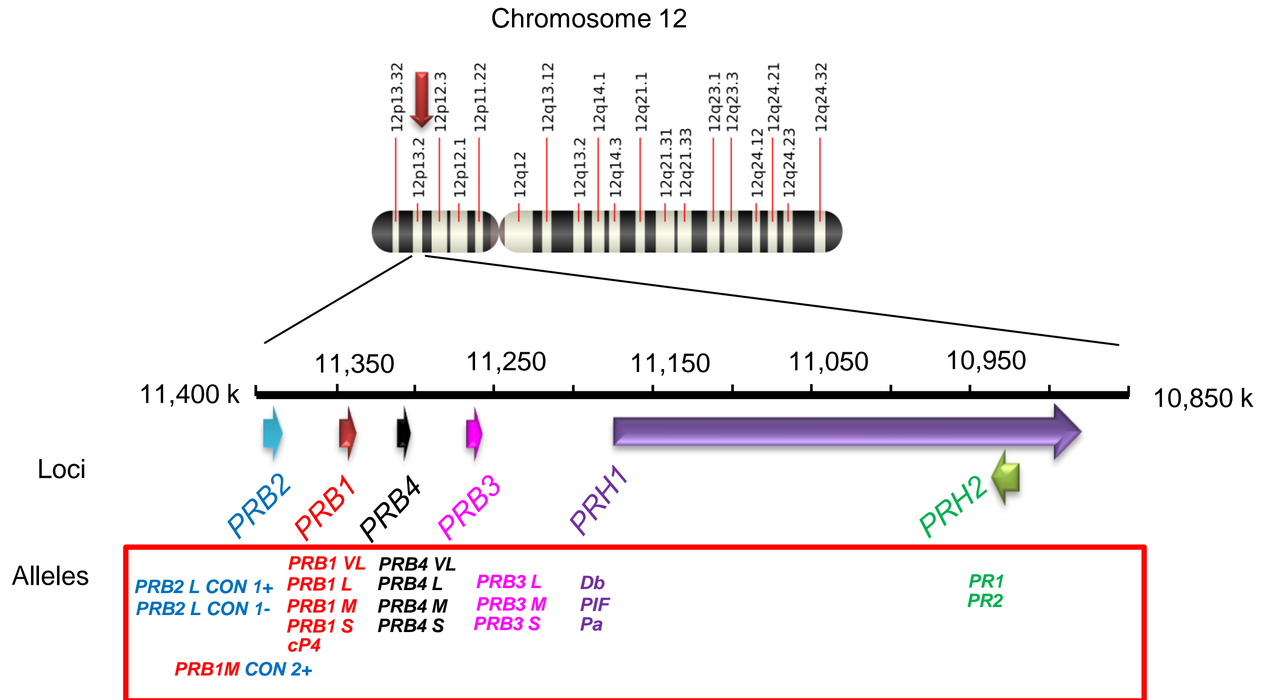


Figure 2

Click here to download high resolution image

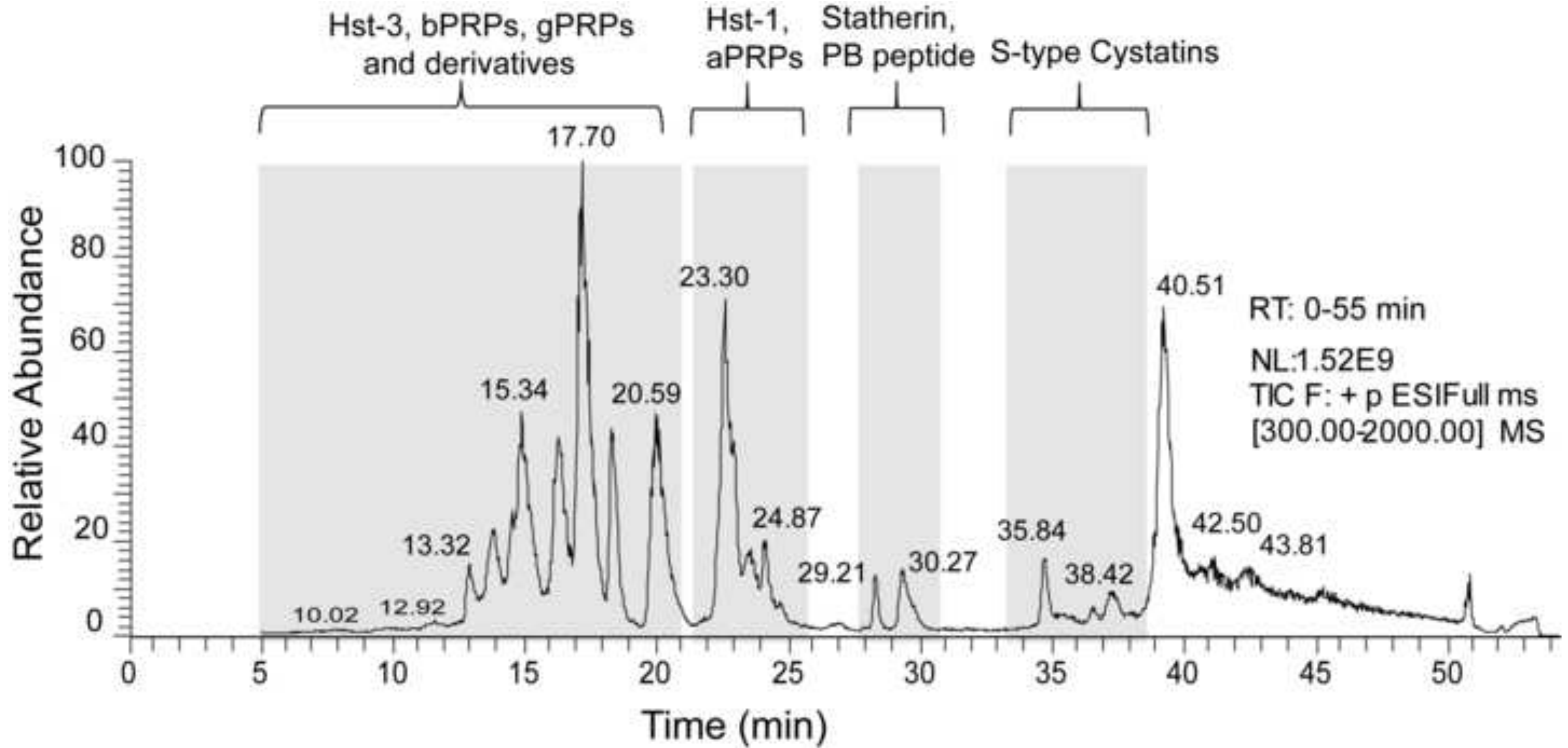


Figure 3

Click here to download high resolution image

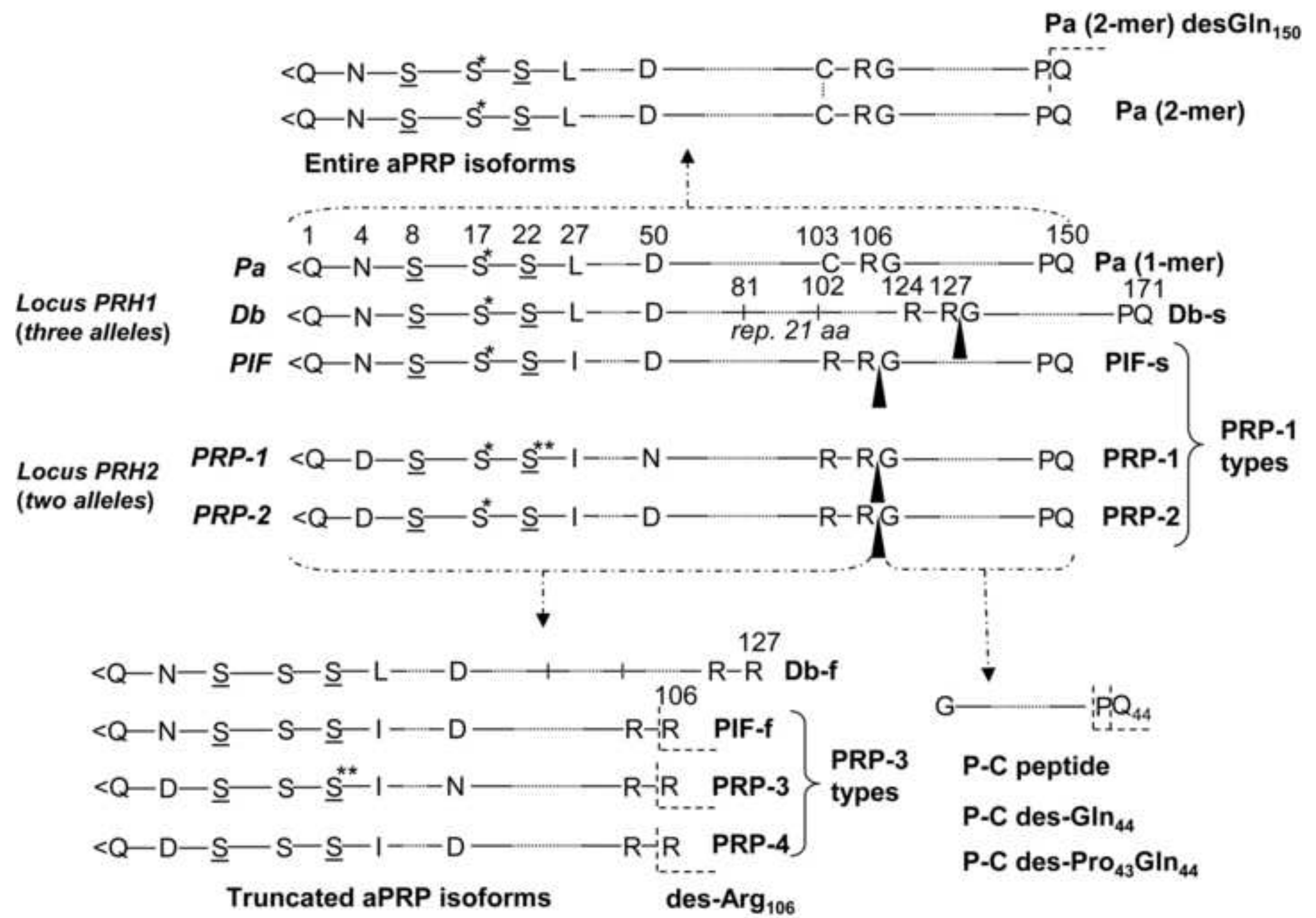
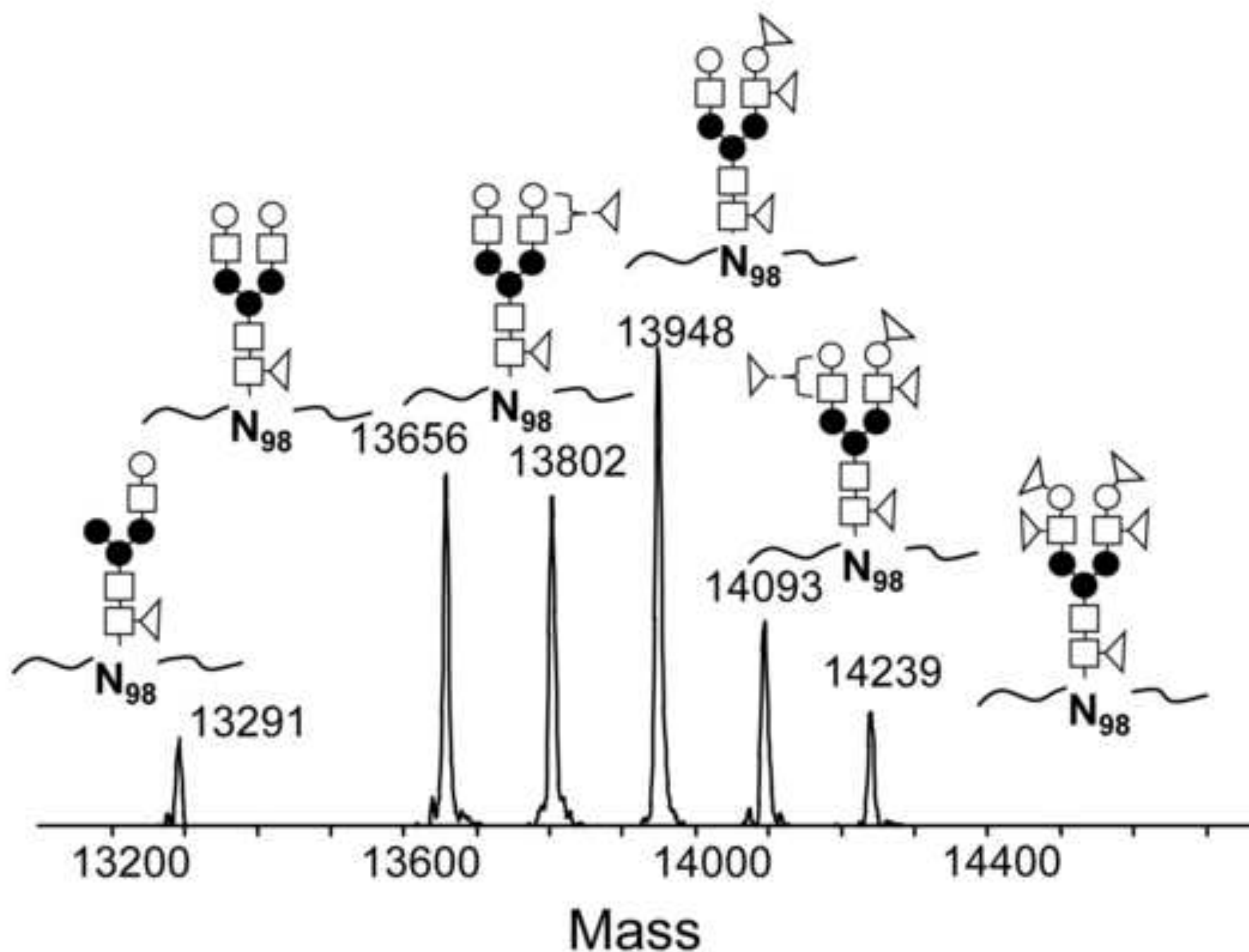


Figure 4

Click here to download high resolution image

SPPGKPQGPP PQGGNQPQGP PPPPGKPQGP PPQGGNKPQG PPPPGKPQGP
10 20 30 40 50
PPQGDNKSQS ARSPPGKPQG PPPQGGNQPQ GPPPPPGKPQ GPPPQGG**N**KS
60 70 80 90 100
QGPPPPGKPQ GPPPQGGSKS RSS
110 120



*Conflict of Interest

Click here to download Conflict of Interest: Author declaration Messana.pdf